

UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOQUÍMICA E IMUNOLOGIA

HEVELINE OLIVEIRA MORAIS ARRUDA

**PRIORIZAÇÃO DE PROTEÍNAS QUINASES DE *TRYPANOSOMA CRUZI* COMO
POTENCIAIS ALVOS TERAPÊUTICOS PARA O DESENVOLVIMENTO DE
INIBIDORES COVALENTES**

Belo Horizonte

2022

Heveline Oliveira Morais Arruda

**PRIORIZAÇÃO DE PROTEÍNAS QUINASES DE TRYPANOSOMA CRUZI COMO
POTENCIAIS ALVOS TERAPÊUTICOS PARA O DESENVOLVIMENTO DE
INIBIDORES COVALENTES**

Dissertação apresentada ao Programa de Pós-Graduação em Bioquímica e Imunologia do Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais como requisito parcial para a obtenção do título de Mestre em Bioquímica e Imunologia.

Orientadora: Profa. Dra. Rafaela Salgado Ferreira

Belo Horizonte

2022

043

Arruda, Heveline Oliveira Morais.

Priorização de proteínas quinases de *Trypanosoma cruzi* como potenciais alvos terapêuticos para o desenvolvimento de inibidores covalentes [manuscrito] / Heveline Oliveira Morais Arruda. – 2022.

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

Orientadora: Profa. Dra. Rafaela Salgado Ferreira.

Dissertação (mestrado) – Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Departamento de Bioquímica e Imunologia.

1. Bioquímica e imunologia. 2. *Trypanosoma cruzi*. 3. Fosfotransferases. 4. Inibidores de Proteínas Quinases. I. Ferreira, Rafaela Salgado. II. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. III. Título.

CDU: 577.1



Universidade Federal de Minas Gerais
Curso de Pós-Graduação em Bioquímica e Imunologia ICB/UFMG
Av. Antônio Carlos, 6627 – Pampulha
31270-901 – Belo Horizonte – MG
e-mail: pg-biq@icb.ufmg.br (31)3409-2615



ATA DA DEFESA DA DISSERTAÇÃO DE MESTRADO DE HEVELINE OLIVEIRA MORAIS ARRUDA. Aos vinte e três dias do mês de junho de 2022 às 13:00 horas, reuniu-se forma “on line” utilizando a plataforma “Zoom”, no Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais, a Comissão Examinadora da dissertação de Mestrado, indicada *ad referendum* do Colegiado do Curso, para julgar, em exame final, o trabalho intitulado "Priorização de proteínas quinases de Trypanosoma cruzi como potenciais alvos terapêuticos para o desenvolvimento de inibidores covalentes", requisito final para a obtenção do grau de Mestre em Bioquímica e Imunologia, área de concentração: Bioquímica. Abrindo a sessão, o Presidente da Comissão, Profa. Rafaela Salgado Ferreira, da Universidade Federal de Minas Gerais, após dar a conhecer aos presentes o teor das Normas Regulamentares do Trabalho Final, passou a palavra à candidata para apresentação de seu trabalho. Seguiu-se a arguição pelos examinadores, com a respectiva defesa da candidata. Logo após a Comissão se reuniu, sem a presença da candidata e do público, para julgamento e expedição do resultado final. Foram atribuídas as seguintes indicações: Dr. Gustavo Costa Bressan (Universidade Federal de Viçosa), aprovada; Dr. Lucas Bleicher (Universidade Federal de Minas Gerais), aprovada; Dra. Rafaela Salgado Ferreira - Orientadora (Universidade Federal de Minas Gerais), aprovada. Pelas indicações a candidata foi considerada:

- APROVADA
 REPROVADA

O resultado final foi comunicado publicamente à candidata pelo Presidente da Comissão. Nada mais havendo a tratar, o Presidente da Comissão encerrou a reunião e lavrou a presente Ata que será assinada por todos os membros participantes da Comissão Examinadora. Belo Horizonte, 23 de junho de 2022.

Dr. Gustavo Costa Bressan (Universidade Federal de Viçosa)

Dr. Lucas Bleicher (UFMG)

Dra. Rafaela Salgado Ferreira - Orientadora (UFMG)

AGRADECIMENTOS

A experiência deste mestrado foi atravessada pela pandemia de Covid 2019 e toda sorte de acontecimentos que vieram junto. Isto moldou significativamente a vivência e o desenvolvimento deste trabalho. Deste contexto ficou a percepção ainda mais nítida da importância do convívio, da troca, no nosso crescimento e aprendizado e, neste sentido, devo muito a muitas pessoas.

Agradeço primeiramente às pessoas diretamente envolvidas neste trabalho, à Rafaela Salgado, minha orientadora, e aos colegas de laboratório, principalmente ao Rafael Rocha, à Viviane Correa e ao Augusto Campos que me propiciaram vivenciar uma nova área do conhecimento e amadurecer cientificamente. Foi um período de grande aprendizado. Agradeço ao Departamento de Bioquímica e Imunologia e à CAPES pela oportunidade e bolsa concedida.

Agradeço também ao Pedro Martins, Wesley Leão, Vinícius Carius e João Locke que gentilmente me auxiliaram em momentos chaves do projeto, trazendo um respiro.

Apesar do convívio limitado tive a oportunidade de conhecer pessoas queridas como o Filipe, a Josiane e a Gôndia que me ajudaram a passar por essa experiência, e espero poder tê-los por perto por muito mais tempo.

Gostaria de agradecer também a algumas pessoas que caminham comigo há muitos anos e que hoje, lidando com a desafiadora depressão, me fazem estar aqui, firme, e seguindo em frente. Agradeço a vocês meus queridos primos Milene e Guilherme, à Raquel, amiga dos tempos de preparatório CEFET/COLTEC, da risada mais famosa do CEFET, à minha querida tríade da química, Marina, Mariana e Cíntia, à Elsie, uma irmã que a vida gentilmente me deu nessas idas e vindas da área de educação, aos queridos colegas de trabalho que hoje são grandes amigos, Rodrigo e Graci, à Sabrina, grande amiga que o tempo de voluntariado no movimento escoteiro me deu, à Elaine, amizade que veio de uma outra área de formação e que me acompanha até os dias de hoje e à Lúcia, professora de inglês que hoje tenho como amiga. Obrigada a vocês pela gentileza, amizade, cuidado e amparo. Sem vocês por perto certamente tudo seria muito mais difícil.

Agradeço a Olorum, Oyá e Oxalá, obrigada, obrigada e obrigada!

RESUMO

A doença de Chagas, uma doença tropical negligenciada causada pelo organismo *Trypanosoma cruzi* (*T. cruzi*), continua gerando grande impacto na saúde pública e prejuízos financeiros, mesmo depois de 100 anos de sua descoberta. O tratamento disponível hoje apresenta baixa eficácia e severos efeitos colaterais. Nos esforços de combater esta doença alguns alvos terapêuticos têm sido explorados. O *T. cruzi* possui 190 proteínas quinases (PKs) no seu genoma que atuam em diversos processos fisiológicos, constituindo uma fonte potencial de alvos terapêuticos. O conhecimento acerca das estruturas das quinases, em conjunto com os estudos de mapeamento de resíduos reativos de cisteínas, permitem o planejamento racional de inibidores covalentes para estas proteínas como uma estratégia que equilibra seletividade, eficácia e segurança. Neste trabalho empregaram-se técnicas de bioinformática para priorizar quinases de *T. cruzi* potencialmente vitais para o organismo e que sejam promissoras alvos para uma inibição covalente. Para tanto, partiu-se do levantamento das quinases das cepas Dm28c 2014 (171 PKs) e CL Brener Esmeraldo-like (175 PKs) em bancos de dados, seguida da classificação das quinases nos grupos e famílias típicos dessa classe de proteínas. Na sequência, foi feito o mapeamento das cisteínas presentes no sítio ativo das quinases das duas. Foram mapeadas 25 posições estratégicas com características estereoquímicas e eletrostáticas favoráveis para uma inibição covalente. Além deste mapeamento, estudos que validaram a importância biológica de quinases de *Trypanosoma brucei* (*T. brucei*) *in vivo*, na proliferação, no ciclo celular e na sobrevivência do parasito, embasaram a etapa de priorização das PKs. Nesta etapa, focada na cepa CL Brener, foram priorizadas 3 quinases de *T. cruzi*: AUK1, CLK1 e CRK9. A AUK1 é reportada como tendo um papel na divisão celular e possivelmente na diferenciação do parasito. A CLK1 e a CRK9, por extrapolação dos estudos em *T. brucei*, supostamente estão envolvidas na divisão celular. Além disso, a CRK9 também estaria envolvida na expressão gênica. Por fim, para possibilitar o início dos estudos de inibição destas quinases, foram construídos os modelos tridimensionais a partir do método de modelagem comparativa, com posterior validação estereoquímica e energética dos modelos gerados. As informações e discussões construídas ao longo deste estudo sobre as proteínas quinases de *T. cruzi* poderão ser utilizadas para embasar novas propostas de alvos terapêuticos e trazer novas perspectivas para a gestão da doença de Chagas.

Palavras-chave: *Trypanosoma cruzi*. Quinases. Inibidores covalentes.

ABSTRACT

Chagas disease, a neglected tropical disease caused by the organism *Trypanosoma cruzi* (*T. cruzi*), continues having a major impact on public health and financial losses, even more than 100 years after its discovery. The treatment available today has low efficacy and causes severe side effects. As an effort to combat this disease, some therapeutic targets have been explored. *T. cruzi* has 190 protein kinases (PKs) in its genome that act in several physiological processes, thus constituting a potential source of therapeutic targets. The knowledge about kinase structures and the studies mapping reactive cysteine residues allow the rational design of covalent inhibitors for these proteins as a strategy that balances selectivity, efficacy, and safety. In this work, we used bioinformatics techniques to prioritize *T. cruzi* kinases that are potentially vital to the organism and promising targets for covalent inhibition. For that, we started with a database survey of the kinases of the Dm28c 2014 (171 PKs) and CL Brener Esmeraldo-like (175 PKs) strains, followed by the classification of the kinases in their typical groups and families. Then, we mapped the cysteines present in the active site of the kinases of the two strains. We mapped twenty-five strategic positions with stereochemical and electrostatic characteristics favorable for covalent inhibition. In addition, studies that validated the biological importance of *T. brucei*'s kinases *in vivo*, in proliferation, in the cell cycle, and in the survival of the parasite, supported the PKs prioritization. In this step, focused on the CL Brener strain, we prioritized three *T. cruzi* kinases: AUK1, CLK1 and CRK9. AUK1 is reported to have a role in cell division and, possibly, in parasite differentiation. We believe that CLK1 and CRK9 participate in cell division based on studies in *T. brucei*. In addition, CRK9 would also be involved in gene expression. Finally, to enable studies on the inhibition of these kinases, we built three-dimensional models employing comparative modeling, with subsequent stereochemical and energetic validation of the models generated. The information and discussions built throughout this study on *T. cruzi* protein kinases can support new proposals for therapeutic targets and bring new perspectives to the management of Chagas disease.

Keywords: *Trypanosoma cruzi*. Kinases. Covalent inhibitors.

LISTA DE FIGURAS

- Figura 1** - Rotas migratórias da América Latina para países não endêmicos e estimativa do número de indivíduos infectados nos Estados Unidos, Canadá, Europa e países da região ocidental do pacífico, Japão e Austrália. 19
- Figura 2** - Ciclo do parasito *Trypanosoma cruzi* no inseto triatomíneo e no hospedeiro homem21
- Figura 3** - Estrutura química dos fármacos disponíveis para o controle e eliminação do parasito *Trypanosoma cruzi*, nifurtimox e benzonidazol.22
- Figura 4** - Esquema da reação de fosforilação de uma proteína quinase.....27
- Figura 5** - Identificação dos subdomínio do domínio catalítico das proteínas quinases. Os 12 subdomínios estão apresentados em diferentes cores juntamente com as estruturas secundárias (β para fitas β e α para α -hélices) enumeradas. Também foram indicadas na figura a localização dos dois lóbulos que correspondem aos grandes agrupamentos de estruturas secundárias do domínio catalítico, e a presença do ATP em bastões na fenda catalítica entre os dois lóbulos.28
- Figura 6** – Trechos dos principais subdomínios do sítio de ligação do ATP nas quinases e resíduos chave para a catálise. Além dos subdomínios, estão destacados em bastões os três resíduos catalíticos das quinases, Lys do subdomínio II, Asp do subdomínio VIB e Asp do subdomínio VII, assim como o Glu conservado da α -hélice C e o resíduo fosforilado na alça de ativação que possuem papel essencial na estruturação do sítio para a catálise. As esferas em roxo representam os dois íons magnésio presentes no sítio ativo da proteína quinase PKA de *Mus musculus* (Código PDB 1ATP)..... 31
- Figura 7** – Principais interações no sítio ativo das quinases. Na figura estão representadas as interações envolvendo alguns resíduos do sítio ativo das quinases, como os resíduos catalíticos K72, D166 e D184, além dos resíduos E91, N171 e K168; os substratos, o ATP (vermelho) e o substrato proteico (laranja); e os íons magnésio, Mg_1 e Mg_2 . A numeração dos resíduos corresponde à da quinase PKA de código PDB 1ATP.32
- Figura 8** - Eixos hidrofóbicos conservados nas proteínas quinases, *C-spine* e *R-spine*. Localização da estruturação espacial da *C-spine* (amarelo) e da *R-spine* (azul) e identificação dos respectivos resíduos que as compõem na quinase PKA (Código PDB:1ATP). (Parte da estrutura tridimensional foi omitida para facilitar a visualização da região analisada)..... 34
- Figura 9** - Esboço das principais estruturas e mudanças conformacionais envolvidas na ativação das quinases. Está representada a movimentação do lóbulo N, da α -hélice C e do motivo DFG de uma conformação *out* para uma conformação *in*, e da alça de ativação de uma conformação fechada (*closed*) para uma aberta (*open*). Também estão identificados alguns dos principais resíduos envolvidos na ativação das quinases, como a Lys catalítica da fita β_3 e o glutamato da α -hélice C, que estabelecem uma ponte salina nas quinases ativas, além do ATP, do resíduo da

alça de ativação cuja fosforilação contribui para o processo de ativação da quinase, e, por fim, da molécula ativadora (ciclina para as quinases da família CDK). 35

Figura 10 - Diagrama esquemático do mecanismo de reação das proteínas quinases. **A.** Esquema dos estados de transição associativo e dissociativo nas reações de transferência do grupo fosforil. **B.** Detalhamento esquemático da reação de fosforilação das proteínas quinases com destaque dos principais resíduos e interações envolvidos na reação no sítio ativo, segundo um mecanismo com formação de complexo ternário e estado de transição dissociativo..... 37

Figura 11 - Localização das regiões e cavidades do sítio ativo das quinases. Estão identificadas na estrutura da quinase EGFR de código PDB 1m14 as regiões e cavidades que integram o sítio ativo das quinases e que são sítios de ligação dos inibidores. Na figura estão destacadas a fenda frontal (azul), a área de portão (ciano), a fenda traseira (cinza), além do seguimento de ativação (preto). 39

Figura 12 - Esquema das etapas da inibição covalente e da reação de adição de Michael e estruturas dos principais grupos eletrofílicos dos inibidores covalentes que se ligam a cisteínas. **A.** Esquema da inibição covalente, envolvendo a formação do complexo não covalente na primeira etapa e a formação de ligação covalente na segunda etapa, resultando em um complexo covalente. A seta pontilhada, indicando a volta do complexo covalente para o complexo não covalente, ocorre apenas para inibidores covalentes reversíveis. **B.** Esquema da reação de adição de Michael entre o grupo eletrofílico de um inibidor e um resíduo nucleofílico de cisteína. Na primeira equação está representada a reação de adição de Michael referente a um inibidor covalente irreversível e, na segunda, referente a um inibidor covalente reversível. (EWG - *Electron-Withdrawing Groups*, grupos retiradores de densidade eletrônica). **C.** Estrutura dos principais grupos eletrofílicos dos inibidores covalentes que se ligam a cisteínas. 44

Figura 13 - Estrutura dos 7 inibidores covalentes de quinases aprovados pela FDA com destaque da região eletrofílica envolvida na reação de adição de Michael com o alvo. 47

Figura 14 - Mapeamento das cisteínas em posições estratégicas para uma inibição covalente em quinases. **A.** Mapeamento de cisteínas na região do sítio ativo de quinases e proximidades elaborado por Zhao e colaboradores com identificação de 17 posições promissoras para uma estratégia de inibição covalente. **B.** Mapeamento de cisteínas na região do sítio ativo de quinases e proximidades elaborado por Chaikwad e colaboradores de 18 posições promissoras para uma estratégia de inibição covalente. 48

Figura 15 - Alinhamento produzido no servidor ClustalΩ entre a sequência da quinase AUK1 de *T. cruzi* da cepa CL Brener NE1 e a sequência do molde (código PDB 2BFY) para execução da modelagem comparativa. 81

Figura 16 - Validação do modelo AUK1.26 refinado construído por modelagem comparativa para a proteína AUK1 de *T. cruzi* da cepa CL Brener NE1, utilizando o molde 2BFY. **A.** Diagrama de Ramachandran gerado no servidor MolProbity para o modelo AUK1.26 refinado. Contornos em azul escuro indicam regiões permitidas, enquanto contornos em azul claro indicam as regiões favoráveis. **B.** Gráfico gerado no servidor ProSA, destacando o valor de Z-score para o modelo AUK1.26 refinado em relação a estruturas determinadas experimentalmente com número de resíduos semelhantes. 83

Figura 17 - Estrutura tridimensional do modelo AUK1.26 refinado da quinase AUK1 de *T. cruzi* da cepa CL Brener NEI gerado por modelagem comparativa em sobreposição com o molde 2BFY, com destaque para os resíduos catalíticos e cisteínas mapeadas. **A.** Sobreposição do modelo AUK1.26 com o molde 2BFY, com a presença do inibidor H1N. **B.** Destaque para os 3 resíduos catalíticos da AUK1, K58, D152 e D170, e para as 2 cisteínas mapeadas C109 (Hinge H2) e C186 (Other O3)..... 84

Figura 18 - Destaque das diferenças no sítio de ligação do ATP entre a estrutura do modelo AUK1.26 refinado da quinase AUK1 de *T. cruzi* da cepa CL Brener NEI e a estrutura do molde 2BFY. **A.** Destaque em bastões dos resíduos diferentes entre o modelo AUK1.26 refinado e o molde 2BFY que se encontram a até 6 Å do inibidor H1N. A estrutura da molécula H1N e parte da estrutura tridimensional foram omitidas para melhor visualização da região do sítio. **B.** Superfície eletrostática da região analisada para o modelo (à esquerda) e o molde (à direita) gerada pelo *plugin* APBS do PyMOL^{106,107}. A coloração em vermelho indica regiões de potencial eletrostático negativo e em azul regiões de potencial positivo..... 85

Figura 19 - Alinhamento produzido no servidor ClustalΩ entre a sequência da quinase AUK1 de *T. cruzi* da cepa CL Brener NEI e a sequência do molde (código PDB 1OL5) para execução da modelagem comparativa. 86

Figura 20 - Validação do modelo AUK1.56 refinado construído por modelagem comparativa para a proteína AUK1 de *T. cruzi* da cepa CL Brener NEI, utilizando o molde PDB 1OL5. **A.** Diagrama de Ramachandran gerado no servidor MolProbity para o modelo AUK1.56 refinado. Contornos em azul escuro indicam regiões permitidas, enquanto contornos em azul claro indicam as regiões favoráveis. **B.** Gráfico gerado no servidor ProSA, destacando o valor de Z-score para o modelo AUK1.56 refinado em relação a estruturas determinadas experimentalmente, com número de resíduos semelhantes..... 88

Figura 21 - Estrutura tridimensional do modelo AUK1.56 refinado da quinase AUK1 de *T. cruzi* da cepa CL Brener NEI gerado por modelagem comparativa em sobreposição com o molde 1OL5 com destaque para os resíduos catalíticos e cisteínas mapeadas. **A.** Sobreposição do modelo AUK1.56 com o molde 1OL5, com a presença do inibidor ADP e de 2 íons magnésio (esferas em verde). **B.** Destaque para os 3 resíduos catalíticos da AUK1, K58, D152 e D170, e para as 2 cisteína mapeadas C109 (Hinge H2) e C186 (Other O3)..... 89

Figura 22 - Destaque das diferenças no sítio de ligação do ATP entre a estrutura do modelo AUK1.56 refinado da quinase AUK1 de *T. cruzi* da cepa CL Brener NEI e a estrutura do molde 1OL5. **A.** Destaque em bastões dos resíduos diferentes entre o modelo AUK1.56 refinado e o molde 1OL5 que se encontram a até 6 Å do ligante ADP. A estrutura da molécula ADP e parte da estrutura tridimensional foram omitidas para melhor visualização da região do sítio. **B.** Superfície eletrostática da região analisada para o modelo (à esquerda) e o molde (à direita), gerada pelo *plugin* APBS do PyMOL^{106,107}. A coloração em vermelho indica regiões de potencial eletrostático negativo e em azul regiões de potencial positivo. 90

Figura 23 - Análise comparativa entre os modelos gerados para a quinase AUK1 de *T. cruzi* da cepa CL Brener NEI, AUK1.26 (molde 2BFY), estrutura semiativa, e AUK1.56 (molde 1OL5), estrutura ativa. **A.** Destaque para a distância entre a lisina catalítica K58 e o Glutamato E77 que estabelecem uma ponte salina na quinase em conformação ativa. O lóbulo N das duas estruturas foram alinhados para melhor evidenciar a variação de distância (parte da estrutura tridimensional foi omitida para despoluir a região analisada). **B.** Ilustração do deslocamento do

lóbulo C entre as estruturas dos dois modelos a partir do alinhamento do lóbulo N. C. Volume do sítio do ATP calculado com a ferramenta DoGSite destacado em azul. À esquerda para a estrutura AUK1.26 e à direita para a estrutura AUK1.56. 92

Figura 24 - Alinhamento das sequências das quinases AUK1 correspondentes em outras cepas de *T. cruzi* com destaque para os resíduos que se encontram no sítio do ATP a até 6 Å de distância dos inibidores H1N (molde 2BFY) e ADP (Molde 1OL5). As sequências foram alinhadas considerando o intervalo do aminoácido 1 ao 180 da quinase AUK1 de CL Brener NEI..... 93

Figura 25 - Alinhamento produzido no servidor ClustalΩ entre a sequência da quinase CLK1 de *T. cruzi* da cepa CL Brener El e a sequência do molde (código PDB 6Q2A) para execução da modelagem comparativa. 97

Figura 26 - Validação do modelo CLK1.35 refinado construído por modelagem comparativa para a proteína CLK1 de *T. cruzi* da cepa CL Brener El, utilizando o molde 6Q2A. **A.** Diagrama de Ramachandran gerado no servidor MolProbity para o modelo CLK1.35 refinado. Contornos em azul escuro indicam regiões permitidas, enquanto contornos em azul claro indicam as regiões favoráveis. **B.** Gráfico gerado no servidor ProSA, destacando o valor de Z-score para o modelo CLK1.35 refinado em relação a estruturas determinadas experimentalmente, com número de resíduos semelhantes. 98

Figura 27 - Estrutura tridimensional do modelo CLK1.35 refinado da quinase CLK1 de *T. cruzi* da cepa CL Brener El gerado por modelagem comparativa em sobreposição com o molde 6Q2A com destaque para os resíduos catalíticos e cisteínas mapeadas. **A.** Sobreposição do modelo CLK1.35 com o molde 6Q2A, com a presença do inibidor 5XH. **B.** Destaque para os 3 resíduos catalíticos da CLK1, K175; D172 e D306, e para as 3 cisteína mapeadas C232 (Front F1); C305 (DFG D1) e C310 (Others O2). 99

Figura 28 - Destaque das diferenças no sítio de ligação do ATP entre a estrutura do modelo CLK1.35 refinado da quinase CLK1 de *T. cruzi* da cepa CL Brener El e a estrutura do molde 6Q2A. **A.** Rotâmeros *outliers* (K228 e E276) na estrutura do modelo CLK1.35 refinado na região do sítio de ligação do ATP em sobreposição com o molde. **B.** Destaque em bastões dos resíduos diferentes entre o modelo CLK1.35 refinado e o molde 6Q2A que se encontram a até 6 Å do inibidor 5XH. **C.** Superfície eletrostática da região analisada para o modelo (à esquerda) e o molde (à direita), gerada pelo *plugin* APBS do PyMOL^{106,107}. A coloração em vermelho indica regiões de potencial eletrostático negativo e em azul regiões de potencial positivo... 100

Figura 29 - Alinhamento das sequências das quinases CLK1 correspondentes em outras cepas de *T. cruzi* com destaque para os resíduos que se encontram no sítio do ATP a até 6 Å de distância do inibidor 5XH (molde 6Q2A). As sequências foram alinhadas considerando o intervalo do aminoácido 123 ao 340 da quinase CLK1 de CL Brener El. 101

Figura 30 - Alinhamento produzido no servidor ClustalΩ entre a sequência da quinase CRK9 de *T. cruzi* da cepa CL Brener El e as sequências dos candidatos a moldes para execução de modelagem comparativa. Neste alinhamento foi mostrado apenas o trecho do resíduo 239 ao 708 da CRK9, englobando o domínio de quinases..... 106

Figura 31 - Validação do modelo CRK9.57 refinado construído por modelagem comparativa para a proteína CRK9 de *T. cruzi* da cepa CL Brener El, utilizando os moldes 4IMY e 3MI9.

A. Diagrama de Ramachandran geral gerado no servidor MolProbity para o modelo CRK9.57 refinado. Contornos em azul escuro indicam regiões permitidas, enquanto contornos em azul claro indicam as regiões favoráveis. Os rotâmeros outliers são indicados por círculos vermelhos. **B.** Diagrama de Ramachandran para isoleucinas e valinas gerado no servidor MolProbity para o modelo CRK9.57 refinado. **C.** Gráfico gerado no servidor ProSA, destacando o valor de Z-score para o modelo CRK9.57 refinado em relação a estruturas determinadas experimentalmente, com número de resíduos semelhantes..... 109

Figura 32 - Estrutura tridimensional do modelo CRK9.57 refinado da quinase CRK9 de *T. cruzi* da cepa CL Brener El gerado por modelagem comparativa em sobreposição com os moldes 4IMY e 3MI9, com destaque para as inserções e para os resíduos catalíticos e cisteínas mapeadas. **A.** Estrutura tridimensional geral do modelo CRK9.57 refinado, **com inserções** dentro do domínio de quinases (destacadas em azul), sobreposta às estruturas dos moldes 4IMY e 3MI9 com a presença do ligante AMP. **B.** Sobreposição do modelo CRK9.57 refinado **sem inserções** com os moldes 4IMY e 3MI9, com a presença do ligante AMP. **C.** Destaque para os 3 resíduos catalíticos da CRK9, K301, D485, D503, e para as 2 cisteína mapeadas C425 (Hinge H2) e C502 (DFG D1) em sobreposição com os respectivos resíduos correspondentes nas estruturas moldes. **D.** Rotâmero *outlier* (I266) na estrutura do modelo CRK9.57 refinado na região do sítio de ligação do ATP em sobreposição com os moldes 4IMY e 3MI9. 111

Figura 33 - Destaque das diferenças no sítio de ligação do ATP entre a estrutura do modelo CRK9.57 refinado da quinase CRK9 de *T. cruzi* da cepa CL Brener El e a estrutura do molde 4IMY. A estrutura do molde 3MI9 não foi representada nesta figura. **A.** Destaque em bastões dos resíduos diferentes entre o modelo CRK9.57 refinado e o molde 4IMY que se encontram a até 6 Å do ligante AMP. A estrutura da molécula AMP foi omitida para melhor visualização do sítio. **B.** Superfície eletrostática da região analisada para o modelo (à esquerda) e o molde (à direita) gerada pelo *plugin* APBS do PyMOL^{106,107}. A coloração em vermelho indica regiões de potencial eletrostático negativo e em azul regiões de potencial positivo..... 112

Figura 34 - Comparativo entre o modelo CRK9.57 gerado por modelagem comparativa com o modelo gerado pelo AlphaFold. **A.** Modelo completo gerado pelo AlphaFold. As cores representam o grau de confiabilidade na predição da estrutura terciária, sendo: laranja confiança muito baixa; amarelo confiança baixa; azul claro boa confiança e azul escuro alta confiança. **B.** Sobreposição do modelo CRK9.57 gerado por modelagem comparativa com o modelo gerado pelo AlphaFold sem os trechos de inserção..... 113

Figura 35 - Alinhamento das sequências das quinases CRK9 correspondentes em outras cepas de *T. cruzi* com destaque para os resíduos que se encontram no sítio do ATP a até 6 Å de distância do ligante AMP (molde 4IMY). As sequências foram alinhadas considerando o intervalo do aminoácido 250 ao 550 da quinase CRK9 cepa CL Brener El. O trecho de grande inserção entre as fitas β4 e β5, do aminoácido 342 ao 412, foi omitido. (ponto de corte marcado pela linha e seta em azul)..... 114

LISTA DE TABELAS

Tabela 1 - Quinases de <i>Trypanosoma cruzi</i> e seus respectivos papéis biológicos descritos na literatura.....	26
Tabela 2 - Tipos de inibidores das proteínas quinases	41
Tabela 3 - Inibidores covalentes de quinases humanas aprovados pela FDA e suas indicações terapêuticas	42
Tabela 4 - Comparativo da classificação das quinases de <i>T. cruzi</i> das cepas Dm28c 2014 e CL Brener El em grupos e famílias, juntamente com classificação feita por Parsons <i>et al.</i> ¹⁵ para a cepa CL Brener.....	60
Tabela 5 - Compilado dos sítios e subsítios dos mapeamentos de cisteínas promissoras para uma estratégia de inibição covalente elaborados por Zhao <i>et al.</i> ⁶¹ e Chaikuad <i>et al.</i> ⁶²	61
Tabela 6 – Classificação em grupos e famílias das 100 quinases de <i>T. cruzi</i> da cepa CL Brener El com cisteínas mapeadas, e detalhamento dos sítios e subsítios de localização das cisteínas	63
Tabela 7 - Síntese do mapeamento das cisteínas das quinases de <i>T. cruzi</i> das cepas Dm28c 2014 e CL Brener El em sítios e subsítios, juntamente com dados do mapeamento de cisteínas de quinases humanas feito por Chaikuad <i>et al.</i> ⁶²	68
Tabela 8 - Quinases de <i>T. brucei</i> destacadas como importantes para a proliferação e/ou ciclo celular do parasito segundo Jones <i>et al.</i> (2014) ³¹ e Fernandez-Cortes <i>et al.</i> (2017) ³²	71
Tabela 9 - Quinases de <i>T. cruzi</i> ortólogas das quinases de <i>T. brucei</i> destacadas por Jones <i>et al.</i> ³¹ e Fernandez-Cortes <i>et al.</i> ³² em estudos de RNAi como importantes para a proliferação e/ou ciclo celular do parasito.....	73
Tabela 10 - Comparativo do número de resíduos, identidade e número de cisteínas mapeadas entre as sequências de quinases AUK1 de diversas cepas de <i>T. cruzi</i> com a sequência AUK1 da cepa CL Brener Esmeraldo-like (CL Brener El)	78
Tabela 11 - Dados dos candidatos a moldes para a modelagem comparativa da quinase AUK1 da cepa CL Brener NEI de <i>T. cruzi</i>	78
Tabela 12 - Dados de validação dos 3 modelos pré-selecionados da modelagem da quinase AUK1 de <i>T. cruzi</i> da cepa CL Brener NEI (AUK1-2BFY) gerados no programa Modeller ...	82
Tabela 13 - Dados de validação dos 3 modelos pré-selecionados da modelagem da quinase AUK1 de <i>T. cruzi</i> da cepa CL Brener NEI (AUK1-1OL5) gerados no programa Modeller....	87
Tabela 14 - Comparativo do número de resíduos, identidade e número de cisteínas mapeadas entre as sequências de quinases CLK1 de diversas cepas de <i>T. cruzi</i> com a cepa CL Brener Esmeraldo-like (CL Brener El).....	94

Tabela 15 - Dados dos candidatos a moldes para a modelagem comparativa da quinase CLK1 da cepa CL Brener El de <i>T. cruzi</i>	96
Tabela 16 - Dados de validação dos 3 modelos pré-selecionados da modelagem da quinase CLK1 de <i>T. cruzi</i> da cepa CL Brener El gerados no programa Modeller	98
Tabela 17 - Comparativo do número de resíduos, identidade e número de cisteínas mapeadas entre as sequências de quinases CRK9 de diversas cepas de <i>T. cruzi</i> com a sequência CLK1 da cepa CL Brener Esmeraldo-like (CL Brener El)	103
Tabela 18 - Dados dos candidatos a molde(s) para a modelagem comparativa da quinase CRK9 da cepa CL Brener El de <i>T. cruzi</i>	103
Tabela 19 - Dados de validação dos 3 modelos pré-selecionados da modelagem da quinase CRK9 de <i>T. cruzi</i> da cepa CL Brener El gerados no programa Modeller.....	108
Tabela 20 - Tabela síntese sobre as 3 quinases de <i>T. cruzi</i> priorizadas e de suas respectivas modelagens comparativas.....	115

LISTA DE ABREVIATURAS E SIGLAS

ADP	Adenosina Difosfato
aPKs	<i>atypical Protein Kinases</i>
ATP	Adenosina Trifosfato
BLAST	<i>Basic Local Alignment Search Tool</i>
cAMP	<i>cyclic Adenosine Monophosphate</i>
CL Brener EI	CL Brener Esmeraldo-like
CL Brener NEI	CL Brener Non-Esmeraldo-like
CPC	<i>Chromosomal Passenger Complex</i>
<i>C-spine</i>	<i>Catalytic spine</i>
DFT	<i>Density Functional Theory</i>
DOPE	<i>Discrete Optimized Protein Energy</i>
ePKs	<i>eukaryotic Protein Kinases</i>
EWG	<i>Electron-Withdrawing Groups</i>
FDA	<i>US Food and Drug Administration</i>
Fs-IFP	<i>Function-site Interaction FingerPrint</i>
hEGFR	<i>human Epidermal Growth Factor Receptor</i>
locPREFMD	<i>Local Protein Structure Refinement via Molecular Dynamics</i>
Molpdf	<i>Molecular probability density function</i>
na	not assigned to a family
NCBI	<i>National Center for Biotechnology Information</i>
OMS	Organização Mundial da Saúde
PAHO	<i>Pan-American Health Organization</i>
PDB	<i>Protein Data Bank</i>
Pfam	<i>Protein families database</i>
PKs	<i>Protein Kinases</i>
ProSA	<i>Protein Structure Analysis</i>
RMN	Ressonância Magnética Nuclear
RMSD	<i>Root Mean Square Deviation</i>

RNAi	<i>RNA interference</i>
<i>R-spine</i>	<i>Regulatory spine</i>
SIB	<i>Swiss Institute of Bioinformatics</i>
<i>T. brucei</i>	<i>Trypanosoma brucei</i>
<i>T. cruzi</i>	<i>Trypanosoma cruzi</i>
UniProt	<i>The Universal Protein Resource</i>
VEuPathDB	Eukaryotic Pathogen, vector and host informatics resource center

SUMÁRIO

1	INTRODUÇÃO	18
1.1	PANORAMA SOBRE O <i>TRYPANOSOMA CRUZI</i> E A DOENÇA DE CHAGAS .	18
1.2	PROTEÍNAS QUINASES DE <i>TRYPANOSOMA CRUZI</i>	24
1.3	CARACTERÍSTICAS ESTRUTURAIS DAS PROTEÍNAS QUINASES.....	27
1.4	INIBIDORES COVALENTES COMO ESTRATÉGIA DE AÇÃO EM PROTEÍNAS QUINASES	38
1.5	MODELAGEM COMPARATIVA.....	49
2	OBJETIVOS	52
2.1	OBJETIVO GERAL.....	52
2.2	OBJETIVOS ESPECÍFICOS	52
3	METODOLOGIA	53
3.1	LEVANTAMENTO E CLASSIFICAÇÃO DAS QUINASES DE <i>TRYPANOSOMA CRUZI</i>	53
3.2	MAPEAMENTO DAS CISTEÍNAS RELEVANTES PARA INIBIÇÃO COVALENTE	54
3.3	PRIORIZAÇÃO DAS QUINASES	54
3.4	MODELAGEM COMPARATIVA.....	55
3.4.1	Obtenção dos modelos.....	55
3.4.2	Validação dos modelos	56
3.4.3	Análises complementares dos modelos gerados.....	58
4	RESULTADO E DISCUSSÃO.....	59
4.1	LEVANTAMENTO E CLASSIFICAÇÃO DAS QUINASES DO <i>TRYPANOSOMA CRUZI</i>	59
4.2	MAPEAMENTO DAS CISTEÍNAS RELEVANTES PARA INIBIÇÃO COVALENTE	61

4.3	PRIORIZAÇÃO DAS QUINASES	70
4.4	MODELAGEM COMPARATIVA DAS QUINASES PRIORIZADAS	76
4.4.1	AUK1	76
4.4.1.1	AUK1-2BFY	80
4.4.1.2	AUK1-1OL5.....	86
4.4.2	CLK1	93
4.4.3	CRK9	102
5	CONCLUSÕES E PERSPECTIVAS	116
	REFERÊNCIAS BIBLIOGRÁFICAS	119
	APÊNDICE A - Tabela de classificação em grupos e famílias das 171 quinases de <i>T. cruzi</i> da cepa Dm28c 2014, e detalhamento do mapeamento das cisteínas em sítios e subsítios	126
	APÊNDICE B - Tabela de classificação das 75 quinases de <i>T. cruzi</i> da cepa CL Brener El que não apresentaram cisteínas em nenhuma das 25 posições mapeadas	132

1 INTRODUÇÃO

1.1 PANORAMA SOBRE O *TRYPANOSOMA CRUZI* E A DOENÇA DE CHAGAS

A doença de Chagas, também conhecida como tripanossomíase americana, e o organismo *Trypanosoma cruzi* (*T. cruzi*), agente etiológico causador da doença, foram primeiramente descritos em 1909 pelo médico pesquisador brasileiro Carlos Justiniano Ribeiro Chagas¹. Desde então, novos estudos ampliaram o conhecimento sobre essa enfermidade e identificaram a doença e a presença do protozoário em diversos países da América Latina, onde a doença é endêmica^{2,3}. Este organismo já está presente nas Américas há milhões de anos e a chegada dos humanos nestes continentes, acompanhado do desenvolvimento da agricultura e da domesticação de animais, tornaram o homem mais um dos hospedeiros deste parasito há cerca de 10 mil anos^{2,4}. Mesmo passados mais de 100 anos da descoberta do *T. cruzi* e da doença de Chagas, esta é uma doença ainda presente nestes países, afetando principalmente áreas rurais e com problemas de saneamento e pobreza, demandando grandes investimentos da saúde pública^{4,5}. Neste contexto, esta doença é classificada como uma Doenças Tropicais Negligenciadas pela Organização Mundial da Saúde (OMS), cujo controle epidemiológico, métodos diagnósticos e terapêuticos carecem de aprimoramento e maiores investimentos⁵. Estima-se que de 6 a 8 milhões de pessoas estejam contaminadas pelo *T. cruzi* e que cerca de 65 a 100 milhões de pessoas vivam em áreas de risco de contaminação^{3,6,7}. A doença de Chagas causa cerca de 12 mil mortes por ano^{5,6} e a média anual de custos com tratamentos de saúde para os casos crônicos, juntamente com perdas econômicas relacionadas à produtividade, fica em torno de 4.660 dólares por pessoa e de mais de 7 bilhões de dólares considerando toda a população afetada⁸.

Grande parte dos esforços de controle da doença focaram no controle e eliminação dos vetores transmissores do *T. cruzi*. Este protozoário é transmitido principalmente pelo contato das fezes contaminadas de insetos triatomíneos hematófagos com o local da picada. No Brasil, estes insetos são popularmente conhecidos como barbeiros. Os barbeiros, ou também chamados *kissing bugs*, *vinchucas*, *pitos*, *chirimachas*, que transmitem o parasito são dos gêneros *Triatoma*, *Panstrongylus* e *Rhodnius*^{5,6}. Considerando o controle da propagação destes insetos e da transmissão da doença, a Organização Panamericana de Saúde - *Pan-American Health*

Organization (PAHO) certificou o controle da transmissão via o vetor *Triatoma infestans* no Uruguai em 1997, no Chile em 1999 e no Brasil em 2006². No Brasil ainda existem focos residuais de *Triatoma infestans* na Bahia e no Rio Grande do Sul⁹. Além disso, devido à grande mobilidade urbana dos tempos atuais, esta doença já está presente em diversos outros países e já não é mais um problema apenas dos países endêmicos e de áreas rurais². Na **Figura 1** está ilustrado o movimento de espalhamento do parasito e da doença por meio das migrações e a estimativa de casos em algumas regiões não endêmicas.

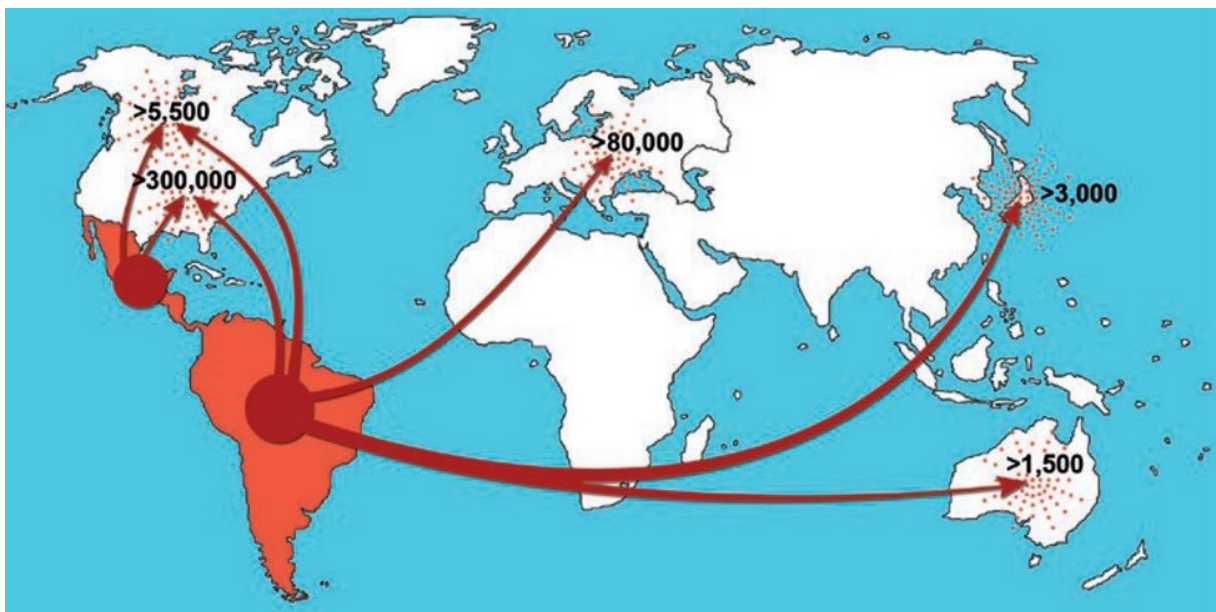


Figura 1 - Rotas migratórias da América Latina para países não endêmicos e estimativa do número de indivíduos infectados nos Estados Unidos, Canadá, Europa e países da região ocidental do pacífico, Japão e Austrália.

Fonte: Coura; Viñas. 2010 (doi: 10.1038/nature09221)⁴

Além da transmissão pela picada do vetor, o parasito *T. cruzi* também pode ser transmitido por transfusão de sangue, transplante de órgãos, acidentes em laboratórios e de forma congênita, sendo estas as formas predominantes de contágio nos países não endêmicos³⁻⁵. Em muitos países não endêmicos não se tem um sistema de rastreio bem estabelecido para este parasito, contribuindo para sua propagação^{7,10}. Outra forma de transmissão é por meio da ingestão de alimentos contaminados. No Brasil, segundo Boletim Epidemiológico do Ministério da Saúde referente ao período de 2012 a 2016, esta é atualmente a principal forma de transmissão do parasito no país. Grande parte dos casos estão concentrados na região amazônica,

principalmente no Estado do Pará, responsável por mais de 80% dos casos registrados no período⁹.

O *T. cruzi* apresenta três formas ou estágios de desenvolvimento durante seu ciclo de vida: forma epimastigota, proliferativa, na qual o protozoário se multiplica no intestino do inseto triatomíneo vetor; duas formas tripomastigotas, a tripomastigota metacíclica, forma infectante, presente no terço final do trato digestivo do inseto vetor, e a tripomastigota sanguínea, forma infectante extracelular no hospedeiro vertebrado; e, por fim, forma amastigota, proliferativa intracelular, também no hospedeiro vertebrado^{3,7}.

Na **Figura 2** está representado de forma esquemática a transmissão e o ciclo de vida do *T. cruzi*, partindo do inseto vetor. (1) O parasito entra no organismo do hospedeiro vertebrado através da lesão da picada do inseto contaminado que, no ato da picada, excreta suas fezes contendo a forma tripomastigota metacíclica. O hospedeiro, ao coçar a região da picada, leva as fezes contaminadas para a lesão ou para as mucosas, que também são porta de entrada para o parasito. Neste tipo de contaminação pode ocorrer a formação de edema no local da picada, o chamado chagoma, ou também edema unilateral da pálpebra, conhecido como sinal romaña, no caso de contato com a mucosa dos olhos. (2) Após a entrada do *T. cruzi* nas células do hospedeiro vertebrado, esse se diferencia na forma proliferativa intracelular amastigota. (3) Em seguida à esta fase de multiplicação do parasito, ocorre a diferenciação para a forma infectante tripomastigota sanguínea. (4) As células contendo os tripomastigotas se rompem e o parasito entra na corrente sanguínea e invade novas células do hospedeiro e inicia assim um novo ciclo. (5) Caso um barbeiro se alimente do sangue de um hospedeiro contaminado, ele irá se contaminar com a forma tripomastigota sanguínea que se diferenciará para a forma epimastigota proliferativa no intestino do inseto. (6) Posteriormente, no terço final do trato digestivo do inseto vetor, o parasito se diferencia para a forma infectante tripomastigota metacíclica, podendo contaminar novos hospedeiros⁷.

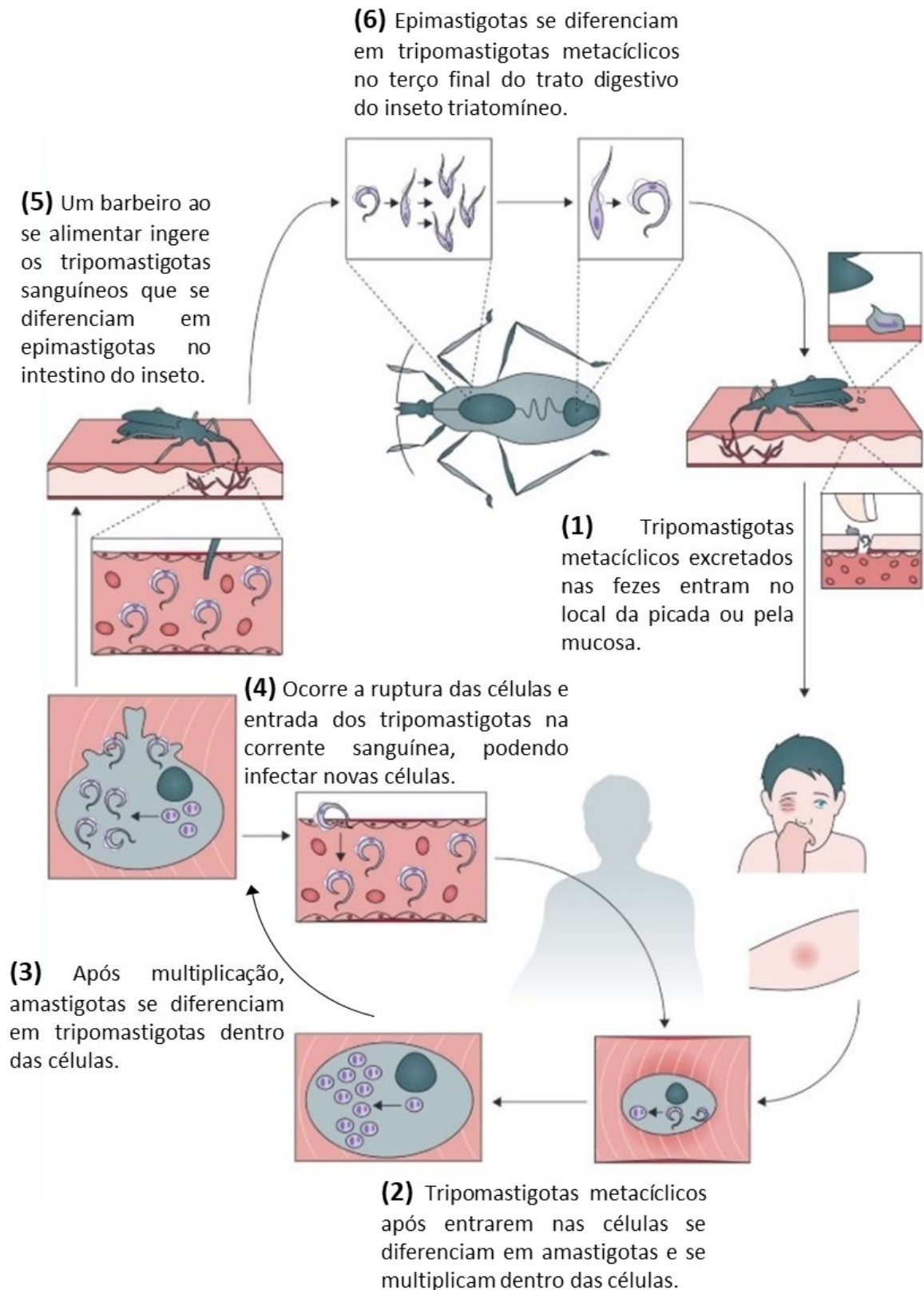


Figura 2 - Ciclo do parasito *Trypanosoma cruzi* no inseto triatomíneo e no hospedeiro homem
 Fonte: adaptada de Pérez-Molina, 2018 (doi: 10.1016/S0140-6736(17)31612-4)⁵

A doença de Chagas é caracterizada por duas fases clínicas, uma fase aguda e uma crônica. Após a contaminação inicia-se a fase aguda, que se estende por cerca de 2 meses, apresentando uma alta parasitemia. Geralmente se manifesta de forma assintomática ou oligossintomática, com um quadro febril leve, dores de cabeça, diarreia, fadiga, dentre outros sintomas menos comuns⁷. Nesta fase o parasito invade o fígado, intestino, baço, gânglios linfáticos, sistema nervoso central e músculos esqueléticos e cardíacos⁷. A falta de sintomas ou a associação destes com outras patologias contribuem para a não detecção precoce da doença, limitando a possibilidade de tratamentos na fase aguda e contribuindo para a evolução da doença para a fase crônica. Na fase crônica, caracterizada por baixa parasitemia, uma parcela dos contaminados seguirão assintomáticos e a fase recebe o nome de indeterminada. Porém, cerca de 30 a 40% desenvolverão alguma condição debilitante nos tecidos e órgãos contaminados e, por esta razão, o diagnóstico muitas vezes acontece apenas neste estágio da doença. Nesta fase, dentre as principais manifestações clínicas tem-se as digestivas, com a síndrome do megacólon e do megaesôfago; manifestações cardíacas (mais comuns), como miocardite, arritmias, insuficiência cardíaca, morte cardíaca repentina, fenômenos tromboembólicos; manifestações cardiodigestivas, e também manifestações neurais, sendo esta última mais rara^{3,7}.

O tratamento disponível hoje para tentar eliminar ou diminuir a parasitemia é baseado apenas em duas medicações licenciadas, nifurtimox e benzonidazol (**Figura 3**)⁵. Estas duas medicações são a base do tratamento da doença de Chagas há quase 50 anos, mas apresentam baixa eficácia nos casos crônicos e severos efeitos adversos. No Brasil, desde os anos 80, apenas o uso do benzonidazol é autorizado¹¹. Neste mesmo período outros países endêmicos também descontinuaram a comercialização do nifurtimox¹¹. O sucesso do tratamento depende em grande medida da detecção precoce do parasito e da idade do paciente. O tratamento também se mostra eficiente nos casos congênitos e de reativação da doença^{5,7}.

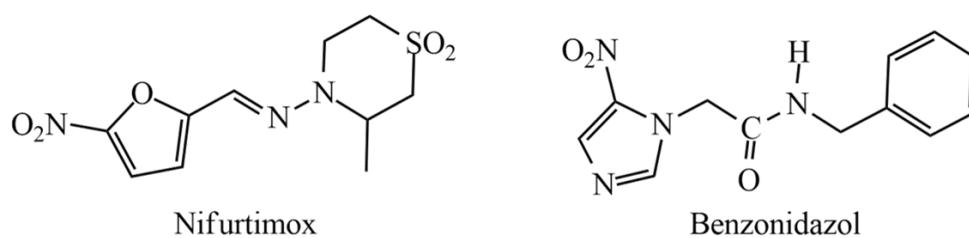


Figura 3 - Estrutura química dos fármacos disponíveis para o controle e eliminação do parasito *Trypanosoma cruzi*, nifurtimox e benzonidazol.

O nifurtimox é ministrado por 2 a 3 meses, apresentando nos casos agudos uma taxa de cura de 65 a 80%⁷ e nos casos crônicos indeterminados uma taxa de cura de 86% para indivíduos abaixo de 14 anos e entre 7 e 8% em adultos⁵. Dentre os diversos efeitos adversos, os mais comuns são anorexia, perda de peso, náusea, vômito e desordens neurológicas, como insônia, irritabilidade, desorientação^{5,7}. O benzonidazol é ministrado por cerca de 2 meses e é descrito como mais tolerável e com melhor penetração nos tecidos. Nos casos agudos de infecção, apresenta cerca de 76% de taxa de cura, enquanto na fase crônica a taxa de cura fica entre 60 e 93% em jovens de até 13 anos e de 2 a 40% em adultos⁵. Apesar de ser mais tolerável que o nifurtimox, o benzonidazol ainda apresenta diversos efeitos adversos, como erupções cutâneas, intolerância digestiva, anorexia, dores de cabeça, fadiga, desordens do sono e, em casos raros, supressão da medula óssea que leva à neutropenia (redução do número de neutrófilos)^{5,7}. O longo tratamento e os efeitos colaterais associado às duas medicações contribuem para uma alta taxa de abandono dos tratamentos, comprometendo ainda mais o controle e eliminação do parasito e da doença de Chagas.

Neste contexto, maiores investimentos em pesquisas voltadas para o desenvolvimento de fármacos que consigam ampliar a taxa de cura, principalmente para os casos crônicos, e que apresentem menos efeitos adversos ou demandem menor tempo de tratamento, são essenciais para um gerenciamento efetivo da doença de Chagas. Para este processo, o primeiro passo é definir um bom alvo terapêutico. Na perspectiva dos parasitos, um bom alvo é aquele que está relacionado com processos fisiológicos importantes e vitais no organismo. O *T. cruzi*, dada sua capacidade de evadir o sistema imune na fase crônica da doença, demanda que os alvos terapêuticos explorados promovam a morte do organismo quando inibidos¹². Além disso, a disponibilidade de informações sobre estes potenciais alvos, como suas características estruturais, interações relevantes, mecanismos reacionais e função biológica são importantes para que se possa desenvolver fármacos¹³. Alguns alvos terapêuticos tradicionais em *T. cruzi* têm sido explorados nos últimos anos, como: a cisteína protease cruzipaina, envolvida na invasão de células hospedeiras, na modulação da resposta de macrófagos e na diferenciação do parasito; enzimas envolvidas na biossíntese do ergosterol, importante para o crescimento do organismo, como a esterol 14-demetilase e a farnesiltransferase; a gliceraldeído-3-fosfato desidrogenase, a hexoquinase e a fosfofrutoquinase, envolvidas da via glicolítica; enzimas envolvidas no metabolismo de tripanotona, que fazem parte do mecanismo de defesa do *T. cruzi* contra o estresse oxidativo, como a tripanotona redutase e a tripanotona sintetase; e a trans-sialidase, enzima que pertence à família de glicoproteínas de superfície^{11,14}. As proteínas

quinases – *Protein Kinases* (PKs) são uma grande e importante classe de enzimas que catalisam reações de fosforilação, coordenando um grande escopo de reações relacionadas com a fisiologia celular em inúmeros organismos, e, dada essa relevância biológica, também são alvos terapêuticos promissores para serem explorados no tratamento da doença de Chagas.

1.2 PROTEÍNAS QUINASES DE *TRYPANOSOMA CRUZI*

Os tripanossomatídeos, como o *Trypanosoma cruzi*, *Trypanosoma brucei*, e *Leishmania major*, apresentam um número de proteínas quinases significativamente maior que outros parasitos intracelulares¹⁵. O *T. cruzi* possui cerca de 190 destas proteínas, sugerindo que a fosforilação seja um mecanismo de regulação celular importante neste organismo, tanto no ciclo celular quanto no processo de diferenciação entre as diferentes formas do parasito¹⁵.

Apesar da escassez de estudos validando o papel biológico de quinases em *T. cruzi*, já foram descritos papéis importantes destas proteínas. Por exemplo, Satori *et al.* (2002) e Cunho *et al.* (2005) demonstraram que a quinase CRK3 atua no ciclo celular do parasito, fosforilando a histona H1 quando a célula progride da fase S para a fase M, tendo também um papel na transição das fases G2/M^{16,17}. Outra quinase, chamada TcVps34 (PI3K), se mostrou importante na adaptabilidade do organismo à variação de ambientes a que o parasito está sujeito durante suas diferentes formas. A TcVps34 está envolvida em processos vitais como osmorregulação, acidificação e tráfego vesicular¹⁸. Outro estudo, referente à quinase TcVps15, demonstrou que esta proteína participa da regulação da atividade enzimática da TcVps34, e que o complexo formado por estas duas proteínas está relacionado com a atividade de autofagia no *T. cruzi*, processo envolvido na diferenciação do organismo¹⁹. A quinase PKA também contribui para a regulação dos processos de diferenciação, além de regulação do crescimento do parasito²⁰. A inibição desta quinase em estudos *in vitro*, utilizando a forma epimastigota, mostrou-se letal para o organismo²⁰. Foram identificados diversos substratos da quinase PKA, incluindo a quinase TcVps34, apontando seu possível papel na regulação do tráfego intracelular, na adaptação às alterações de ambientes, na regulação da homeostasia do monofosfato cíclico de adenosina (cAMP), e também na invasão das células do hospedeiro e evasão do seu sistema imune^{20,21}. A quinase TcMAPK2, que também é um dos substratos da PKA, é relatada como tendo importante papel na resposta ao stress, o que indica sua possível importância na regulação

da proliferação e diferenciação do parasito²². A interação da TcMAPK2 com a PKA e outras enzimas apontam para um provável papel desta quinase também na homeostasia do cAMP²². Outras quinases foram relacionadas com a modulação da resposta ao stress, como as quinases TcK1 e TcK2. Malvezzi *et al.* (2020) analisaram o papel da TcK1 em resposta ao stress nutricional e demonstraram que esta enzima afeta a formação dos grânulos de stress e modula o destino do parasito durante sua diferenciação, invasão e proliferação na condição de inanição nutricional²³. Em relação à TcK2, Augusto *et al.* (2015) mostraram que o seu silenciamento levou a um comprometimento do crescimento e da capacidade infecciosa do organismo²⁴. Moléculas heme advindas da digestão do sangue do hospedeiro se ligam à quinase TcK2 inativando-a e mantendo o crescimento celular da forma proliferativa do *T. cruzi*. A ausência da quinase TcK2 em presença de moléculas heme acarretou na geração e acúmulo de espécies reativas de oxigênio, causando dano celular, além de inibir a metaciclologênese, processo no qual o parasito na forma proliferativa epimastigota se diferencia para a forma infecciosa tripomastigota metacíclica no inseto vetor²⁴. A metaciclologênese se mostrou dependente da presença da forma ativa da TcK2. Outra quinase que também teve sua função associada com a molécula heme foi a CAMKII. Esta quinase, diferentemente da TcK2, é ativada na presença de molécula heme e está envolvida na via de sinalização que media o processo de proliferação da forma epimastigota do parasito²⁵. A CAMKII faz parte da via de sinalização redox intracelular na presença de molécula heme, com produção de espécies reativas de oxigênio que estão relacionadas com o desencadeamento da proliferação do *T. cruzi* no inseto²⁶. Mais uma quinase descrita como tendo um papel na diferenciação do *T. cruzi* é a PKC. Esta quinase também está envolvida na via de sinalização da metaciclologênese²⁷. Uma proteína que se mostrou importante no processo de divisão celular e possivelmente na diferenciação do parasito é a AUR1. Estudos conduzidos por Fassolari e Alonso (2019) demonstraram que esta quinase está relacionada com a montagem do fuso mitótico e segregação cromossômica, além de atuarem na iniciação da duplicação do cinetoplasto, uma rede de DNA circular associada à mitocôndria que coordena sua divisão e segregação com a mitose nuclear²⁸. Por fim, no que diz respeito à expressão gênica, mais especificamente ao processamento do RNA mensageiro, a quinase TcSRPK foi descrita como responsável pela regulação da atividade da *Trypanosoma cruzi* serine/arginine protein (TcSR) através da fosforilação dos resíduos de serina. As proteínas SR são fatores essenciais no *splicing* durante a expressão gênica e o ciclo celular²⁹. Na **Tabela 1** estão reunidas as informações sobre as quinases de *T. cruzi* estudadas em relação aos processos biológicos nos quais estão envolvidas.

Tabela 1 - Quinases de *Trypanosoma cruzi* e seus respectivos papéis biológicos descritos na literatura.

Quinase	Processos biológicos
CRK3	Ciclo celular ^{16,17}
TcVps34	Osmorregulação, acidificação, tráfego celular e autofagia ¹⁸
TcVps15	Regulação da TcVps34, autofagia ¹⁹
PKA	Tráfego intracelular, homeostasia cAMP, adaptabilidade a alterações ambientais, invasão das células do hospedeiro e evasão do sistema imune ^{20,21}
TcMAPK2	Resposta ao stress, homeostasia do cAMP ²²
TcK1	Resposta ao stress nutricional ²³
TcK2	Resposta ao stress nutricional, metaciclogênese ²⁴
CAMKII	Via de sinalização redox, proliferação da forma epimastigota ^{25,26}
PKC	Metaciclogênese ²⁷
AUR1	Divisão celular ²⁸
TcSRPK	Expressão gênica ²⁹

Diante das evidências apresentadas sobre a função e importância biológica de algumas das quinases de *T. cruzi* fica notório que esta classe de proteínas se mostra promissora para ser explorada como alvo terapêutico para o controle e cura da doença de Chagas. Devido à complexidade e particularidade de seu genoma, muitas ferramentas de estudos para definição de alvos terapêuticos e de candidatos a inibidores se mostram limitadas para o *T. cruzi*^{12,30}. Um exemplo é a técnica de RNA de interferência - *RNA Interference* (RNAi). Enquanto o *T. cruzi* não possui um sistema funcional de RNAi, estudos de RNAi em *Trypanosoma brucei* (*T. brucei*), permitiram validar as quinases como alvos promissores neste organismo^{12,30,31,32}. Nestes estudos foram determinadas quais quinases eram mais relevantes para a proliferação, ciclo celular e sobrevivência deste organismo através do silenciamento de cada quinase. No trabalho de Jones *et al.* (2014)³¹ identificaram 42 linhagens cuja proliferação foi afetada devido ao silenciamento das quinases. Destas, 24 demonstraram ter um papel também no ciclo celular e 11 quinases silenciadas levaram à morte do organismo. Em outro estudo, de Fernandez-Cortes *et al.* (2017)³², além do estudo *in vitro* foi realizado o silenciamento das quinases *in vivo*, em camundongos. Seus resultados foram convergentes com os de Jones *et al.*, com 35 quinases destacadas em ambos. A partir dos resultados *in vivo*, além destas 35 quinases, outras 14 se mostraram biologicamente relevantes para o organismo. Grande parte destas 14 quinases não apresentaram perda de *fitness* significativa nos estudos *in vitro* ou mesmo apresentaram alterações apenas no estudo *in vivo*, indicando um possível papel destas quinases na evasão do

sistema imune do hospedeiro e na resposta ao stress. Estes resultados em quinases de *T. brucei* reforçam a importância dessa classe de proteínas para tripanossomatídeos.

1.3 CARACTERÍSTICAS ESTRUTURAIS DAS PROTEÍNAS QUINASES

As fosforilações mediadas pelas quinases podem resultar na ativação ou inativação de enzimas, criar sítio de reconhecimento para o recrutamento de outras proteínas e gerar ajustes conformacionais nas estruturas das proteínas³³. Sua relevância bioquímica fica evidenciada pelo grande número de proteínas quinases presentes nos organismos eucarióticos. O quinoma humano, por exemplo, é constituído por mais de 550 destas proteínas³⁴.

Esta classe de proteínas é subdividida de acordo com o resíduo de aminoácido que fosforilam em seus substratos. Comumente são ou serinas e treoninas ou tirosinas, e as quinases são denominadas como serina/treonina quinases ou tirosina quinases, respectivamente³⁵. Contudo, algumas PKs são capazes de fosforilar estes três resíduos de aminoácidos, sendo denominadas quinases duais^{36,37}. As quinases atuam conduzindo a transferência de um grupo fosforil de uma molécula doadora, geralmente o ATP, para um substrato proteico aceptor³⁶ (**Figura 4**). Durante a reação de fosforilação ocorre a transferência do fosfato γ do ATP para a hidroxila alcoólica de uma serina ou treonina ou para a hidroxila fenólica de uma tirosina de um substrato³⁸.

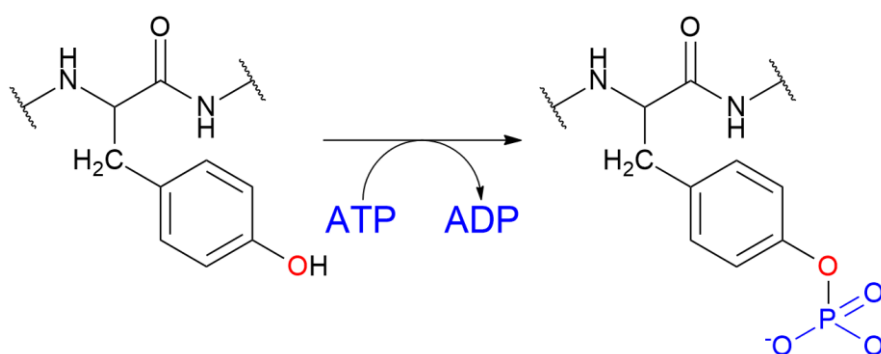


Figura 4 - Esquema da reação de fosforilação de uma proteína quinase

As PKs apresentam um enovelamento característico, possuindo uma estrutura tridimensional central conservada e um domínio catalítico constituído por cerca de 250 a 300 aminoácidos, que pode se estender caso existam trechos de inserção^{35,38,39}. O domínio catalítico das quinases não é conservado de maneira uniforme ao longo da sequência de aminoácidos, ele apresenta regiões com graus alternados de conservação. Devido a este fato, o domínio catalítico é dividido em 12 subdomínios, indicados por algarismo romano, majoritariamente constituídos de fitas β e α -hélices, nas quais há maior conservação, e pelas regiões de alça presentes entre estas estruturas secundárias, geralmente possuindo menor conservação^{35,38,39} (**Figura 5**).

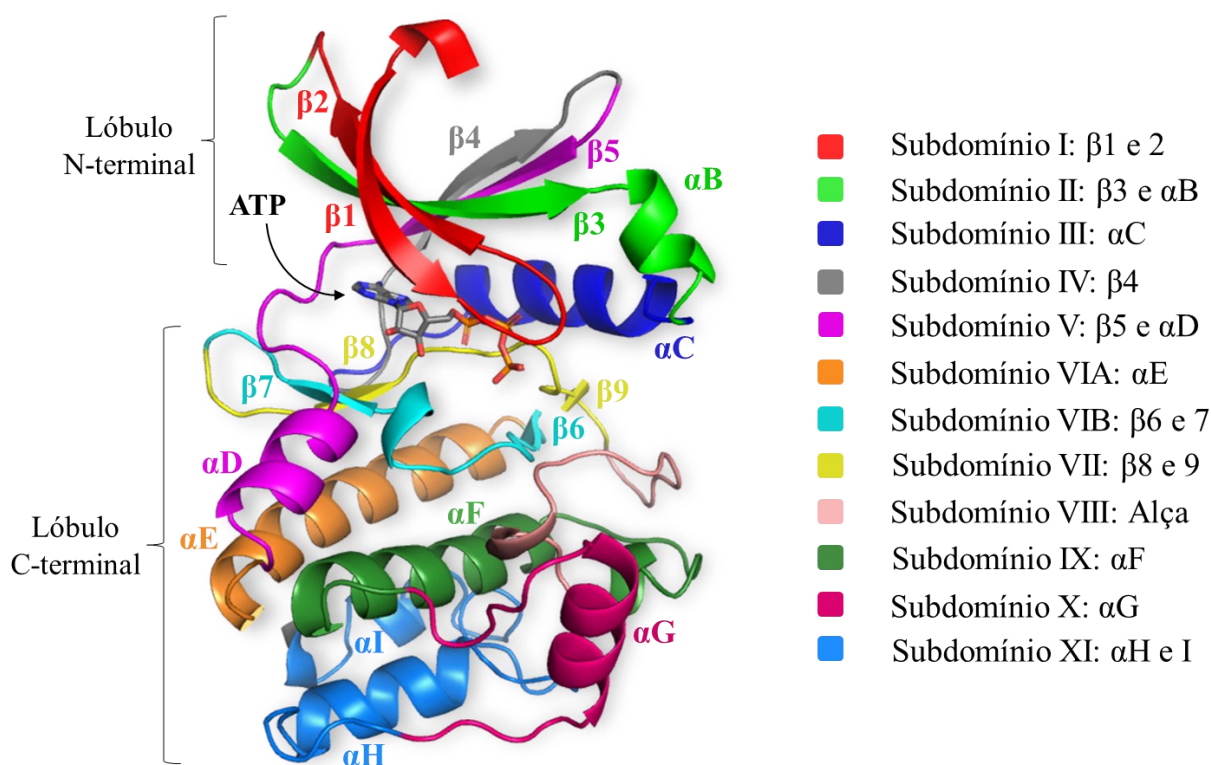


Figura 5 - Identificação dos subdomínios do domínio catalítico das proteínas quinases. Os 12 subdomínios estão apresentados em diferentes cores juntamente com as estruturas secundárias (β para fitas β e α para α -hélices) enumeradas. Também foram indicadas na figura a localização dos dois lóbulos que correspondem aos grandes agrupamentos de estruturas secundárias do domínio catalítico, e a presença do ATP em bastões na fenda catalítica entre os dois lóbulos.

A região rica em fitas β antiparalelas localizada na porção N-terminal, também chamada de lóbulo N-terminal ou lóbulo N, é formada pelos subdomínios I, II, III e IV (**Figura 5**). Este lóbulo está envolvido na orientação e no ancoramento da molécula doadora do grupo fosforil (ATP)^{38,40}. A região rica em α -hélices localizada na porção C-terminal é um lóbulo maior

formado pelos subdomínios VIA ao XI e é chamado de lóbulo C-terminal ou Lóbulo C (**Figura 5**). Esta região da proteína está envolvida na ligação do substrato proteico e iniciação da fosfotransferência^{38,40}. O ATP, ou outra molécula doadora de grupo fosforil, se liga na fenda profunda situada entre estes dois lóbulos. As fitas β são classificadas de 1 a 9 e as estruturas de α -hélices da letra A à I. Adotar esta enumeração permite ter uma referência de localização destas estruturas secundárias ao longo do domínio catalítico e discutir a importância e o papel de cada parte da proteína na sua ativação e na atividade catalítica das quinases. Naturalmente, outras proteínas quinases apresentam algum grau de variação destas estruturas. A α -hélice A não foi representada por não estar dentro da estrutura base do domínio de quinases, segundo Hanks e Hunter (1995). Porém, estes autores mencionam que esta estrutura parece ser importante em algumas quinases na estabilização dos dois lóbulos.

As quinases com as características estruturais descritas são chamadas de proteínas quinase eucarióticas - *eukaryotic Protein Kinases* (ePKs), nomenclatura bastante empregada na literatura apesar de não serem exclusivas de organismos eucarióticos³⁸. Existe um conjunto reduzido de quinases que não apresenta uma similaridade de sequência com o domínio das ePKs, sendo chamado de proteínas quinases atípicas – *atypical Protein Kinases* (aPKs). Apesar da ausência dos motivos das quinases eucarióticas, as aPKs apresentam atividade catalítica de fosforilação como as demais proteínas quinases⁴¹. Além disso, comumente as aPKs compartilham um enovelamento típico das ePKs e, nestes casos, também são classificadas como *Protein Kinase-Like* (PKL)³⁴.

Estudos que analisaram estruturas, alinhamentos múltiplos e árvores filogenéticas destas enzimas, permitiram a identificação de motivos e outros detalhes estruturais. Estas informações possibilitaram classificar as quinases segundo modos de regulação, especificidade de substratos e funções, dividindo-as em grupos que por sua vez são subdivididos em famílias e subfamílias. Nos trabalhos de Hunter (1991)³⁶, Hanks e Hunter (1995)³⁸ e Manning *et al.* (2002)⁴¹ é apresentada a classificação das PKs. As ePKs são classificadas nos grupos AGC, CAMK, CK1, CMGC, Other, RGC, STE, TK, e TKL e estes se subdividem em diversas famílias como, por exemplo, na família das *Cyclin-Dependent Kinases* (CDKs) pertencente ao grupo CMGC e, novamente, a família das CDKs se subdivide em subfamílias, como CDK1, CDK2, etc^{34,41}. As aPKs, que são tidas como um grupo, são classificadas em 7 famílias^{41,42,43,44}: ABC1, Alpha, Brd, PDHK, PIKK, RIO e TIF1.

Estes estudos estruturais, de alinhamentos múltiplos e árvores filogenéticas também permitiram reconhecer e descrever o papel dos motivos e resíduos conservados na reação de fosforilação e na estabilização da estrutura das PKs. Estas informações serão descritas a seguir baseadas no trabalho de Hanks e Hunter (1995)³⁸ com identificação dos resíduos na quinase PKA de *Mus musculus* representada pelo código PDB 1ATP (PDB: *Protein Data Bank*). Esta descrição se refere às ePKs e seus 12 subdomínios no estado ativo. No subdomínio I, tem-se o motivo Gly⁵⁰-Xxx-Gly⁵²-Xxx-Xxx-Gly⁵⁵, também conhecido como *G-rich loop* e *P-loop*, localizado na alça entre as fitas β 1 e 2. Esta alça auxilia no posicionamento dos fosfatos β e γ do ATP para a catálise e nas trocas entre ATP/ADP durante os ciclos catalíticos^{45,46}. Outros resíduos deste subdomínio contribuem para a formação de uma cavidade hidrofóbica que envolve o anel de adenina do ATP. O *G-rich loop* (**Figura 6**) é bastante flexível na ausência do ATP, característica que facilita a ligação de pequenas moléculas inibidoras⁴⁷. No subdomínio II está situado o resíduo catalítico Lys⁷² (**Figura 6**), pertencente à fita β 3, que também atua no ancoramento e orientação do ATP dentro do sítio catalítico, estabelecendo interações essenciais com os fosfatos α e β ⁴⁷. Além disso, este resíduo forma uma interação iônica com o grupo carboxila da cadeia lateral do Glu⁹¹ conservado (**Figura 6**) localizado na α -hélice C do subdomínio III nas quinases em estado ativo. Outros resíduos da fita β 3 contribuem para a formação da cavidade hidrofóbica. As fitas β 4 (subdomínio IV) e 5 (subdomínio V) possuem características hidrofóbicas. Alguns resíduos no subdomínio V, na alça entre a fita β 5 e a α -hélice D, são importantes para a ancoragem do ATP, formando ligações de hidrogênio tanto com o anel da adenina quanto com o da ribose. Outros resíduos na alça, próximo à fita β 5, também contribuem para a cavidade hidrofóbica. A α -hélice E do subdomínio VIA tem características hidrofóbicas e aparentemente atua mais como um suporte estrutural no domínio das quinases. As pequenas fitas β 6 e 7 intercaladas por uma alça constituem o subdomínio VIB. Neste subdomínio encontram-se alguns resíduos conservados dentre eles o Asp¹⁶⁶ e Asn¹⁷¹ pertencentes ao motivo HRDXXXXN (YRDXXXXN na PKA). O Asp¹⁶⁶ é mais um resíduo catalítico das quinases (**Figura 6**), sendo a base catalítica que abstrai o próton do grupo hidroxila dos resíduos de serina, treonina e tirosina, facilitando o ataque nucleofílico do substrato ao átomo de fósforo γ do ATP^{40,46}. Devido a este papel catalítico do Asp¹⁶⁶, a alça entre as fitas β 6 e 7 também é denominada alça catalítica (*catalytic loop*). Outros resíduos nesta alça, como a Lys¹⁶⁸, podem facilitar a transferência do fosfato, neutralizando a carga negativa do fosfato γ durante a reação. Em tirosina quinases, o resíduo com o mesmo papel da Lys¹⁶⁸ é uma arginina localizada a 4 resíduos de distância do aspartato catalítico³³. O resíduo Asn¹⁷¹ estabiliza a alça catalítica estabelecendo ligação de hidrogênio com o Asp¹⁶⁶, além de quelar

um dos íons magnésio (Mg_2) presente no sítio ativo, que realiza ponte salina com os fosfatos α e γ do ATP³³. Os íons Mg^{2+} (ou Mn^{2+}) têm um papel relevante na formação do estado de transição na reação de transferência do grupo fosforil do ATP para o substrato da quinase, devido às interações eletrostáticas estabelecidas com os fosfatos do ATP⁴⁸. As quinases possuem a capacidade de se ligarem a dois Mg^{2+} , contudo, enquanto a atividade catalítica de algumas quinases parece ser otimizada pela ligação de dois íons magnésio, para outras parece ser com a ligação de apenas um⁴⁹. Mas, ainda sim, a maioria das quinases demandam a presença dos dois íons magnésio para a atividade catalítica⁵⁰.

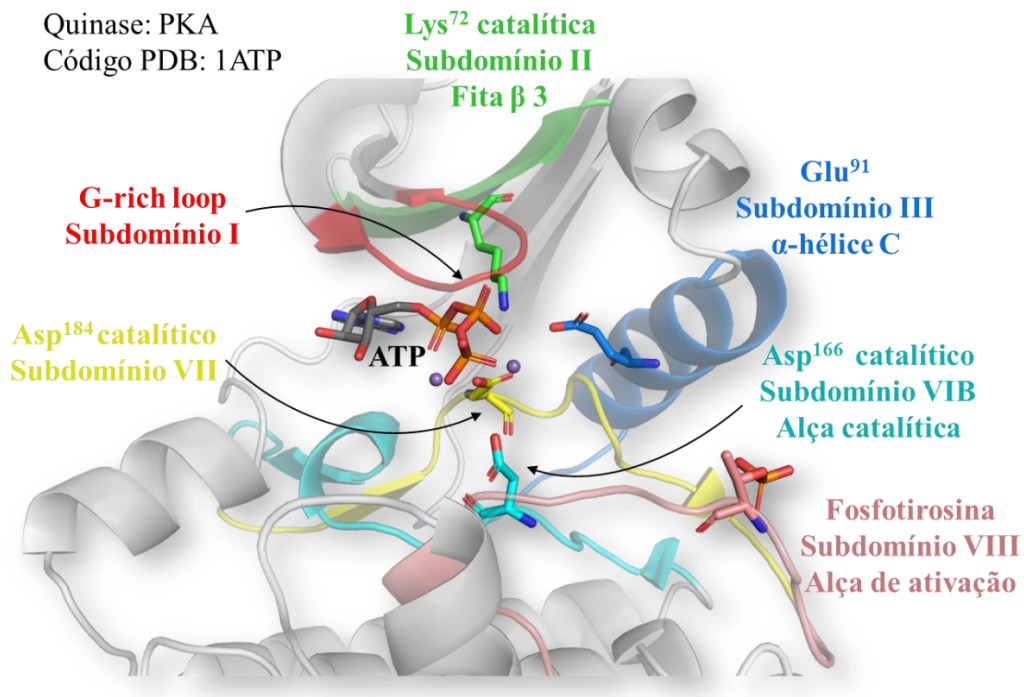


Figura 6 – Trechos dos principais subdomínios do sítio de ligação do ATP nas quinases e resíduos chave para a catálise. Além dos subdomínios, estão destacados em bastões os três resíduos catalíticos das quinases, Lys do subdomínio II, Asp do subdomínio VI e Asp do subdomínio VII, assim como o Glu conservado da α -hélice C e o resíduo fosforilado na alça de ativação que possuem papel essencial na estruturação do sítio para a catálise. As esferas em roxo representam os dois íons magnésio presentes no sítio ativo da proteína quinase PKA de *Mus musculus* (Código PDB 1ATP).

O subdomínio VII, formado pelas fitas $\beta 8$ e 9 intercaladas por uma alça, contém o motivo conservado DFG, Asp¹⁸⁴-Phe¹⁸⁵-Gly¹⁸⁶. O Asp¹⁸⁴, o terceiro resíduo catalítico (**Figura 6**), coordena tanto o íon magnésio Mg_2 , que quela os fosfatos α e γ do ATP, quanto o Mg_1 , que interage com os fosfatos β e γ , contribuindo na orientação do fosfato γ para a transferência^{33,40}.

A Gly¹⁸⁶ estabiliza a alça entre as fitas $\beta 8$ e 9 ao estabelecer ligações de hidrogênio com o Asp¹⁸⁴, e a Phe¹⁸⁵ é responsável pela orientação adequada do Asp¹⁸⁴ no sítio e pela acomodação da α -hélice C, auxiliando na interação entre a Lys⁷² catalítica e o Glu⁹¹ da α -hélice C⁴⁰. A histidina (tirosina na PKA) do motivo HRD na alça catalítica atua como suporte central ao interagir com a carbonila do Asp¹⁸⁴ e realizar contato hidrofóbico com a Phe¹⁸⁵⁴⁰. O subdomínio VIII, constituído por uma alça, está envolvido no reconhecimento dos substratos, e sua posição está relacionada com estados ativos e inativos das proteínas quinases. Parte do processo de ativação de muitas quinases acontece a partir da fosforilação de resíduo(s) de serina, treonina ou tirosina presente(s) nesta alça (Thr¹⁹⁷ na estrutura considerada) (**Figura 6**). Devido a esta característica uma parte desta alça é denominada de alça de ativação (*activation loop*)^{47,51}. Esta, juntamente com parte do subdomínio VII, desde o motivo DFG até o motivo APE do subdomínio VIII, englobando cerca de 20 a 35 resíduos, é denominada de seguimento de ativação.⁵¹ Os demais subdomínios e seus resíduos conservados auxiliam na estabilização da estrutura das PKs e na acomodação de seus substratos. Na **Figura 7** estão representadas algumas das principais interações aqui descritas presentes no sítio ativo das quinases juntamente com os resíduos e substratos envolvidos.

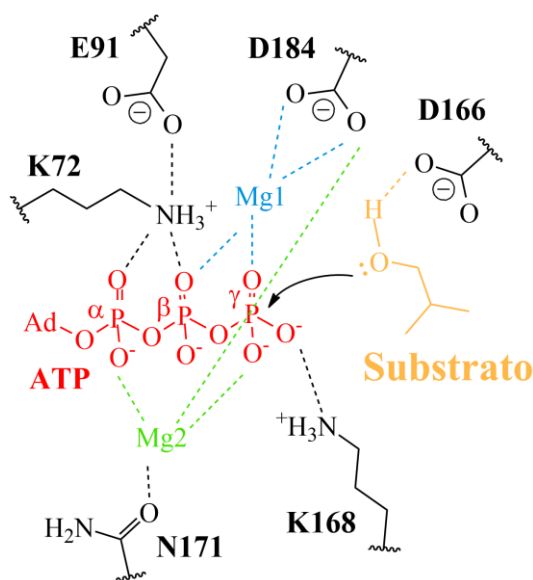


Figura 7 – Principais interações no sítio ativo das quinases. Na figura estão representadas as interações envolvendo alguns resíduos do sítio ativo das quinases, como os resíduos catalíticos K72, D166 e D184, além dos resíduos E91, N171 e K168; os substratos, o ATP (vermelho) e o substrato proteico (laranja); e os iões magnésio, Mg₁ e Mg₂. A numeração dos resíduos corresponde à da quinase PKA de código PDB 1ATP.

Fonte: adaptada de Endicott *et al.*, 2012 (doi: 10.1146/annurev-biochem-052410-090317)³³; e Adams, 2001 (doi: 10.1021/cr000230w)⁵²

A conformação ativa das proteínas quinases é bastante similar. Por outro lado, na forma inativa apresentam maior plasticidade, sendo observadas distintas conformações influenciadas pelas interações estabelecidas com domínios regulatórios e com outras proteínas⁴⁷. Em relação à fosforilação da alça de ativação, a quantidade de resíduos fosforilados para a ativação pode variar de uma quinase para outra, podendo até mesmo não ocorrer em algumas quinases⁵¹. Esta alça sofre grandes mudanças conformacionais quando a quinase alterna entre os estados ativos e inativos. Sua fosforilação estabiliza sua conformação num estado aberto (*open*) e estendido que permite o correto posicionamento do ATP e do aspartato catalítico do motivo DFG (Asp¹⁸⁴, subdomínio VII), que fica orientado para o sítio ativo^{47,51,53}. Esta conformação ativa do motivo DFG é classificada como DFG *in*. Devido à diversidade de substratos, o seguimento de ativação, mais especificamente a alça de ativação, apresenta uma diversidade conformacional e de sequência entre as quinases, mesmo na conformação ativa⁵¹. A estabilização da conformação aberta do seguimento de ativação promove um conjunto de interações na interface de ligação com os substratos proteicos. A interrupção dessa rede de interações pode comprometer a atividade das quinases⁵¹. A α -hélice C situada no subdomínio III, no lóbulo N, faz contato direto com uma porção do seguimento de ativação. Desta forma, a conformação desta α -hélice está frequentemente relacionada com a do seguimento e a do motivo DFG. A movimentação da α -hélice C durante a ativação altera a interface entre os lóbulos N e C por meio de suas rotações e modifica o acesso à fenda catalítica^{47,51}. A conformação ativa da α -hélice C é denominada α -hélice C *in*. A interação iônica do glutamato conservado desta α -hélice com a lisina catalítica do subdomínio II no estado ativo das quinases auxilia na estabilização das interações entre a lisina e os fosfatos α e β do ATP e no correto posicionamento do nucleotídeo no sítio catalítico^{47,53}. Em suma, os 3 principais elementos regulatórios no domínio das quinases são a posição da α -hélice C (posição *in*, ativa, *out* inativa), posição do motivo DFG (posição *in*, ativa, *out* inativa) e posição do seguimento de ativação (aberto/*open*, ativo, fechado/*closed* inativo)⁴⁶. Toda essa movimentação da proteína, ajustando a posição das estruturas secundárias para uma conformação ativa com a otimização das interações e o correto posicionamento dos resíduos importantes para a catálise enzimática, propagam este ajuste ao longo dos dois lóbulos, formando as estruturas chamadas de *spines*. Estas estruturas são caracterizadas pelo alinhamento de diversos resíduos hidrofóbicos conservados ao longo de dois eixos, estabelecendo uma arquitetura interna coesa e flexível que conecta os dois lóbulos^{40,45,53}. Uma dessas estruturas é chamada de *Catalytic spine* (C-*spine*) (**Figura 8**) que se ancora na α -hélice F e se prolonga até o sítio de ligação do ATP⁴⁵. Os resíduos pertencentes a esta *spine* na porção do lóbulo C são a Leu²²⁷ e a Met²³¹ da α -hélice F, a Met¹²⁸ da α -hélice D,

a Leu¹⁷², Leu¹⁷³ e Ile¹⁷⁴ da fita $\beta 7$, e na porção do lóbulo N são a Val⁵⁷ da fita $\beta 2$ e a Ala⁷⁰ da fita $\beta 3$ ⁴⁵. O anel de adenina do ATP também integra a *C-spine*⁴⁶. O outro eixo, também ancorado na α -hélice F, recebe o nome de *Regulatory spine* (*R-spine*) (**Figura 8**) por sua formação depender do estado de ativação das quinases, estando intimamente relacionada com a conformação do seguimento de ativação^{40,45}. No estado inativo os 4 resíduos hidrofóbicos que a compõem, a Leu⁹⁵ da α -hélice C, a Leu¹⁰⁶ da fita $\beta 4$, a Tyr¹⁶⁴ do motivo YRD (HRD) da alça catalítica e a Phe¹⁸⁵ do motivo DFG (seguimento de ativação), encontram-se desordenados^{40,45}. A *R-spine* participa do posicionamento do substrato proteico enquanto a *C-spine* está envolvida no posicionamento do ATP e promoção da catálise⁴⁶.

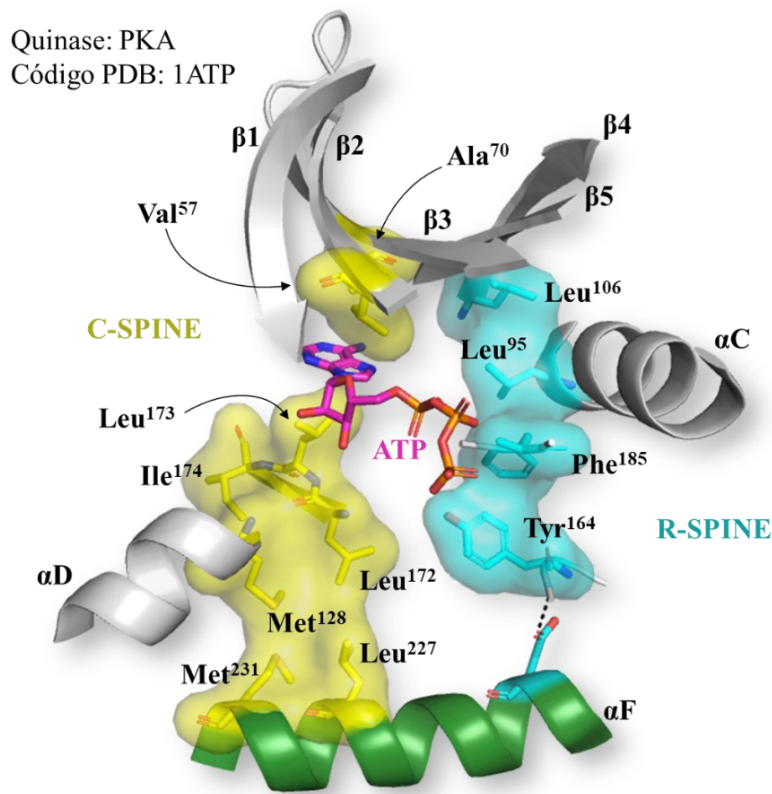


Figura 8 - Eixos hidrofóbicos conservados nas proteínas quinases, *C-spine* e *R-spine*. Localização da estrutura espacial da *C-spine* (amarelo) e da *R-spine* (azul) e identificação dos respectivos resíduos que as compõem na quinase PKA (Código PDB:1ATP). (Parte da estrutura tridimensional foi omitida para facilitar a visualização da região analisada).

O controle da ativação das proteínas quinases muitas vezes depende também de outros mecanismos. A presença de domínios regulatórios e a associação de moléculas reguladoras e subunidades das proteínas pode sustentar um estado inativo em algumas quinases, enquanto em

outras pode ser essencial para uma completa ativação^{47,51}. Muitas das moléculas reguladoras da atividade das quinases se ligam na região da α -hélice C, alterando sua conformação e de outras estruturas próximas. Esta movimentação estrutural pode comprometer ou promover a formação da interação iônica entre a lisina catalítica e o glutamato da α -hélice C e assim ativar ou impedir a ativação da quinase por uma regulação alostérica⁴⁷. A presença de inserções e extensões no domínio das quinases e variações nas sequências também são descritas como aspectos importante no processo de ativação, influenciando, por exemplo, na correta conformação do seguimento de ativação⁵¹. Algumas PKs adotam a forma inativa por um mecanismo autoinibitório causado por interações de segmentos de suas estruturas na região do sítio ativo ou vizinhanças, gerando um impedimento estérico ou efeito alostérico negativo⁴⁷. A conjugação destes mecanismos com a fosforilação da alça de ativação é necessária para a ativação completa de muitas quinases. Porém, para outras, o estado ativo pode ser alcançado apenas com algum destes modos regulatórios, existindo quinases que não demandam a fosforilação da alça⁵¹. Na **Figura 9** está esboçado de forma simplificada as movimentações estruturais chaves na ativação das quinases, como a do lóbulo N, da α -hélice C e seu resíduo de glutamato conservado, da lisina catalítica e da alça de ativação (seguimento de ativação) para uma quinase CDK regulada positivamente por uma molécula de ciclina.

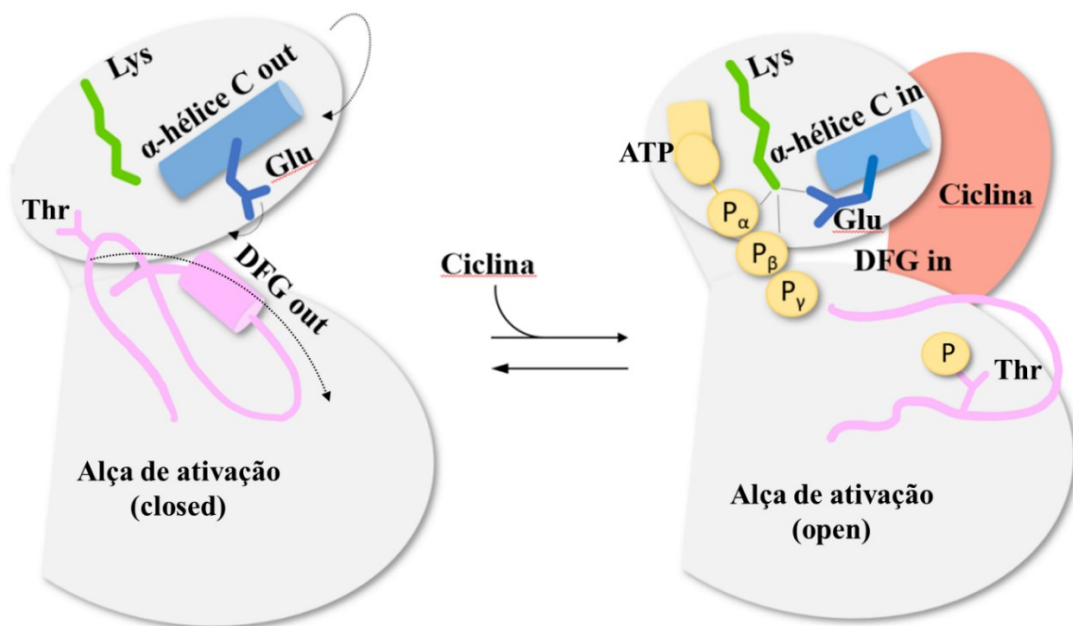


Figura 9 - Esboço das principais estruturas e mudanças conformacionais envolvidas na ativação das quinases. Está representada a movimentação do lóbulo N, da α -hélice C e do motivo DFG de uma conformação *out* para uma conformação *in*, e da alça de ativação de uma conformação fechada (*closed*) para uma aberta (*open*). Também estão identificados alguns dos principais resíduos envolvidos na

ativação das quinases, como a Lys catalítica da fita $\beta 3$ e o glutamato da α -hélice C, que estabelecem uma ponte salina nas quinases ativas, além do ATP, do resíduo da alça de ativação cuja fosforilação contribui para o processo de ativação da quinase, e, por fim, da molécula ativadora (ciclina para as quinases da família CDK).

Fonte: adaptada de Huse e Kuriyan, 2002 (doi: 10.1016/S0092-8674(02)00741-9)⁴⁷

A partir deste conhecimento estrutural sobre as PKs é possível discutir com mais detalhes o seu mecanismo de reação. Para a catálise, as proteínas quinases precisam de pelo menos um íon Mg^{2+} ou Mn^{2+} presente no sítio ativo, como já mencionado⁵⁴. As serina/treonina quinases parecem seguir um mecanismo reacional de complexo ternário, em que ocorre uma reação direta entre os substratos, o ATP e a proteína a ser fosforilada⁵⁴. Ambos os substratos se ligam no sítio ativo, formando o complexo ternário, no qual ocorre a transferência do grupo fosforil do ATP para o substrato proteico, liberando o ADP e a proteína fosforilada. Porém, para uma quinase atípica foi reportado o mecanismo de *ping-pong*, em que ocorre a formação de um intermediário covalente enzimático, uma fosfoquinase⁵⁴. Durante a reação, num primeiro momento, ocorre a transferência do grupo fosfato do ATP para um resíduo da quinase, para somente depois ocorrer a transferência para o substrato proteico.

Definir o estado de transição da transferência do grupo fosfato é uma questão chave para compreender o mecanismo reacional das quinases. Na literatura são descritos dois estados de transição, um chamado de dissociativo e outro de associativo (**Figura 10**)^{33,54}. No primeiro, durante a reação ocorre o completo rompimento da ligação covalente entre o fósforo e o grupo abandonador, para posteriormente ocorrer a ligação com o nucleófilo, formando um intermediário *metaphosphate-like*⁵⁴. No outro estado de transição, o associativo, o nucleófilo forma uma ligação com o fósforo à medida que este se desliga do grupo abandonador, ocorrendo a formação de um intermediário fosforano pentavalente^{33,54}. Wang e Cole (2014)⁵⁴ pontuam que analisar o estado de transição de enzimas é mais desafiador devido às suas dimensões, diversidade de proteínas quinases, associação de moléculas reguladoras, fosforilações, influência das condições reacionais adotadas nos experimentos e limitações dos métodos disponíveis. Porém, segundo estes autores, as evidências a partir de experimentos em solução apoiam fortemente um estado de transição dissociativo para as reações das proteína quinase que foram estudadas e, dada a conservação das características das quinases, é plausível extrapolar que o mecanismo dissociativo ocorra para grande parte destas proteínas. Nesta lógica, nas PKs ocorreria o rompimento da ligação covalente entre o fósforo γ e a molécula de ATP, para posteriormente ocorrer a ligação com o nucleófilo, a hidroxila alcóolica ou fenólica do substrato

proteico, formando um intermediário *metaphosphate-like*. Na **Figura 10** está representado um diagrama esquemático do mecanismo catalítico das PKs com destaque das principais interações presentes nas quinases ativas que contribuem para a correta orientação dos substratos e prosseguimento da reação.

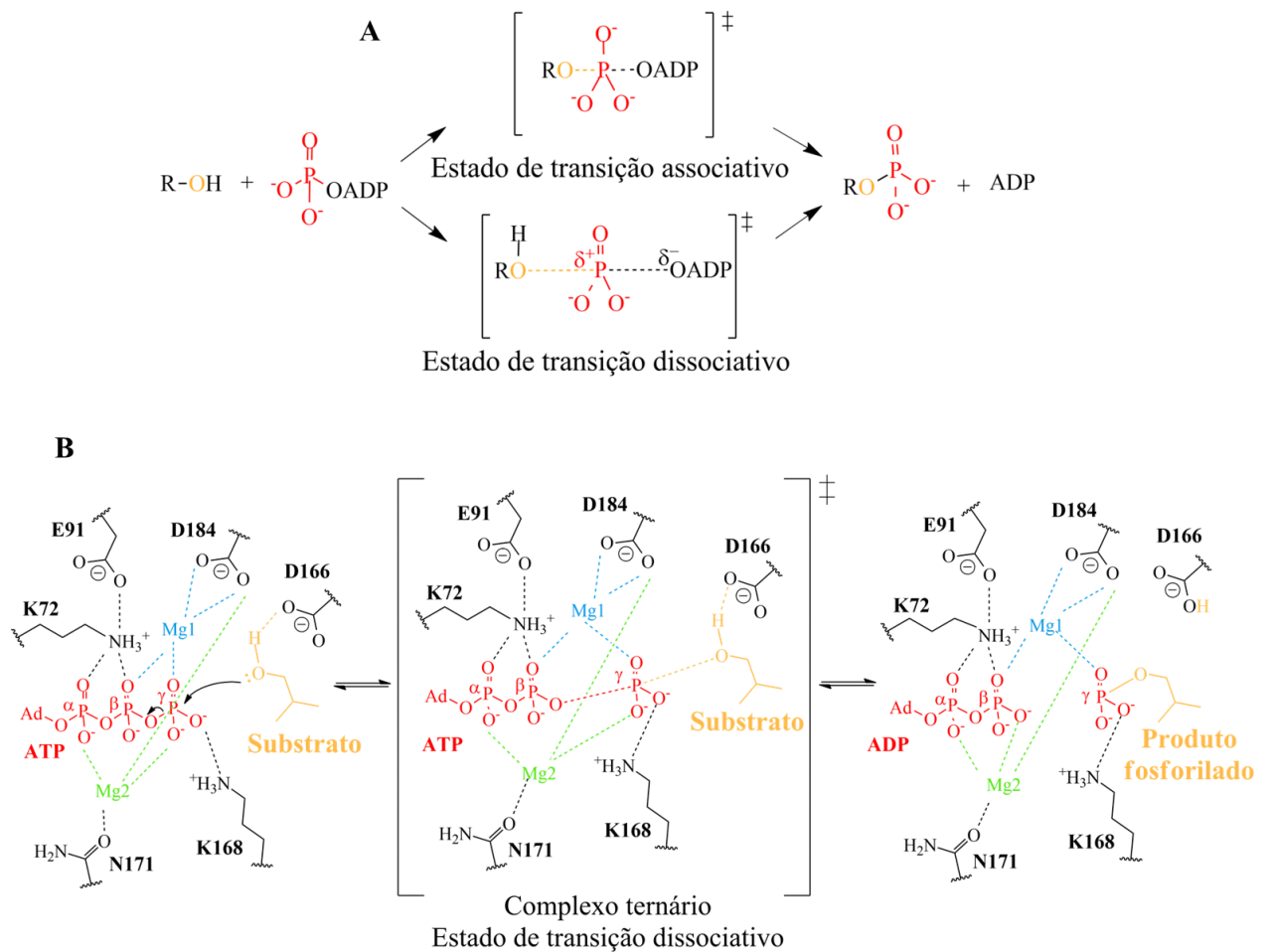


Figura 10 - Diagrama esquemático do mecanismo de reação das proteínas quinases. **A.** Esquema dos estados de transição associativo e dissociativo nas reações de transferência do grupo fosforil. **B.** Detalhamento esquemático da reação de fosforilação das proteínas quinases com destaque dos principais resíduos e interações envolvidos na reação no sítio ativo, segundo um mecanismo com formação de complexo ternário e estado de transição dissociativo.

Fonte: Painel A - adaptada de Wang e Cole, 2014 (doi: 10.1016/B978-0-12-397918-6.00001-X)⁵⁴;

Painel B - adaptada de Endicot *et al.*, 2012 (doi: 10.1146/annurev-biochem-052410-090317)³³; e Adams, 2001 (doi: 10.1021/cr000230w)⁵²

1.4 INIBIDORES COVALENTES COMO ESTRATÉGIA DE AÇÃO EM PROTEÍNAS QUINASES

As proteínas quinases correspondem a cerca de 2% do proteoma de grande parte dos organismos eucariotos, estando envolvidas na fosforilação de mais de 30% das proteínas celulares⁵⁵. Dada esta significativa relevância das PKs, disfunções envolvendo estas proteínas e suas atividades catalítica, como mutações, superexpressão, translocação e desregulação, são causas de uma diversidade de doenças associadas a distúrbios de desenvolvimento e metabólicos^{55,56,50}. Em humanos, alterações envolvendo quinases estão relacionadas a doenças inflamatórias ou de resposta proliferativa, como o câncer, artrite reumatóide, desordens cardiovasculares e neurológicas, doenças autoimune, dentre outras^{55,57}. Uma abordagem para controlar e tratar estas doenças consiste na inibição das proteínas quinases envolvidas nestes distúrbios.

Grande parte dos inibidores enzimáticos, não somente das quinases, se ligam de forma reversível a seus alvos terapêuticos, estabelecendo interações intermoleculares¹³. Contudo, existem também inibidores que apresentam um mecanismo inibitório covalente, com formação de uma ligação covalente entre o inibidor e um resíduo da proteína no sítio de ligação¹³. Nas discussões sobre os sítios de ligação dos inibidores nas quinases é apresentada uma divisão espacial do sítio catalítico localizado entre os lóbulos N e C (**Figura 11**). Segundo Roskoski (2021), o sítio catalítico é dividido em duas regiões, a região frontal, também chamada de fenda frontal, e a região traseira, ou cavidade hidrofóbica II, que se subdivide nas cavidades área de portão (*gate area*) e fenda traseira (*back cleft*)³⁷. A região frontal inclui os três últimos resíduos da fita $\beta 1$, o *G-rich loop (P-Loop)*, os quatro primeiros resíduos da fita $\beta 2$, os resíduos da alça localizada após a fita $\beta 5$, chamados de resíduos de dobradiça (*hinge residues*), o seguimento que conecta os resíduos de dobradiça à α -hélice D, a α -hélice D, a alça catalítica e a fita $\beta 7$ ^{37,50}. A área de portão é composta pelos três últimos resíduos da fita $\beta 3$, os dois primeiros resíduos da alça entre a fita $\beta 3$ e a α -hélice C e pelo primeiro resíduo antes do seguimento de ativação juntamente com os próximos quatro resíduos deste seguimento, incluindo o motivo DFG^{37,50}. O nome desta região se refere ao seu papel no controle do acesso à fenda traseira⁴⁶. A fenda traseira é constituída pela porção média da α -hélice C, incluindo a alça que conecta esta α -hélice à fita $\beta 4$, pela fita $\beta 4$ e porções da fita $\beta 5$, pelos dois primeiros e quinto resíduos da α -hélice E e pelos dois resíduos anteriores à alça catalítica^{37,50}.

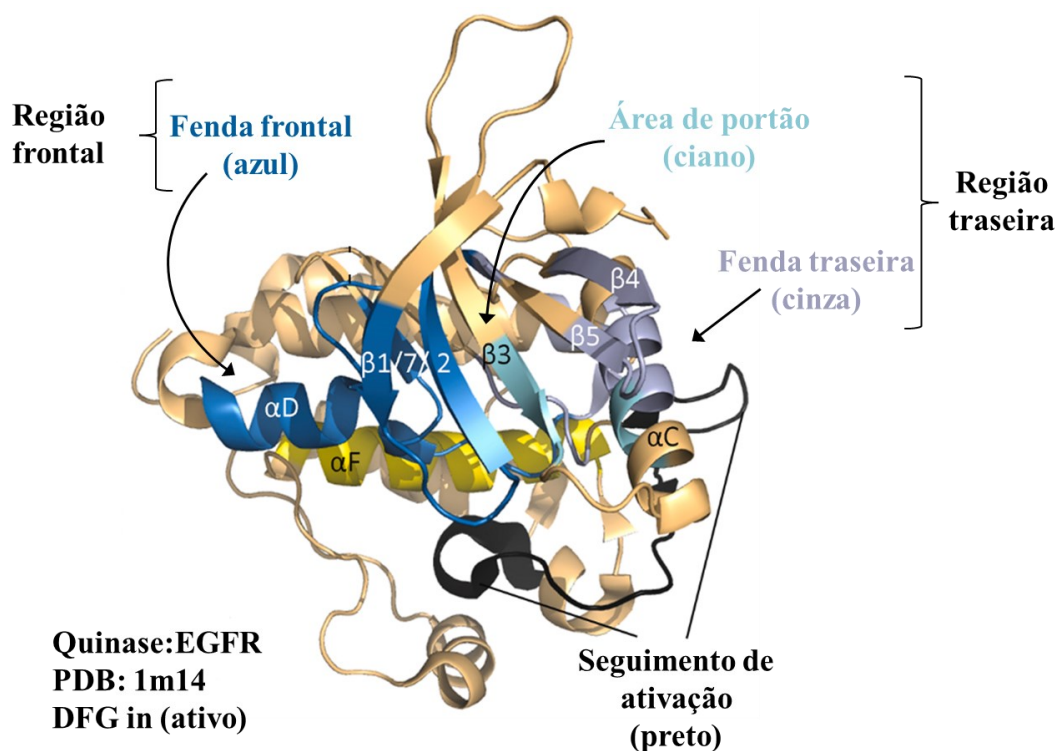


Figura 11 - Localização das regiões e cavidades do sítio ativo das quinases. Estão identificadas na estrutura da quinase EGFR de código PDB 1m14 as regiões e cavidades que integram o sítio ativo das quinases e que são sítios de ligação dos inibidores. Na figura estão destacadas a fenda frontal (azul), a área de portão (ciano), a fenda traseira (cinza), além do seguimento de ativação (preto).

Fonte: adaptada de Roskoski, 2021 (doi: 10.1016/j.phrs.2021.105422)³⁷

A classificação dos tipos de inibidores de quinases considera seu local de ligação e as características estruturais das proteínas quanto ao seu estado de ativação. Estes inibidores são classificados em 7 tipos, de I a VI, incluindo a variação II/2, com subdivisões dentro dos tipos II/2 e II. Os inibidores do tipo I são aqueles que se ligam no sítio de ligação do ATP e redondezas, geralmente na fenda frontal, em quinases com conformação ativa^{37,46,50}. Os do tipo II também se ligam no sítio de ligação do ATP e redondezas (fenda frontal), porém em quinases inativas, com DFG *out*^{46,50}. Inibidores dos tipos III e IV se ligam a sítios alostéricos. Os do tipo III se ligam em sítios alostéricos próximos ao sítio de ligação do ATP e os do tipo IV em locais fora da fenda catalítica, afastados dos sítios de ligação dos substratos^{37,46,50}. Aqueles inibidores que se ligam abrangendo duas regiões do domínio de quinases são classificados como do tipo V ou inibidores bivalentes⁵⁰. Por fim, os inibidores do tipo VI são aqueles que formam ligação covalente com suas enzimas alvo⁵⁰. Em relação às divisões dos inibidores do tipo I, a variação II/2 consiste nos compostos que inibem quinases inativas, mas cujo DFG está na conformação *in*. Estes inibidores se subdividem nas variações A e B, em que

os classificados como A são aqueles inibidores que ocupam a fenda frontal, a área de portão e se estendem até a fenda traseira, e os inibidores I1/2 B ocupam a fenda frontal e a área de portão, mas não se estendem até a fenda traseira^{37,46,50}. A divisão dos inibidores do tipo II também utiliza a classificação A e B, seguindo a explicação dos inibidores do tipo I1/2. Estas e outras informações complementares destes 7 tipos de inibidores estão apresentadas na **Tabela 2**.

Atualmente, cerca de um quarto dos esforços na descoberta de novos fármacos estão direcionadas para as quinases³⁷. Até 16 de abril de 2022 já encontravam-se aprovadas pela *US Food and Drug Administration* (FDA) 71 pequenas moléculas inibidoras de quinases, que atuam em cerca de 20 diferentes PKs (<http://www.brimr.org/PKI/PKIs.htm>). Grande parte destes inibidores, 39 no total, atuam em receptores tirosina quinases, 16 atuam em tirosina quinases não receptoras, 12 em serina/treonina quinases e 4 em quinases de especificidade dual. Além disso, até esta mesma data, cerca de 230 novos candidatos a inibidores de quinases encontravam-se em ensaios clínicos no mundo (<https://www.icoa.fr/pkidb/>)^{37,58}. Dentre os inibidores aprovados pela FDA, 7 apresentam um mecanismo de inibição covalente: afatinib, dacomitinib, neratinib, osimertinib, acalabrutinib, ibrutinib e zanubrutinib³⁷. Estes fármacos têm como alvos terapêuticos majoritariamente quinases envolvidas em doenças oncológicas. O ibrutinib, além de ser empregado no tratamento de câncer, também é utilizado no tratamento da doença do enxerto contra o hospedeiro, condição em que após um transplante ocorre uma reação imunológica do material transplantado (enxerto) contra o organismo que recebeu o transplante (hospedeiro)³⁷. Na **Tabela 3** estão apresentadas informações complementares referentes a estes 7 inibidores covalentes de quinases aprovados pela FDA.

Tabela 2 - Tipos de inibidores das proteínas quinases

Propriedades	Tipos de inibidores						
	Tipo I	Tipo I1/2 (A/B)	Tipo II (A/B)	Tipo III	Tipo IV	Tipo V	Tipo VI
Região de ligação	Sítio de ligação do ATP (quinase ativa)	Sítio de ligação do ATP (quinase inativa, DFG <i>in</i>)	Sítio de ligação do ATP (quinase inativa, DFG <i>out</i>)	Sítio alostérico próximo ao sítio de ligação do ATP	Sítio alostérico fora da fenda catalítica	Inibidores bivalentes, abrangendo duas regiões do domínio	Forma dutos covalentes com a enzima alvo
Estende até a fenda traseira	Não	A (sim)/B (não)	A (sim)/B (não)	Sim	Não	Variável	Variável
Conformação DFG	In	In	Out	Variável	Variável	Variável	Variável
Conformação Segui. de ativ.	Out	Variável	Variável	Variável	Variável	Variável	Variável
Conformação α -hélice C	In	Variável	Variável	Out	Variável	Variável	Variável
Conformação S- <i>spine</i>	linear	Distorcida	Usualmente distorcida	Distorcida	Variável	Variável	Variável
ATP-competitivo	Sim	Sim	Sim	Não	Não	Variável	Não
Inibição reversível	Sim	Sim	Sim	Sim	Sim	Sim	Usualmente não

Fonte: adaptada de Roskoski, 2016 (doi: 10.1016/j.phrs.2015.10.021)⁴⁶

Tabela 3 - Inibidores covalentes de quinases humanas aprovados pela FDA e suas indicações terapêuticas

Fármaco/ Nome comercial	Ano de aprovação	Alvos primários	Indicação terapêutica
Afatinib/ Tovok	2013	ErbB1/2/4	Câncer de pulmão de células não pequenas
Dacomitinib/ Visimpro	2018	ErbB1/2/4	Câncer de pulmão de células não pequenas mutantes de EGFR
Neratinib/ Nerlynx	2017	HER2/EGFR	Câncer de mama HER2-positivo
Osimertinib/ Tagrisso	2015	EGFR T970M/ ErbB3	Câncer de pulmão de células não pequenas
Acalabrutinib/ Calquence	2017	BTK	Linfomas de células do manto leucemias linfocíticas crônicas Pequenas leucemias linfocíticas
Ibrutinib/ Imbruvica	2013	BTK	Linfomas de células do manto leucemias linfocíticas crônicas Pequenas leucemias linfocíticas Doença do enxerto contra o hospedeiro Macroglobulinemia de Waldenstrom
Zanabrutinib/ Brukinsa	2019	BTK	Linfomas de células do manto

ErbB/HER: linhagem de proteínas que pertencem à família de receptores do fator de crescimento epidérmico humano

EGFR: *epidermal growth factor receptor* (receptores do fator de crescimento epidérmico humano)

BTK: *Bruton tyrosine kinase* (tirosina quinase de Bruton)

Fonte: adaptada de Roskoski, 2021 (doi: 10.1016/j.phrs.2021.105422)³⁷

A estrutura dos inibidores covalentes é constituída por uma região que realiza interações intermoleculares com o alvo, sendo importante para a complementariedade com o sítio de ligação, e por uma região eletrofílica que estabelece a ligação covalente com o alvo⁴⁶. Dentre os grupos eletrofílicos, a acrilamida (**Figura 12**, destacado em vermelho) é o grupo mais comum nos inibidores covalentes de quinases e possui capacidade de se ligar em resíduos reativos do alvo em diferentes localizações⁵⁷. A acrilamida, seus derivados e outros grupos eletrofílicos apresentam em suas estruturas carbono insaturado ligado a uma carbonila ou carboxila e são chamados de aceptores de Michael⁵⁹. A variedade de grupos eletrofílicos aceptores de Michael, dentre outros tipos, como tióis, epóxidos, alcinos, etc (**Figura 12**), permitiram direcionar a inibição para determinadas regiões do alvo e melhorar a seletividade⁵⁷. A inibição covalente é estabelecida com um resíduo reativo da proteína com característica nucleofílica. Os potenciais nucleófilos são os resíduos de cisteína, lisina, aspartato, serina, tirosina e histidina^{13,57}. Dentre estes, cisteínas pouco conservadas na região do sítio de ligação

do ATP estão frequentemente envolvidas na ligação covalente com os inibidores^{13,57}. As demais opções de nucleófilos apresentam maior conservação e abundância quando comparados com as cisteínas, sendo menos interessantes para a abordagem de inibição covalente^{13,57}. A inibição covalente se processa em duas etapas (**Figura 12**). A primeira envolve a formação do complexo não covalente, com associação reversível do inibidor com seu alvo, por meio de interações intermoleculares, orientando o grupo eletrofílico do inibidor próximo ao resíduo nucleofílico do alvo. A segunda etapa consiste no estabelecimento da ligação covalente entre ambos, formando o complexo covalente.³⁷ No caso de aceptores de Michael, a reação que ocorre nesta segunda etapa é conhecida como adição de Michael^{37,57}. Nesta reação, o grupo tiol de uma cisteína (ou outro resíduo nucleofílico) ataca o grupo aceptor de Michael, formando a ligação covalente e o produto final tioéter (**Figura 12**)³⁷. Dependendo do grupo eletrofílico, a inibição covalente pode ser reversível ou irreversível. A associação de grupos retiradores de densidade eletrônica à porção eletrofílica dos inibidores é uma maneira de modular a reatividade dos compostos e obter inibidores covalentes reversíveis (**Figura 12**)¹³. Um exemplo é a associação do grupo funcional -CN à acrilamida, formando a cianoacrilamida, grupo eletrofílico mais comum entre os inibidores covalentes reversíveis^{13,57}.

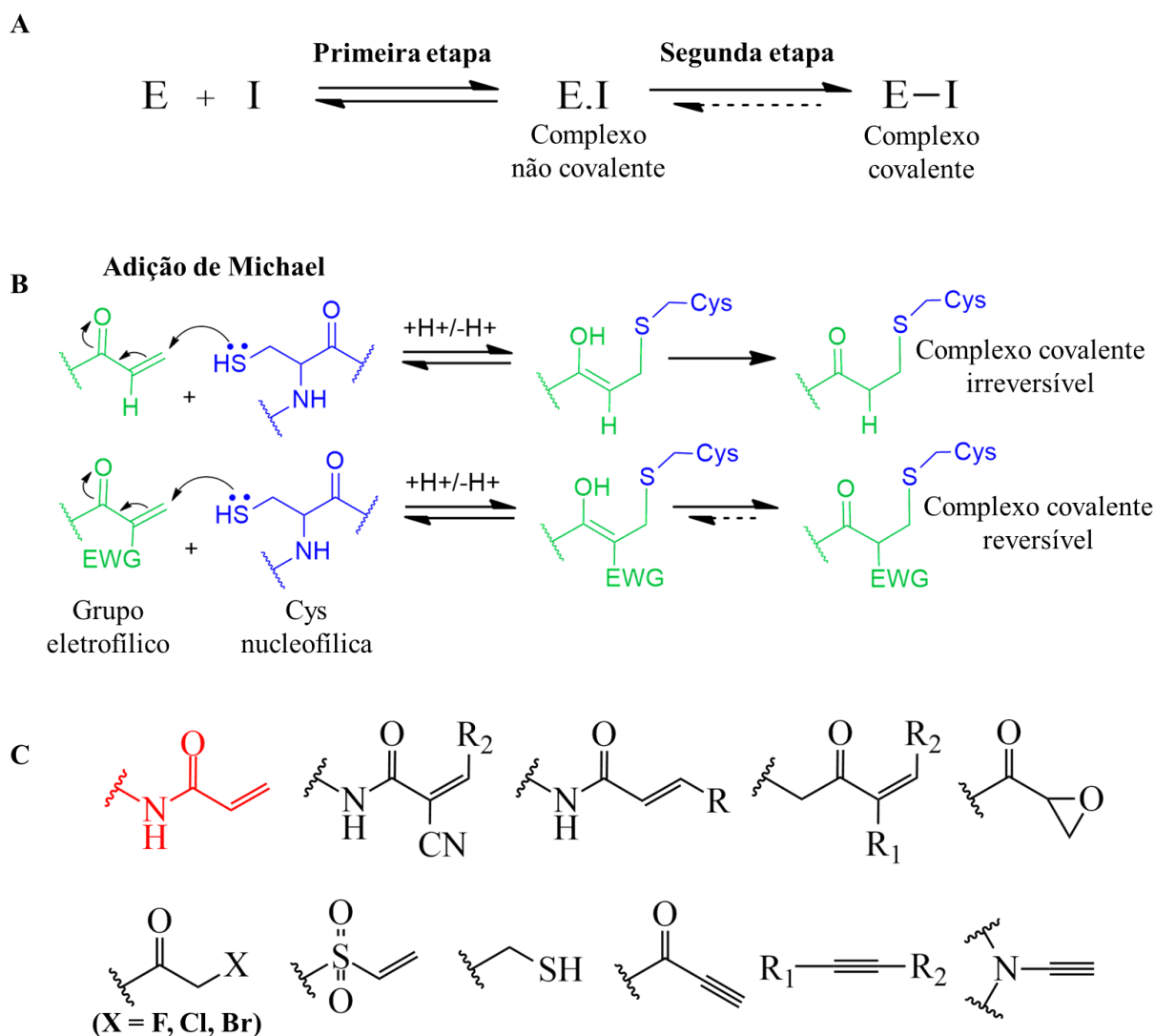


Figura 12 - Esquema das etapas da inibição covalente e da reação de adição de Michael e estruturas dos principais grupos eletrofílicos dos inibidores covalentes que se ligam a cisteínas. **A.** Esquema da inibição covalente, envolvendo a formação do complexo não covalente na primeira etapa e a formação de ligação covalente na segunda etapa, resultando em um complexo covalente. A seta pontilhada, indicando a volta do complexo covalente para o complexo não covalente, ocorre apenas para inibidores covalentes reversíveis. **B.** Esquema da reação de adição de Michael entre o grupo eletrofílico de um inibidor e um resíduo nucleofílico de cisteína. Na primeira equação está representada a reação de adição de Michael referente a um inibidor covalente irreversível e, na segunda, referente a um inibidor covalente reversível. (EWG - *Electron-Withdrawing Groups*, grupos retiradores de densidade eletrônica). **C.** Estrutura dos principais grupos eletrofílicos dos inibidores covalentes que se ligam a cisteínas.

Fonte: Figura B – adaptada de Abdeldayem *et al.*, 2020 (doi: 10.1039/c9cs00720b)¹³; Figura C – adaptada de Zhao e Bourne, 2018 (doi: 10.1016/j.drudis.2018.01.035)⁵⁷

Uma série de vantagens no uso dos inibidores covalentes são descritas. Uma delas diz respeito à potência do fármaco. A formação de uma ligação covalente com o alvo contribui significativamente no ganho de potência sem que para isso seja necessário aumentar o tamanho

da molécula para ampliar as interações intermoleculares com o alvo¹³. Os inibidores não covalentes têm sua potência associada às interações intermoleculares e, apesar de ser possível aumentar o tamanho da molécula para incrementar as interações, esta alternativa impacta em parâmetros farmacocinéticos, alterando, por exemplo, a estabilidade metabólica do inibidor e o perfil de permeabilidade¹³. Uma outra vantagem no emprego dos inibidores covalentes é o prolongamento da duração de sua ação no alvo em comparação com os inibidores não covalentes. Esta característica, somada à elevada potência, permite administrar doses menores do inibidor e com menores frequências, reduzindo efeitos adversos¹³. O desenvolvimento de resistência a fármacos devido às mutações nos alvos terapêuticos é um outro ponto que compromete a potência dos inibidores. Neste contexto, Abdeldayem *et al.* (2020)¹³ mencionam que estudos em inibidores covalentes validaram a habilidade destes de evadir aos eventos de mutação, mantendo suas potências contra os alvos mutados. Enquanto a mutação no sítio ativo do alvo pode impedir a ocorrência de alguma interação importante entre o alvo e o inibidor, a ligação covalente estabelecida ainda pode sustentar a capacidade inibitória, desde que não tenha ocorrido mutação no resíduo reativo do sítio¹³. Zhao e Bourne (2018)⁵⁷ mencionam que no tratamento de cânceres estes inibidores têm se mostrado mais promissores do que inibidores não covalentes de quinases, devido à ocorrência de resistência aos fármacos não covalentes. Existem fármacos bastante utilizados e seguros cujo mecanismo de inibição é covalente. A aspirina, os inibidores de bomba de prótons como os prazóis, empregados em distúrbios gastrointestinais, medicações para o tratamento da doença de Parkinson, como selegiline e rasagiline, e medicações utilizadas no tratamento de mieloma, como carfilzomib são alguns exemplos de sucesso de inibidores que atuam por inibição covalente em alvos sem ser quinases³⁷.

As desvantagens do uso dos inibidores covalentes envolvem o controle da potência do inibidor, a produção de metabólitos que podem causar respostas imunogênicas, a maior probabilidade de ligação a outras moléculas não alvo, e a limitação de aplicar esta estratégia de inibição apenas para quinases que tenham resíduos reativos e que estes estejam em posições acessíveis e de interesse^{13,57}. Apesar de ser um objetivo produzir um fármaco que apresente boa potência contra seu alvo, em situações em que é necessária uma regulação negativa transitória ou parcial da atividade do alvo, a inibição covalente pode ser problemática¹³. A preocupação a respeito de uma possível resposta imunogênica se deve à possibilidade de na metabolização do inibidor covalente serem gerados eletrófilos reativos promíscuos que tenham a capacidade de modificar outras proteínas¹³. Em relação à inibição de moléculas não alvo, esta é uma preocupação que

faz parte do desenvolvimento de qualquer inibidor. Mas, no caso da inibição covalente, devido à maior reatividade dos inibidores, há uma maior probabilidade de proteínas não alvo serem afetadas, podendo trazer complicações e aumentar a citotoxicidade¹³.

O desenvolvimento de inibidores covalentes reversíveis tem sido uma alternativa para contornar ou minimizar estes possíveis problemas associados aos inibidores covalentes. O controle da reatividade dos grupos eletrofilicos dos inibidores covalentes e da complementariedade adequada das interações intermoleculares reduz a habilidade destes de se ligarem de forma definitiva ou prolongada a moléculas não alvo^{13,60}. Desta forma, esta variação de inibidor possibilita minimizar a citotoxicidades associadas à ligação a outras proteínas, contribuindo para o ganho de seletividade, além de possibilitar gerenciar a atividade inibitória transitória ou parcial contra alvos que demandem este tipo de regulação^{13,60}. Somado a isto, evitar a construção de inibidores com grupos eletrofilicos muito reativos e não específicos, buscando utilizar grupos de reatividade moderada a fraca, e o uso de doses menores ou iguais a 10 mg kg^{-1} permitem contornar ou controlar parte das respostas indesejadas ao uso dos inibidores covalentes¹³. Apesar do design de pequenas moléculas inibidoras para quinases ser um desafio, em função da similaridade do sítio do ATP entre as quinases, principal sítio explorado para inibição, o avanço do conhecimento sobre a estrutura destas proteínas tem possibilitado a descoberta de pequenas diferenças no sítio e na sua vizinhança, conferindo novas perspectivas nesta abordagem⁵⁷. Este conhecimento, juntamente com o planejamento racional de compostos covalentes permitem direcionar a inibição para os alvos de maneira seletiva e eficiente, administrando a eficácia e a toxicidade^{13,57}. Na **Figura 13** a seguir estão representadas as estruturas dos 7 inibidores covalentes de quinases aprovados pela FDA com destaque das regiões eletrofílicas constituídas de derivados da acrilamida envolvidas na reação de adição de Michael com os alvos.

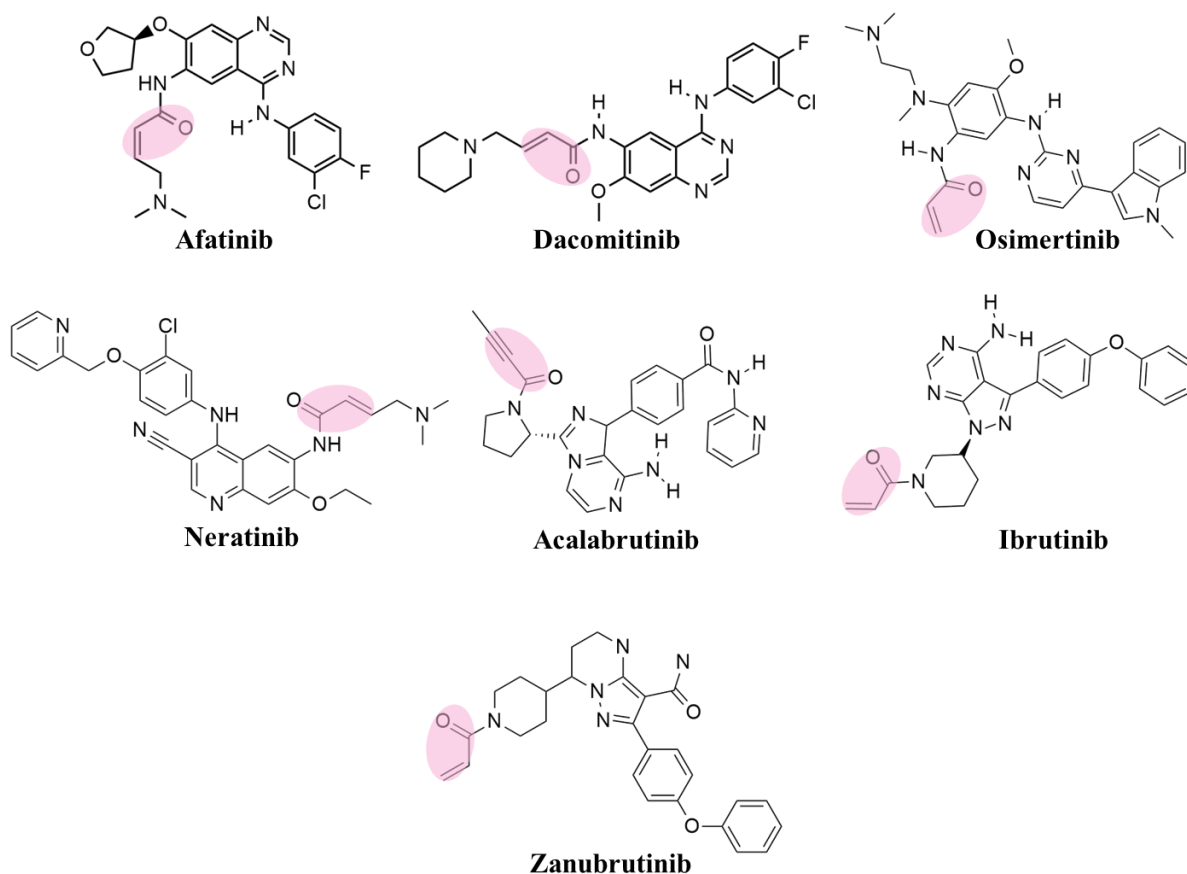


Figura 13 - Estrutura dos 7 inibidores covalentes de quinases aprovados pela FDA com destaque da região eletrofílica envolvida na reação de adição de Michael com o alvo.

Fonte: adaptada de Roskoski, 2021 (doi: 10.1016/j.phrs.2021.105422)³⁷

Para que quinases sejam candidatas a alvos terapêuticos para uma inibição covalente é necessária a presença de resíduos reativos em posições estratégicas no sítio de ligação do inibidor. Ao longo dos anos foram mapeadas posições no sítio ativo das ePKs nas quais a presença de cisteínas permite a formação de ligação covalente com inibidores. Zhao *et al.* (2017)⁶¹ combinaram os métodos *Function-site Interaction FingerPrint* (Fs-IFP) e *Density Functional Theory* (DFT) para obter os modos de ligação nos complexos quinase-ligante, além de obter detalhes das interações de cada átomo na interface do sítio de ligação e quantificar o efeito dos microambientes na reatividade das cisteínas além de avaliar aspectos estéricos⁶¹. Zhao *et al.* analisaram 2774 estruturas de complexos de quinases humanas com inibidores diversos, covalentes e não covalentes, identificando 1599 estruturas, referentes a 169 quinases distintas, que possuíam pelo menos uma cisteína localizada nos seus respectivos sítios de ligação. Neste processo mapearam 63 posições, destacando 17 (**Figura 14**) nas quais havia pelo

menos um contato (interação intermolecular ou ligação covalente) entre uma cisteína e o inibidor presentes nas estruturas. Outro mapeamento foi realizado por Chaikuad *et al.* (2018)⁶², baseado na orientação espacial das cadeias laterais das cisteínas na região do sítio e proximidades. Foram mapeadas 18 posições com cisteínas em orientações favoráveis para uma inibição covalente, considerando a conformação ativa das quinases (**Figura 14**).

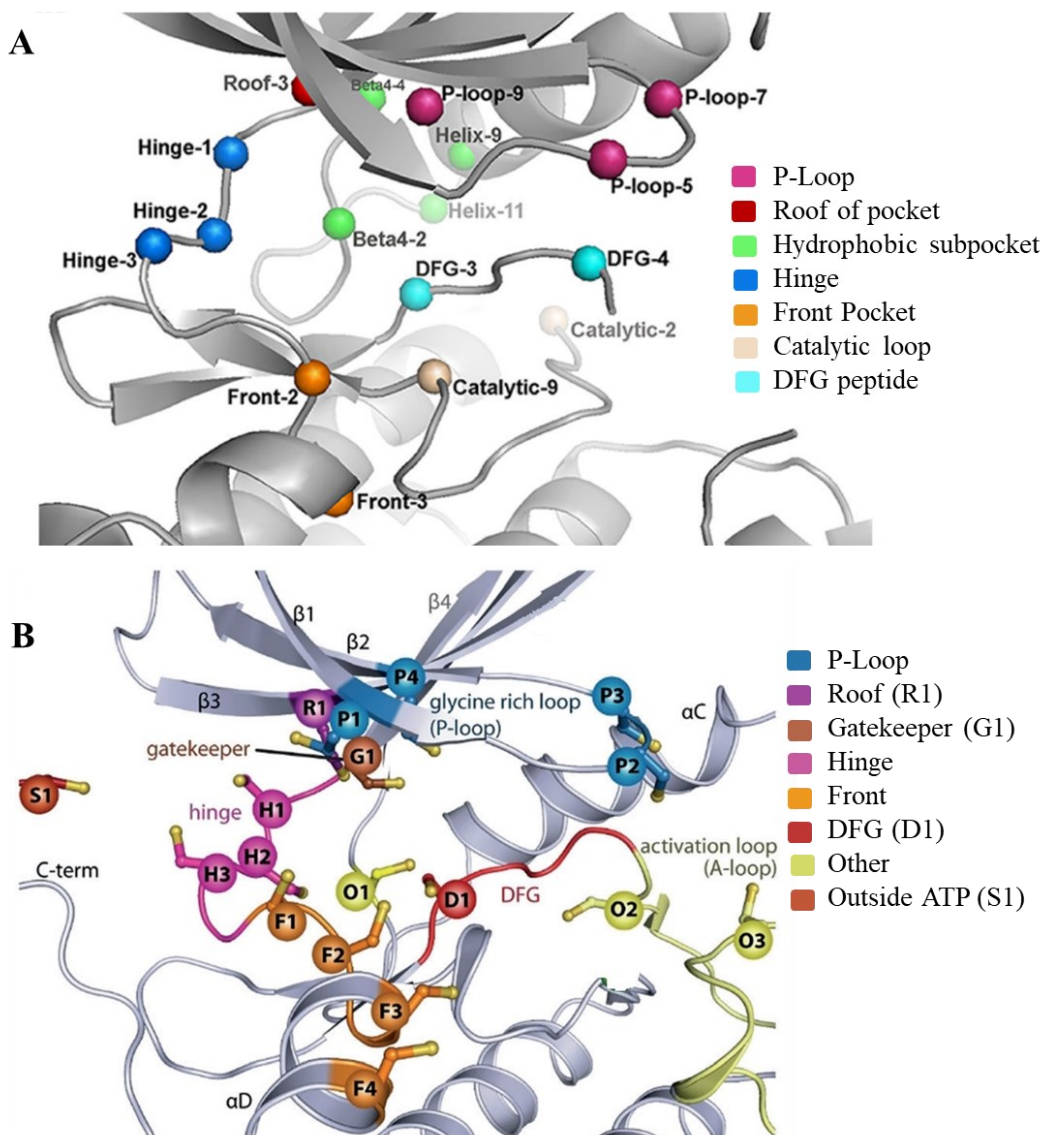


Figura 14 - Mapeamento das cisteínas em posições estratégicas para uma inibição covalente em quinases. **A.** Mapeamento de cisteínas na região do sítio ativo de quinases e proximidades elaborado por Zhao e colaboradores com identificação de 17 posições promissoras para uma estratégia de inibição covalente. **B.** Mapeamento de cisteínas na região do sítio ativo de quinases e proximidades elaborado por Chaikuad e colaboradores de 18 posições promissoras para uma estratégia de inibição covalente.

Fonte: Figura A – Adaptada de Zhao *et al.*, 2017 (doi:10.1021/acs.jmedchem.6b01815)⁶¹; Figura B - Adaptada de Chaikuad *et al.*, 2018 (doi:10.1002/anie.201707875)⁶²

1.5 MODELAGEM COMPARATIVA

A predição de estruturas tridimensionais de proteínas cujas estruturas ainda não foram elucidadas experimentalmente é uma alternativa de grande importância para possibilitar o estudo das propriedades e funções destas macromoléculas, seus mecanismos enzimáticos, e o desenvolvimento e descoberta de novos fármacos direcionados a elas^{63,64}. Além disso, apesar da determinação experimental ser considerada o melhor processo, em alguns casos, devido ao elevado custo financeiro, o tempo e a complexidade de algumas proteínas, o uso de métodos computacionais pode ser a alternativa mais acessível⁶⁵. Como ainda não existem estruturas resolvidas experimentalmente para as proteínas quinases de *T. cruzi*, a construção de modelos teóricos computacionais é essencial para conduzir estudos de inibição direcionados para estas proteínas.

O enovelamento de uma proteína é definido por informações contidas na sua sequência de aminoácidos. A rede de interações intramoleculares formada entre os resíduos de aminoácidos, como ligações de hidrogênio e interações de van der Waals, além de ligações dissulfetos e o efeito hidrofóbico estabilizam a conformação adotada pelas proteínas⁶⁵. Além das informações contidas na estrutura primária das proteínas, um outro aspecto que torna a construção dos modelos tridimensionais teóricos possível é a característica de que a estrutura terciária das proteínas é mais conservada do que as sequências de aminoácidos⁶³. Portanto, sequências com pouca similaridade podem possuir enovelamentos semelhantes. Os métodos desenvolvidos para predição da estrutura de proteínas são divididos entre os métodos independentes de estruturas molde (*template free*) que incluem a predição *ab initio* e predição *de novo*, métodos baseados em estruturas molde (*template based*), constituídos pelas predições de enovelamento, também chamada de *threading*, e modelagem comparativa, sendo esta última a predição utilizada neste estudo e aprofundada nesta discussão⁶⁵, além dos métodos baseados em aprendizagem de máquinas. A determinação de qual dos métodos será usado dependerá da disponibilidade ou não de estruturas resolvidas experimentalmente em bancos de dados para serem usadas como moldes, e também do grau de identidade entre estas estruturas candidatas a molde e a proteína a ser modelada⁶⁵. A proposta destas modelagens computacionais é tentar aproximar a estrutura construída daquela que seria a estrutura nativa da proteína modelada, explorando a energia potencial do espaço conformacional até se chegar em uma estrutura com um mínimo de energia global⁶⁶.

A modelagem comparativa é um dos métodos mais utilizados para predição de estruturas tridimensionais de proteínas e é considerado o mais acurado para este fim^{63,65}. Para empregar este método é necessário que a identidade entre a(s) sequência(s) da(s) estrutura(s) do(s) molde(s) e a sequência da proteína a ser modelada seja de pelo menos de 25 a 30%⁶⁵. Uma identidade inferior a esta leva à construção de um modelo de baixa acurácia⁶⁷. Em contrapartida, havendo uma identidade de cerca de 50% ou mais o modelo construído terá alta qualidade, podendo ser empregado em estudos de mecanismos catalíticos, triagem virtual, *docking* de pequenas moléculas e no design e melhoramento de ligantes⁶⁷. Um dos programas utilizados para este tipo de modelagem é o programa Modeller, que se baseia na satisfação das restrições espaciais presentes na(s) estrutura(s) do(s) molde(s) e no alinhamento feito entre esta(s) sequência(s) molde(s) e a sequência alvo da modelagem^{68,69,70}. No geral, estas restrições espaciais correspondem a limitações estereoquímicas associadas ao comprimento das ligações, ângulo de ligação, ângulos diedrais e distâncias de contato de van der Waals. A modelagem comparativa é constituída de cinco etapas básicas^{63,65,66}:

Etapa 1: Identificação e seleção das estruturas que servirão de molde para a modelagem. Nesta etapa, utilizando ferramentas como o *Basic Local Alignment Search Tool* (BLAST)⁷¹ e a sequência da proteína alvo, é feita a busca em bancos de dados como o PDB⁷² por estruturas moldes. Nesta fase é importante buscar por estruturas que tenham uma boa resolução, de cerca de 2Å ou menos, que apresentem uma identidade satisfatória e boa cobertura com a sequência a ser modelada e que contenham outras moléculas associadas que sejam importantes na definição da conformação buscada (inativa, ativa, etc). Estas moléculas podem ser substratos, inibidores, moléculas reguladoras, dentro outras.

Etapa 2: Alinhamento entre a(s) sequência(s) da(s) estrutura(s) molde(s) selecionada(s) e a sequência da proteína que será modelada. Um bom alinhamento é essencial para a condução de uma modelagem comparativa de qualidade. Nesta etapa é importante avaliar a qualidade geral do alinhamento e verificar se existem regiões sem alinhamento que possam comprometer a modelagem de áreas que venham a impedir ou limitar o uso do modelo em estudos posteriores.

Etapa 3: Construção do modelo tridimensional. Nesta etapa, com o uso de um servidor ou de um programa de modelagem comparativa, as informações extraídas da(s) estrutura(s) do(s) molde(s) e do alinhamento é transferida para os modelos em construção. No fim, serão gerados um determinado número de modelos predeterminados. Nesta fase é possível fazer uma filtragem e pré-seleção dos modelos gerados avaliando a qualidade a partir de funções calculadas pelos próprios servidores/programas, como as funções *Discrete Optimized Protein*

Energy (DOPE score)⁷³ e *Molecular probability density function* (Molpdf)⁶⁸ presentes no programa Modeller.

Etapa 4: Otimização dos modelos. Após a obtenção dos modelos é feita uma etapa de otimização com o intuito de minimizar a energia dos modelos. Neste processo é feita uma otimização geométrica, promovendo, por exemplo, a relaxação de *clashes* (contato inapropriado entre átomos). A dinâmica molecular é uma das possibilidades de otimização.

Etapa 5: Validação dos modelos. Na verificação da qualidade dos modelos tridimensionais gerados para a seleção do melhor modelo, além do DOPE score e do Molpdf, por exemplo, são avaliados outros aspectos estereoquímicos e energéticos. Em relação aos aspectos estereoquímicos, o diagrama de Ramachandran é um dos parâmetros mais importantes. Neste diagrama são verificados os ângulos diedrais ϕ (phi) ψ (psi), permitindo avaliar a qualidade da estrutura secundária e identificar a existência de resíduos *outliers* que estão relacionados com desvios nos ângulos diedrais. Em alguns servidores, como no MolProbity^{74,75}, é possível também avaliar outros parâmetros estereoquímicos como a presença de rotâmeros *outliers* e de *clashes*. Energeticamente, no servidor Saves, é possível avaliar a qualidade da relação 1D-3D construída para os modelos utilizando ferramentas como Errat⁷⁶, Verify3D^{77,78}, dentre outras, e também através do Z-score, calculado no servidor *Protein Structure Analysis* (ProSA)^{79,80}, que permite comparar a qualidade do modelo em relação a estruturas de tamanho semelhante resolvidas experimentalmente. A avaliação da estrutura dos modelos em relação à estrutura molde pode ser feita através do cálculo do *Root Mean Square Deviation* (RMSD)⁸¹. Valores de RMSD menores ou iguais a 2Å correspondem a uma boa sobreposição estrutural entre modelo e molde⁶⁵. Após o alinhamento e sobreposição das estruturas do modelo e do molde, também é interessante analisar a qualidade da modelagem de resíduos e interações chaves da classe de proteína em estudo. Feitas todas estas análises, é selecionado o modelo que representará a estrutura teórica da proteína de interesse.

Após a execução destas etapas, sendo identificada alguma falha ou qualidade ruim nos modelos finais, ou mesmo nos resultados de cada etapa, é necessário reavaliar e modificar as etapas prévias e refazer o processo a fim de se corrigir o erro e obter o modelo final com a qualidade necessárias para os estudos posteriores.

2 OBJETIVOS

2.1 OBJETIVO GERAL

Priorizar quinases de *Trypanosoma cruzi*, a partir de estudos de bioinformática e revisão da literatura, para serem utilizadas como alvos terapêuticos no combate a infecções causadas por este parasito através de estratégias de inibição covalente no sítio de ligação do ATP.

2.2 OBJETIVOS ESPECÍFICOS

- I. Levantar em bancos de dados o conjunto de proteína quinases de *Trypanosoma cruzi*, classificando-as em grupos e famílias.
- II. Realizar o mapeamento das cisteínas relevantes para uma abordagem de inibição covalente nas quinases levantadas, a partir de conhecimentos descritos na literatura.
- III. Identificar e priorizar, dentre as quinases levantadas, aquelas que se mostrem potencialmente promissoras como alvos terapêuticos e que possuam cisteínas mapeadas na região do sítio ativo.
- IV. Predizer a estrutura tridimensional das quinases priorizadas por meio da modelagem comparativa.

3 METODOLOGIA

3.1 LEVANTAMENTO E CLASSIFICAÇÃO DAS QUINASES DE *TRYPANOSOMA CRUZI*

Inicialmente, foi feito o levantamento das quinases de *T. cruzi* no *The Universal Protein Resource* (UniProt)⁸² (<https://www.uniprot.org/>). Neste banco de dados, realizando a busca por meio do *Protein families database* (Pfam)^{83,84} que representa o domínio das proteínas quinases, PF00069, chegou-se num conjunto de quinases da cepa Dm28c 2014⁸⁵. Posteriormente, fez-se um levantamento no banco de dados TriTrypDB⁸⁶ (<https://tritrypdb.org/tritrypdb/app/>), também utilizando o Pfam PF00069. Este é um banco de dados voltado para parasitos cinetoplastídeos que está vinculado ao *Eukaryotic Pathogen, vector and host informatics resource center* (VEuPathDB). O TriTrypDB fornece primariamente dados sobre os organismos *Leishmania braziliensis*, *L. infantum*, *L. major*, *L. tarentolae*, *Trypanosoma brucei* e *T. cruzi*⁸⁶. Neste segundo levantamento focou-se na cepa CL Brener Esmeraldo-like (CL Brener El)⁸⁷, cepa cujos dados constam como curados e como cepa referência no TriTrypDB.

As quinases levantadas para as duas cepas foram classificadas quanto ao grupo e família a que pertencem e nome (este último apenas para a cepa CL Brener El) a partir das anotações dos bancos de dados consultados e da realização de BLAST⁷¹ conduzido na plataforma Kinase.com (<http://kinase.com/web/current/>). Esta plataforma é produzida e mantida pelo laboratório de Gerard Manning na Genentech Inc na Califórnia. Neste processo foram utilizadas as sequências no formato FASTA obtidas no UniProt e no TriTrypDB. Os parâmetros adotados durante o BLAST foram os padrões do servidor, utilizando o *database 4-kinome proteins* que engloba o quinoma de 4 organismos: *Homo Sapiens* (Hsap), *Saccharomyces cerevisiae* (Scer), *Drosophila melanogaster* (Dmel), *Caenorhabditis elegans* (Cele). O critério para efetivar a classificação foi que houvesse uma diferença de pelo menos 5log entre os E-values dos grupos e famílias dos ortólogos do *T. cruzi* retornados nos resultados¹⁵. Não existindo essa diferença de 5log entre os grupos e as famílias optou-se pela classificação “unique” e “not assigned to a family (na)”, respectivamente. Esse valor de referência para o E-Value, assim como as classificações mencionadas, unique e na, são os mesmos utilizados por Parsons *et al.* (2005)¹⁵.

3.2 MAPEAMENTO DAS CISTEÍNAS RELEVANTES PARA INIBIÇÃO COVALENTE

As sequências das quinases das cepas Dm28c e CL Brener El de *T. cruzi* foram alinhadas no programa HMMalign, que faz parte do pacote HMMER (<http://hmmer.org/>). Este programa executa um alinhamento múltiplo utilizando um perfil dos principais padrões de domínios e motivos da família das proteínas de interesse, construído a partir de um conjunto de proteínas representativas desta família (*seeds*).

A partir das sequências alinhadas, o mapeamento das cisteínas promissoras para uma inibição covalente das quinases de *T. cruzi* foi feito a partir de dois trabalhos que analisaram e mapearam estas posições no quinoma humano. Foram mapeadas 17 posições destacadas no estudo de Zhao *et al.* (2017)⁶¹ juntamente com 18 posições levantadas por Chaikwad *et al.* (2018)⁶², totalizando 25 posições mapeadas, sendo 10 posições comuns aos dois estudos.

3.3 PRIORIZAÇÃO DAS QUINASES

A partir desta etapa focou-se apenas na cepa CL Brener El por ser a cepa referência de *T. cruzi* no banco de dados TriTrypDB. Para a priorização das quinases de *T. cruzi* partiu-se do conhecimento sobre as quinases de *T. brucei* destacadas em estudos de RNA de interferência (RNAi), como os trabalhos de Jones *et al.*³¹ e de Fernandez-Cortes *et al.*³². Foram elencadas 56 quinases de *T. brucei*, a partir das quais foi feita a busca pelas proteínas ortólogas em *T. cruzi* através de BLAST conduzido no servidor TriTrypDB, utilizando as sequências das quinases de *T. brucei* levantadas neste mesmo servidor. Além disso, consultou-se a seção de ortólogos dentro da página de cada quinase de *T. brucei* no TriTrypDB. Na sequência, dentre as quinases ortólogas de *T. cruzi*, foram selecionadas aquelas que possuíam cisteínas mapeadas em posições favoráveis para uma inibição covalente, conforme descrito previamente. Por fim, para seguir para a etapa de predição da estrutura tridimensional, dentre as quinases que possuíam cisteínas mapeadas, fez-se mais um refinamento selecionando aquelas que foram destacadas nos dois estudos de RNAi mencionados e que levaram à morte do organismo *T. brucei*.

3.4 MODELAGEM COMPARATIVA

Nesta etapa do trabalho foi feita a predição da estrutura tridimensional por modelagem comparativa das três proteínas priorizadas, AUK1, CLK1 e CRK9. A modelagem da AUK1 foi conduzida com a quinase da cepa CL Brener haplótipo Non-Esmeraldo-like (CL Brener NEI) e as demais com as quinases da cepa CL Brener haplótipo Esmeraldo-like (CL Brener EI). Até o momento da escrita deste trabalho a estrutura destas três proteínas ainda não tinham sido elucidadas experimentalmente.

3.4.1 Obtenção dos modelos

A busca de moldes para as sequências das quinases priorizadas foi feita por meio de BLAST, mais especificamente por BLASTp, direcionado para proteínas, disponível no site do *National Center for Biotechnology Information* (NCBI)^{88,89} (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>), utilizando o banco de dados de estruturas PDB⁷². Também se utilizou da ferramenta de busca por moldes do servidor Swiss-Model^{90,91}, disponibilizado pelo *Swiss Institute of Bioinformatics* (SIB)⁹² (<https://swissmodel.expasy.org/interactive>). Os candidatos a moldes foram analisados quanto ao grau de identidade, similaridade e cobertura com a sequência alvo, quanto à resolução da estrutura cristalográfica, à existência de resíduos faltantes na estrutura e presença de inibidores. Outro aspecto considerado para a escolha dos moldes foi a presença de elementos necessários para a conformação ativa das quinases, como fosforilações e associação com outras proteínas ou peptídeos, ou mesmo a descrição desta informação no artigo referente à estrutura cristalográfica em análise.

Para a modelagem das quinases foi utilizado o programa Modeller versão 9.25^{68,69,70}. O alinhamento foi produzido no ClustalΩ⁹³ versão 1.2.4 (<https://www.ebi.ac.uk/Tools/msa/clustalo/>), obtendo no fim arquivo com as sequências alinhadas em formato FASTA que posteriormente foi editado e adequado para o formato lido pelo programa Modeller. Em cada processo de modelagem de uma sequência foram gerados 100 modelos, com transferência do inibidor ligado no sítio do ATP do molde para os modelos. Foram pré-selecionados para a etapa de validação os 10 modelos com a menor pontuação

DOPE⁷³, com posterior seleção dos 3 com menor pontuação Molpdf⁶⁸, ambas funções calculadas pelo programa Modeller. DOPE é um potencial estatístico dependente da distância atômica, é uma função de pontuação que permite avaliar a energia dos modelos gerados, cujos resultados com os menores valores correspondem aos modelos hipoteticamente mais próximos à estrutura nativa. O parâmetro Molpdf está relacionado com as violações das restrições espaciais previamente construídas a partir do alinhamento entre o alvo e o molde. Molpdf é a soma das contribuições de todas as violações ocorridas, e, semelhante ao DOPE, quanto menor o valor melhor é a qualidade do modelo gerado. Posteriormente, foi feito o refinamento dos 3 melhores modelos no servidor *Local Protein Structure Refinement via Molecular Dynamics* (locPREFMD)⁹⁴ disponibilizado pelo *Feig Computational Biophysics Lab* da *Michigan State University* (<http://feig.bch.msu.edu/web/services/locprefmd/>). O refinamento realizado neste servidor é local, conduzido por simulações de dinâmica molecular, visando melhorar parâmetros estereoquímicos da proteína sem comprometer o seu enovelamento geral.

3.4.2 Validação dos modelos

A validação dos modelos selecionados e dos refinamentos foi feita através da avaliação de alguns parâmetros além do DOPE e do Molpdf mencionados anteriormente, como o RMSD⁸¹, e o diagrama de Ramachandran⁹⁵. O RMSD é calculado a partir do alinhamento e sobreposição das estruturas dos modelos pré-selecionados com seus respectivos moldes utilizados na modelagem. Simplificadamente, consiste no cálculo do desvio quadrático médio das coordenadas x, y e z dos pares de átomos equivalentes alinhados e sobrepostos, tendo como resultado final a sobreposição que minimize o valor do RMSD⁸¹. O cálculo do RMSD pode ser equacionado como:

$$RMSD = \sqrt{\frac{1}{n} \sum_{i=1}^n d_i^2}$$

em que n corresponde aos n pares de átomos equivalentes e d_i é a distância entre os dois átomos alinhados e sobrepostos⁹⁶.

Estes valores de RMSD foram obtidos no programa PyMOL versão 2.4.1⁹⁷. Quanto ao diagrama de Ramachandran, foi avaliada a proporção de resíduos com ângulos ϕ (phi) ψ (psi) em regiões favoráveis, regiões permitidas e regiões não permitidas (*outliers*), através do servidor MolProbity^{74,75} versão 4.5.1 (<http://molprobity.biochem.duke.edu/>, mantido pelo *Department of Biochemistry da Duke University School of Medicine*). Os valores de corte para uma estrutura de boa qualidade para o MolProbity são de 98,0% ou mais dos resíduos em regiões favoráveis, 99,8% ou mais em regiões permitidas e até 0,5% em regiões não permitidas. Outras ferramentas empregadas na avaliação da qualidade dos modelos foram o Errat⁷⁶ e o Verify3D^{77,78}, disponibilizados no servidor Saves versão 6.0 (<https://saves.mbi.ucla.edu/>), e o ProSA⁸⁰, desenvolvido pelo *Center of Applied Molecular Engineering da Paris Lodron Universität Salzburg* (<https://prosa.services.came.sbg.ac.at/prosa.php>). O Errat é uma ferramenta que inspeciona o padrão de pares de interações não covalentes entre os átomos de carbono (C), oxigênio (O) e nitrogênio (N) dos resíduos, podendo apresentar seis pares de interações, CC, CN, CO, NN, NO e OO. Essas interações não acontecem de forma aleatória e este aspecto permite identificar possíveis irregularidades estruturais e compreender a qualidade 3D da proteína⁷⁶. O Errat retorna um fator de qualidade geral que corresponde à porcentagem da proteína para a qual o valor do erro fica abaixo do limite de rejeição de 95%. Estruturas com resoluções altas costumam apresentar um fator de qualidade de 95% ou superior e estruturas de resoluções mais baixas, entre 2,5 e 3,0Å, geralmente apresentam um fator de qualidade geral em torno de 91%. A ferramenta Verify 3D verifica a compatibilidade da estrutura tridimensional das proteínas com as informações presentes nas suas sequências de aminoácidos, como o ambiente e características dos resíduos. Os critérios considerados são as áreas dos resíduos que se encontram enterradas, a fração de átomos polares presente nas cadeias laterais e a estrutura secundária local⁷⁸. Quando pelo menos 80% dos resíduos em análise apresentam uma pontuação maior ou igual a 0,2, há um indício de que a estrutura terciária da proteína está adequada. O ProSA também é uma ferramenta cujos cálculos de pontuações e gráficos avaliam a acurácia dos modelos. Esta ferramenta fornece a pontuação Z-score que indica a qualidade geral da estrutura da proteína, e um gráfico que referencia esta pontuação com as de estruturas proteicas determinadas experimentalmente por cristalografia de Raios X e Ressonância Magnética Nuclear (RMN) depositadas no PDB. Sendo assim, é possível verificar se o Z-score da estrutura em análise está dentro da faixa de pontuação normalmente encontrada para proteínas nativas de tamanho semelhante⁸⁰. O processo de validação como um todo foi feito utilizando apenas as partes das proteínas que foram cobertas pelos moldes nos alinhamentos e modeladas pelo programa Modeller, com exceção do cálculo do RMSD entre os modelos

gerados e os moldes para a proteína CRK9 em que foram mantidos os trechos de inserção dentro do domínio de quinases que não ficaram cobertos e que não foram modelados. Por fim, foi selecionado o modelo final para cada quinase priorizada.

3.4.3 Análises complementares dos modelos gerados

Para a modelagem da quinase AUK1 foram necessárias análises complementares para compreender as diferenças entre os dois modelos construídos, um correspondendo à forma ativa e o outro inativa dos moldes utilizados. O deslocamento entre os lóbulos N e C entre os dois modelos foi calculado através do servidor *Protein Domain Movement Analysis*⁹⁸ (<http://dyndom.cmp.uea.ac.uk/dyndom/runDyndom.jsp>). Um outro aspecto analisado foi o volume do sítio do ATP destes dois modelos, a partir da ferramenta DoGSite⁹⁹ disponível no servidor *ProteinsPlus - Structure-Based Modeling Support Server* (<https://proteins.plus/>) mantida pela *Universität Hamburg*.

Em relação à modelagem da quinase CRK9 foi realizada a comparação entre o modelo construído através do programa Modeller com o modelo gerado pelo Alphafold¹⁰⁰, preditor de estrutura tridimensional que se baseia no aprendizado de máquinas, incorporando nova arquitetura de rede neural. O cálculo de RMSD entre as duas estruturas foi feito utilizando o programa PyMOL versão 2.4.1⁹⁷.

4 RESULTADO E DISCUSSÃO

4.1 LEVANTAMENTO E CLASSIFICAÇÃO DAS QUINASES DO *TRYPANOSOMA CRUZI*

No sequenciamento do genoma da cepa CL Brener de *T. cruzi* foram descritas 167 quinases ePKs e 19 aPKs⁸⁷. Já segundo Parsons *et al.* (2005)¹⁵, o organismo *T. cruzi* possui 171 ePKs e 19 aPKs, totalizando 190 quinases¹⁵. Na busca feita no Uniprot, para a cepa Dm28c 2014, foram levantadas 171 quinases, enquanto para a cepa CL Brener EI obteve-se um conjunto de 180 quinases no banco de dados TriTrypDB. Destas 180 quinases 5 estão classificadas como pseudogenes e foram desconsideradas, restando um total de 175 proteínas. Era esperado que as PKs levantadas para as duas cepas correspondessem apenas às ePKs, uma vez que as aPKs não apresentam sequências intimamente relacionadas com as proteínas quinases eucaritóticas devido à falta dos 11 subdomínios que as definem¹⁵ e, portanto, não estão representadas pelo Pfam PF00069.

A classificação das quinases de *T. cruzi* foi conduzida com o intuito de melhor conhecer os conjuntos de quinases levantados para cada cepa e eventualmente auxiliar nas etapas posteriores deste projeto. A tabela completa, discriminando os grupos, famílias e nomes das 175 quinases levantadas para a cepa CL Brener EI está apresentada na seção seguinte (Mapeamento das cisteínas relevantes para estratégia de inibição covalente) em conjunto com o mapeamento das cisteínas. A tabela referente a estes mesmos dados para a cepa Dm28c 2014 está disponível no Apêndice A. Uma síntese dos resultados obtidos para as duas cepas, em contraste com a classificação feita por Parsons *et al.* (2005) referente às quinases da cepa CL Brener está representada na **Tabela 4**. Parsons *et al.* não discriminaram se se tratava do haplótipo CL Brener Esmeraldo-like (CL Brener EI) ou CL Brener Non-Esmeraldo-like (CL Brener NEI). Nenhum dos grupos e famílias retornados pelo BLAST no kinase.com e nas anotações dos bancos de dados se referiram às aPKs, conforme já era esperado. Optou-se por buscar pelas sequências das quinases atípicas de *T. cruzi* caso houvesse necessidade nas etapas seguintes deste trabalho.

Tabela 4 - Comparativo da classificação das quinases de *T. cruzi* das cepas Dm28c 2014 e CL Brener El em grupos e famílias, juntamente com classificação feita por Parsons *et al.*¹⁵ para a cepa CL Brener.

Grupo	Família	Nº de quinases			Grupo	Família	Nº de quinases		
		Dm28c 2014	CL Brener El	Parsons <i>et al.</i> ¹⁵			Dm28c 2014	CL Brener El	Parsons <i>et al.</i> ¹⁵
AGC	(na)	8	3	7	Other	(na)	2	0	0
	NDR	2	1	1		AUR	3	4	3
	PKA	1	1	3		CAMKK	4	4	4
	RSK	0	2	1		NAK	2	2	1
	PKB	0	2	0		NEK	19	20	22
	Total	11	9	12		PEK	2	1	2
CAMK	(na)	7	7	7		PLK	2	2	2
	CAMKL	6	6	6		TLK	1	1	1
	CDPK1	1	0	0		ULK	2	1	1
	Total	14	13	13		VPS	0	0	1
CK1	(na)	0	0	0		WEE	1	2	2
	CK1	2	8	7	SCY1	1	1	0	
	TTBK	1	1	1	Total	39	38	39	
	Total	3	9	8	STE	(na)	12	6	8
CMGC	(na)	4	1	1		STE7	2	2	2
	CDK	10	17	10		STE11	18	15	18
	CLK	3	5	5		STE20	3	3	3
	DYRK	6	6	7		Total	35	26	31
	GSK	2	2	2		Unique		27	33
	CDKL	0	0	1	Total ePK			171	175
	MAPK	11	10	11					
	RCK	3	2	3					
	SRPK	1	2	2					
	CK2	2	2	2					
Total	42	47	44						

(na) - not assigned to a Family

De maneira geral percebe-se uma convergência da classificação entre as cepas analisadas neste trabalho e a classificação apresentada por Parsons *et al.* O processo neste último foi mais elaborado, envolvendo não somente análises por BLAST no servidor kinase.com, mas também análise filogenética. É provável que parte das quinases das cepas Dm28c 2014 e CL Brener El aqui classificadas como *unique* sejam realocadas para outros grupos e famílias ao conduzir uma análise mais refinada. Algumas famílias presentes na **Tabela 4** não estavam assinaladas na classificação de Parsons *et al.*, como a família PKB do grupo AGC, família CDPK1 do grupo CAMK e família SCY1 do grupo Other, possivelmente por não terem sido identificadas quinases pertencentes a elas na ocasião. A família CK2 no trabalho de Parsons *et al.* estava

associada ao grupo Other e hoje esta família se encontra alocada no grupo CMGC. Uma pequena observação quanto ao número de quinases do grupo Other no trabalho de Parsons *et al.* é que, segundo estes, o somatório de quinases deste grupo resultaria em 40, desconsiderando as da família CK2. Mas, aparentemente, houve um equívoco na contagem e a soma de quinases deste grupo resulta em 39, levando a um somatório geral de 170 quinases e não 171. Os tripanossomatídeos não possuem quinases dos grupos TK e TK1, grupos compostos por enzimas que fosforilam tirosinas e que comumente representam receptores em diversos organismos^{87,15}. Mas, apesar deste desfalque, as quinases da família WEE fosforilam tirosina e existem quinases de especificidade dual, como as das famílias CLK e DYRK do grupo CMGC e da família STE7, do grupo STE¹⁵. Em relação ao quinoma humano, os tripanossomatídeos são pobremente representados pelos grupos CAMK e AGC e, em contraste, apresentam boa representatividade de quinases dos grupos CMGC e STE e da família NEK do grupo Other^{87,15}.

4.2 MAPEAMENTO DAS CISTEÍNAS RELEVANTES PARA INIBIÇÃO COVALENTE

Após a classificação das quinases das cepas Dm28c 2014 e CL Brener El fez-se o mapeamento das cisteínas mais promissoras para uma estratégia de inibição covalente baseado nos mapeamentos em sítios e subsítios elaborados por Zhao *et al.* (17 posições)⁶¹ e Chaikuad *et al.* (18 posições)⁶². O compilado destas 25 posições mapeadas a partir destes dois estudos com suas respectivas nomenclaturas para os sítios e subsítios e a indicação do resíduo correspondente nas estruturas adotadas como “mapas”, código PDB 3BYU no trabalho de Zhao *et al.* e código PDB 1ATP no de Chaikuad *et al.* está apresentado na **Tabela 5**.

Tabela 5 - Compilado dos sítios e subsítios dos mapeamentos de cisteínas promissoras para uma estratégia de inibição covalente elaborados por Zhao *et al.*⁶¹ e Chaikuad *et al.*⁶².

Sítio	Subsítio Chaikuad	Subsítio Zhao	Resíduo PDB 1ATP ^a	Resíduo PDB 3BYU ^b
Glycine-rich loop region (P-loop) ^a P-Loop ^b	P1		L49	L251
	P2		S53	Q255
	P3	P-Loop-7	F54	F256
	P4	P-Loop-9	V57	V259

Sítio	Subsítio Chaikuad	Subsítio Zhao	Resíduo PDB 1ATP ^a	Resíduo PDB 3BYU ^b
DFG region ^a DFG Peptide ^b	D1	DFG-3	T183	A381
Roof region ^a Roof of Pocket ^b	R1	Roof-3	A70	A271
Gatekeeper ^a	G1		M120	T316
Hinger region ^a Hinge ^b	H1	Hinge-1	Y122	Y381
	H2	Hinge-2	V123	M319
	H3	Hinge-3 (A124)	A124 G125	E320 N321
Front region ^a Front of Pocket ^b	F1		G126	G322
	F2	Front-2	E127	S323
	F3	Front-3	F129	V325
	F4		R133	K329
Outside ATP ^a Catalytic Loop (Catalytic-7) ^b	S1		P169	A367
Others ^a Beta Sheet 4 (Beta 4-2 (V301)) ^b DFG Peptide (DFG-5 (L385)) ^b	O1	Beta 4-2	V104	V301
	O2		F187	L385
	O3		A188 C199	A386 E398
Beta Sheet 4 ^b		Beta 4-4	L106	L303
P-Loop ^b		P-Loop-5	G52	G254
Helix ^b		Helix-9	L95	M292
		Helix-11	V98	L295
Catalytic Loop ^b		Catalytic-2	L162	Y360
		Catalytic-9	L172	I370
DFG Peptide ^b		DFG-4	G186	G384

^a Nomenclatura e estrutura cristalográfica (“mapa”) utilizada por Chaikuad *et al.*⁶².

^b Nomenclatura e estrutura cristalográfica (“mapa”) utilizada por Zhao *et al.*⁶¹

Para a cepa Dm28c 2014 foram mapeadas 172 cisteínas distribuídas nos diversos sítios e subsítios de 112 quinases (**Apêndice A**). Das 171 quinases levantadas, 59 não possuem cisteínas em posições relevantes (**Apêndice A**). No mapeamento das cisteínas das quinases da cepa CL Brener El foi identificado um total de 164 cisteínas distribuídas em 100 quinases (**Tabela 6**), restando 75 de 175 proteínas sem cisteínas promissoras para inibição covalente (**Apêndice B**). Na **Tabela 6** e nos Apêndices A e B, adotou-se a nomenclatura de Chaikuad *et al.* para os sítios e subsítios que são comuns aos dois mapeamentos utilizados como referência.

Nas demais posições, exclusivas de cada mapeamento, foram utilizadas as respectivas nomenclaturas. Além disso, optou-se pela estrutura de código PDB 1 ATP usada por Chaikuad *et al.* para indicar as 25 posições mapeadas. Estas definições foram convencionadas ao longo do texto. Nestas tabelas também estão presentes as classificações em grupos e famílias de cada quinase, assim como o nome atribuído às proteínas da cepa CL Brener El, quando definido.

Tabela 6 – Classificação em grupos e famílias das 100 quinases de *T. cruzi* da cepa CL Brener El com cisteínas mapeadas, e detalhamento dos sítios e subsítios de localização das cisteínas

Código TriTrypDB	Código Uniprot	Classificação			Mapeamento Cisteínas			
		Grupo	Família	Nome	Sítio	Subsítio	Resíduo PDB 1ATP ^a	Resíduo CL Brener El
TcCLB.506401.110	Q4E3T2	STE	STE11	-	P-Loop P	P2	S53	C651
					P-Loop P	P3	F54	C652
					Hinge	H2	V125	C726
					Others	O2	F187	C872
TcCLB.503613.10	Q4CYK6	Other	CAMKK	-	P-Loop P	P3	F54	C208
TcCLB.506529.340	Q4E5P6	STE	STE11	-	P-Loop P	P3	F54	C301
					Front	F3	F129	C376
TcCLB.510407.9	Q4D7E4	Other	NEK	NEK11	P-Loop P	P4	V57	C18
TcCLB.506513.160	Q4DPI3	Other	PLK	PLK	P-Loop P	P4	V57	C62
					Hinge	H2	V123	C126
					Others	O3	C199	C205
TcCLB.511127.320	Q4E3Z1	CMGC	CLK	CLK1 KKT10	DFG	D1	T183	C305
					Front	F1	G126	C232
					Others	O2	A188	C310
TcCLB.506943.10	Q4DCP6	CMGC	MAPK	MAPK5	DFG	D1	T183	C188
TcCLB.509099.150	Q4DNX1	CMGC	CDK	CRK9	DFG	D1	T183	C502
					Hinge	H2	V123	C425
TcCLB.507993.80	Q4E1X0	CMGC	GSK	GSK3	DFG	D1	T183	C170
TcCLB.510745.40	Q4CTC3	STE	STE20	-	DFG	D1	T183	C196
					Others	O2	A188	C201
TcCLB.506483.69	Q4D3Y2	CMGC	MAPK	MPK4	DFG	D1	T183	C173
TcCLB.510257.130	Q4DBK1	CAMK	CAMKL	-	DFG	D1	T183	C154
TcCLB.507547.80	Q4D8V8	CMGC	CLK	-	DFG	D1	T183	C175
					Others	O1	V104	C86
					Helix	Helix-9	L95	C77
TcCLB.503575.60	Q4D201	CMGC	MAPK	-	DFG	D1	T183	C156
					Beta Sheet 4	Beta 4-4	L106	C76
TcCLB.506857.10	Q4D9J9	CMGC	GSK	GSK3A	DFG	D1	T183	C238

Código TriTrypDB	Código Uniprot	Classificação			Mapeamento Cisteínas			
		Grupo	Família	Nome	Sítio	Subsítio	Resíduo PDB 1ATP ^a	Resíduo CL Brener EI
TcCLB.511573.40	Q4CU32	CMGC	MAPK	MPK10	DFG	D1	T183	C175
TcCLB.506525.140	Q4DK21	STE	STE11	MRK1	DFG	D1	T183	C444
					Others	O2	F187	C448
TcCLB.511001.60	Q4DZ37	CAMK	CAMKL	-	DFG	D1	T183	C161
TcCLB.509103.4	Q4CV59	Unique			DFG	D1	T183	C17
TcCLB.506737.20	Q4DGS9	CMGC	MAPK	-	DFG	D1	T183	C282
TcCLB.506953.30	Q4CZI6	Other	NAK	-	DFG	D1	T183	C189
TcCLB.504883.10	Q4D6L8	Other	NAK	-	DFG	D1	T183	C212
					Hinge	H2	V123	C147
TcCLB.510105.120	Q4DZU2	Unique			DFG	D1	T183	C1181
					Others	O1	V104	C1038
TcCLB.511127.330	Q4E3Z0	CMGC	CLK	KKT19	DFG	D1	T183	C313
					Front	F1	G126	C240
					Others	O2	A188	C318
TcCLB.510123.20	Q4CZQ7	CMGC	MAPK	MAPK11	DFG	D1	T183	C167
TcCLB.506869.60	Q4D2D8	CMGC	DYRK	PK4	Roof R	R1	A70	C287
TcCLB.510285.70	Q4DIT1	Unique		KKT2	Hinge	H1	Y122	C195
					Hinge	H2	V123	C196
					Catalytic Loop	Catalytic-2	L162	C234
TcCLB.506493.70	Q4DE65	STE	STE11	-	Hinge	H1	Y122	C1342
TcCLB.511751.170	Q4DVU5	Unique		EIF2AK3	Hinge	H2	V123	C730
					Catalytic Loop	Catalytic-2	L162	C768
TcCLB.511751.50	Q4DVW0	CMGC	CDK	CRK11	Hinge	H2	V123	C126
TcCLB.504229.30	Q4DCW6	CMGC	CDK	-	Hinge	H2	V123	C83
					Catalytic Loop	Catalytic-9	L172	C134
TcCLB.506493.50	Q4DE67	CAMK	CAMKL	-	Hinge	H2	V123	C107
					Others	O2	A188	C194
					Others	O3	C199	C211
TcCLB.511055.50	Q4D1Y3	CMGC	CDK	-	Hinge	H2	V123	C128
					Catalytic Loop	Catalytic-9	L172	C179
TcCLB.510077.30	Q4CS57	Other	NEK	NEK22	Hinge	H2	V123	C100
					Others	O3	C199	C184
TcCLB.511351.30	Q4CQ91	Other	NEK	-	Hinge	H2	V123	C109
					Others	O3	C199	C190
TcCLB.504133.30	Q4CXT1	Unique			Hinge	H2	V123	C336
TcCLB.506801.120	Q4DH28	Other	CAMKK	-	Hinge	H2	V123	C386
					Catalytic Loop	Catalytic-9	L172	C433

Código TriTrypDB	Código Uniprot	Classificação			Mapeamento Cisteínas			
		Grupo	Família	Nome	Sítio	Subsítio	Resíduo PDB 1ATP ^a	Resíduo CL Brener EI
TcCLB.508207.20	Q4DVN7	CMGC	CDK	-	Hinge	H2	V123	C128
					Catalytic Loop	Catalytic-9	L172	C179
TcCLB.508277.20	Q4E0Z5	AGC	(na)	PDK1	Hinge	H2	V123	C98
					Others	O3	C199	C203
					Helix	Helix-11	V98	C73
TcCLB.508325.130	Q4E0I1	CMGC	CDK	-	Hinge	H2	V123	C128
TcCLB.510175.160	Q4DRP6	CMGC	CDK	-	Hinge	H2	V123	C128
					Catalytic Loop	Catalytic-9	L172	C179
TcCLB.503799.4	Q4CMJ2	Other	AUR	AUK1	Hinge	H2	V123	C109
					Others	O3	C199	C186
TcCLB.510089.170	Q4DQZ5	CK1	CK1	-	Hinge	H2	V123	C112
					Helix	Helix-11	V98	C87
TcCLB.508325.320	Q4E0H6	CMGC	CDK	-	Hinge	H2	V123	C128
					Catalytic Loop	Catalytic-9	L172	C179
TcCLB.506743.160	Q4DPN9	STE	STE20	-	Hinge	H2	V123	C98
					Helix	Helix-11	V98	C73
TcCLB.507485.90	Q4DDD1	Other	WEE	WEE1	Hinge	H2	V123	C270
					Others	O2	F187	C338
TcCLB.510345.9	Q4CTL6	Other	NEK	NEK15	Hinge	H2	V123	C108
TcCLB.507611.40	Q4E3H3	CMGC	CDK	-	Hinge	H2	V123	C128
					Catalytic Loop	Catalytic-9	L172	C179
TcCLB.509169.39	Q4CSM2	Unique		NEK10	Hinge	H2	V123	C227
TcCLB.511471.60	Q4D838	CMGC	CDK	-	Hinge	H2	V123	C128
					Catalytic Loop	Catalytic-9	L172	C179
TcCLB.506697.30	Q4D396	Other	NEK	NEK6	Hinge	H2	V123	C143
					Others	O2	A188	C212
TcCLB.507211.20	Q4CUG8	Other	NEK	-	Hinge	H2	V123	C109
					Others	O3	C199	C190
TcCLB.508987.30	Q4CZA3	Unique			Hinge	H2	V123	C756
					Others	O2	A188	C865
TcCLB.508831.110	Q4DX17	Other	NEK	NEK20	Hinge	H2	V123	C89
					Beta Sheet 4	Beta 4-4	L106	C70
TcCLB.510055.160	Q4DMM6	STE	STE11	-	Hinge	H2	V123	C137
					Others	O2	F187	C205
TcCLB.503685.10	Q4CQY7	Other	AUR	AUK2	Hinge	H2	V123	C189
					Others	O3	C199	C300

Código TriTrypDB	Código Uniprot	Classificação			Mapeamento Cisteínas			
		Grupo	Família	Nome	Sítio	Subsítio	Resíduo PDB 1ATP ^a	Resíduo CL Brener EI
TcCLB.503627.20	Q4CS54	Unique			Hinge	H2	V123	C1352
TcCLB.506945.350	Q4DX65	Other	PLK	-	Hinge	H2	V123	C257
					Others	O3	C199	C336
TcCLB.507881.60	Q4D8Q9	CMGC	CDK	-	Hinge	H2	V123	C128
					Catalytic Loop	Catalytic-9	L172	C179
TcCLB.507275.10	Q4D4W6	CMGC	CDK	CRK4	Hinge	H2	V123	C74
TcCLB.510565.70	Q4DPG2	STE	STE11	-	Hinge	H2	V123	C812
TcCLB.507837.20	Q4D4L7	Other	NEK	-	Hinge	H2	V123	C115
TcCLB.506457.120	Q4DG94	Other	SCY1	-	Hinge	H2	V123	C104
					Beta Sheet 4	Beta 4-4	L106	C87
TcCLB.507683.40	Q4D2W9	CMGC	DYRK	-	Hinge	H2	V123	C788
TcCLB.509007.60	Q4DSM2	AGC	(na)	RAC	Hinge	H3	A124	C431
					Others	O3	C199	C509
TcCLB.506211.220	Q4DYJ6	CAMK	(na)	AKB1	Hinge	H3	G125	C215
					Front	F3	F129	C219
TcCLB.511127.220	Q4E401	CMGC	CDK	CRK12	Hinge	H3	G125	C462
					Others	O2	A188	C524
TcCLB.503715.40	Q4CZQ9	STE	STE11	-	Hinge	H3	G125	C962
					Others	O2	F187	C1024
TcCLB.510525.10	Q4D6W8	CAMK	CAMKL	-	Front	F3	F129	C90
TcCLB.506205.40	Q4CU00	STE	STE11	-	Others	O1	V104	C231
TcCLB.506773.100	Q4DQF3	Unique			Others	O2	F187	C266
					Beta Sheet 4	Beta 4-4	L106	C172
TcCLB.506925.180	Q4E4M9	Unique			Others	O2	A188	C394
					Others	O3	C199	C423
TcCLB.506885.120	Q4E4I5	CMGC	RCK	MOK	Others	O2	A188	C150
TcCLB.510533.70	Q4DX07	Unique			Others	O2	A188	C166
TcCLB.510323.10	Q4D9C1	Unique			Others	O2	A188	C350
TcCLB.511621.110	Q4DT55	STE	(na)	-	Others	O2	A188	C696
					Beta Sheet 4	Beta 4-4	L106	C610
TcCLB.511801.14	Q4D7I3	CAMK	(na)	LDK	Others	O3	C199	C173
TcCLB.509599.150	Q4DP62	Other	NEK	NEK17	Others	O3	C199	C205
					Beta Sheet 4	Beta 4-4	L106	C95
					Helix	Helix-11	V98	C87
TcCLB.508943.20	Q4CUK2	Unique			Others	O3	C199	C508
TcCLB.510525.139	Q4D6V5	Other	NEK	-	Others	O3	C199	C230
					Helix	Helix-11	V98	C100

Código TriTrypDB	Código Uniprot	Classificação			Mapeamento Cisteínas			
		Grupo	Família	Nome	Sítio	Subsítio	Resíduo PDB 1ATP ^a	Resíduo CL Brener EI
TcCLB.503599.10	Q4D2U3	CAMK	(na)	-	Others	O3	C199	C209
TcCLB.510861.140	Q4DPG9	CAMK	CAMKL	-	Others	O3	C199	C173
TcCLB.506885.280	Q4DWX0	Other	NEK	-	Others	O3	C199	C229
					Helix	Helix-11	V98	C119
TcCLB.509599.130	Q4DP63	Other	NEK	-	Others	O3	C199	C205
					Beta Sheet 4	Beta 4-4	L106	C95
					Helix	Helix-11	V98	C87
TcCLB.509805.10	Q4E007	AGC	PKA	PKAC3	Others	O3	C199	C39
TcCLB.506465.40	Q4D6M3	CAMK	(na)	-	Others	O3	C199	C742
TcCLB.464807.10	Q4CNP2	CAMK	CAMKL	-	Others	O3	C199	C124
TcCLB.510347.60	Q4D056	CAMK	(na)	-	Others	O3	C199	C302
TcCLB.506513.50	Q4DPJ2	CAMK	(na)	-	Others	O3	C199	C189
					Catalytic Loop	Catalytic-9	L172	C159
TcCLB.509965.260	Q4E0T4	AGC	PKB	PKB	Others	O3	C199	C245
TcCLB.508205.30	Q4CUA3	CAMK	(na)	-	Others	O3	C199	C219
TcCLB.510349.80	Q4D3R7	Other	AUR	AUK3	Others	O3	C199	C172
TcCLB.503953.30	Q4D0W2	Other	NEK	NEK6	Others	O3	C199	C208
TcCLB.506715.10	Q4CQ74	Other	AUR	AUK3	Others	O3	C199	C172
TcCLB.510971.60	Q4DAR4	Other	NEK	NRKB	Others	O3	C199	C184
					Helix	Helix-11	V98	C73
TcCLB.476421.9	Q4CLT4	STE	STE11	-	Others	O3	C199	C284
					Beta Sheet 4	Beta 4-4	L106	C169
TcCLB.510409.10	Q4D1T8	STE	STE11	-	Beta Sheet 4	Beta 4-4	L106	C278
TcCLB.509599.157	Q4DP61	Other	NEK	-	Helix	Helix-11	V98	C182
TcCLB.506695.30	Q4CXD1	Other	WEE	-	Helix	Helix-11	V98	C108
TcCLB.509551.10	Q4DJK1	STE	STE20	-	Catalytic Loop	Catalytic-2	L162	C72

(na) - not assigned to a Family

Alguns sítios e subsítios não apresentaram quinases contendo cisteínas no mapeamento das duas cepas de *T. cruzi* analisadas, como os sítios *Gatekeeper* e *Outside ATP*, o subsítio F2 do sítio *Front region* e o subsítio *P-Loop-5* do sítio *P-loop*. Outras posições ficaram ausentes em somente uma das cepas, como o subsítio P1 do sítio *Glycine-rich loop region (P-loop)*, o subsítio F4 do sítio *Front Region (Front of Pocket)* e subsítio DFG-4 do sítio *DFG Peptide*

para a cepa CL Brener El e o subsítio *Helix-9* do sítio *Helix* para a cepa Dm28c 2014. A síntese do mapeamento das cisteínas das quinases das duas cepas está apresentada na **Tabela 7**, juntamente com os dados do mapeamento feito por Chaikuad *et al.* para as quinases humanas.

Tabela 7 - Síntese do mapeamento das cisteínas das quinases de *T. cruzi* das cepas Dm28c 2014 e CL Brener El em sítios e subsítios, juntamente com dados do mapeamento de cisteínas de quinases humanas feito por Chaikuad *et al.*⁶²

Sítio	Subsítio	Resíduo PDB 1ATP	Nº de cisteínas		
			<i>T. cruzi</i> Dm28c 2014	<i>T. cruzi</i> CL Brener El	<i>Homo sapiens</i> ^a
Glycine-rich loop region (P-loop)	P1	L49	1	0	5
	P2	S53	1	1	10
	P3	F54	4	3	4
	P4	V57	2	2	11
DFG region	D1	T183	21	20	47
Roof region	R1	A70	3	1	5
Gatekeeper	G1	M120	0	0	2
Hinge region	H1	Y122	4	2	5
	H2	V123	37	41	87
	H3	A124	2	1	9
		G125	4	3	
Front region	F1	G126	1	2	3
	F2	E127	0	0	12
	F3	F129	3	3	3
	F4	R133	4	0	1
Outside ATP	S1	P169	0	0	3
Others	O1	V104	4	3	5
	O2	F187	9	6	44
	O3	A188	10	12	
		C199	32	31	26
Beta Sheet 4	Beta 4-4	L106	12	9	
P-Loop	P-Loop-5	G52	0	0	
Helix	Helix-9	L95	0	1	
	Helix-11	V98	10	10	
Catalytic Loop	Catalytic-2	L162	4	3	
	Catalytic-9	L172	3	10	
DFG Peptide	DFG-4	G186	1	0	
Total de cisteínas mapeadas			172	164	282
Total de quinases com cisteína(s)			112	100	
Total de quinases sem cisteína			59	75	

^a Dados de Chaikuad *et al.* (2018)⁶². As demais regiões sem indicação da quantidade de cisteínas em quinases humanas são regiões exclusivas do mapeamento de Zhao *et al.* (2017)⁶¹

Ao comparar o número de cisteínas mapeadas no quinoma humano e no de *T. cruzi* percebe-se que há uma semelhança na distribuição das cisteínas nos diversos sítios e subsítios, com a ressalva de que o quinoma humano (555 quinases³⁴) é cerca de duas vezes e meia maior que o de *T. cruzi*. Fica evidente na **Tabela 7** que as regiões com maior abundância de cisteínas são comuns aos dois organismos, como as regiões DFG, *Hinge* e *Others*. Este comparativo das quantidades de quinases em cada região priorizada pode ser útil na tentativa de se buscar uma seletividade na inibição das quinases de *T. cruzi*. Os subsítios P3, H1, F1, F3, H1 e O1 parecem promissores para este fim, visto que são posições com poucas quinases mapeadas em humanos e que em termos proporcionais representam uma maior fatia do quinoma de *T. cruzi*. Uma observação interessante deste mapeamento das cisteínas das quinases de *T. cruzi* é a ausência de cisteínas na posição F2 (*Front-2*) do sítio *Front region* (*Front of Pocket*). No estudo de Zhao *et al.* é apontado que das 43 estruturas que apresentaram contato covalente entre o inibidor e uma cisteína, 29 ocorreram nesta posição. A justificativa para este fato é que grande parte das quinases humanas inibidas nesta região são quinases do grupo TK (*tyrosine kinase*), que não está presente no quinoma dos tripanossomatídeos. Contudo, há cisteínas mapeadas no subsítio F3 deste mesmo sítio pelo mapeamento de Zhao e colaboradores, assim como nas posições F1 e F4 exclusivas do mapeamento de Chaikuad *et al.* (**Tabela 5**). Além da aparente relevância do sítio *Front region* e principalmente do subsítio F2 para uma inibição covalente, Zhao *et al.* elenca outras regiões que se sobressaem dentre as 17 mapeadas e que apresentaram cisteínas mapeadas no quinoma de *T. cruzi*, são elas: sítio *P-Loop*, ou seja, subsítios P3 (*P-Loop-7*), P4 (*P-Loop-9*), *P-Loop-5*, e subsítios R1 (*Roof-3*), *Catalytic-2* e D1 (*DFG-3*). Zhao e colegas ainda constataram que dos 119 inibidores covalentes de quinases levantados em bancos de dados, com informações quantitativas das interações entre as cisteínas e os inibidores que se ligavam na região do sítio do ATP, 118 se localizavam nestas posições destacadas como mais promissoras dentre as 17. Em relação a este ponto, Chaikuad *et al.* ampliam a discussão questionando que o foco e os esforços no desenvolvimento de inibidores de quinases têm sido restritos a um conjunto limitado de proteínas relacionadas com distúrbios oncológicos, justificando, por exemplo, o grande número de inibidores atuando da posição F2. Desta forma, diversos outros alvos e inibidores covalentes focados em outras regiões deixam de ser explorados.

Ainda analisando os dados da **Tabela 7**, ao comparar os resultados do mapeamento das duas cepas de *T. cruzi*, percebe-se que há uma significativa semelhança na distribuição das cisteínas, destoando um pouco nas posições F4 e *Catalytic-9*. Esta correspondência entre os mapeamentos

das duas cepas é um indicativo da conservação destas cisteínas nas sequências das quinases das diversas cepas. Possivelmente, em muitos casos, o inibidor de uma quinase de uma determinada cepa também promoverá a inibição de quinases correspondentes nas demais cepas. A partir deste mapeamento das cisteínas das quinases de *T. cruzi*, tem-se a base das informações necessárias para auxiliar no processo de priorização de algumas quinases como alvos terapêuticos contra este parasito e auxiliar no planejamento e desenvolvimento de inibidores covalentes contra estes alvos.

4.3 PRIORIZAÇÃO DAS QUINASES

Nesta etapa buscou-se selecionar as proteínas mais promissoras para atuarem como bons alvos terapêuticos no controle e/ou eliminação do parasito, com foco nas quinases da cepa CL Brener El. O primeiro filtro do conjunto de 175 quinases desta cepa foi baseado nos resultados dos trabalhos de silenciamento de quinases de Jones *et al.* (2014)³¹ e Fernandez-Cortes *et al.* (2017)³² conduzidos com a forma sanguínea do organismo *T. brucei*. No primeiro, Jones e colegas identificaram 42 linhagens cujo silenciamento afetou a proliferação do organismo, sendo que, destas, 24 apresentaram também um comprometimento no ciclo celular e 11 culminaram na morte do organismo. No segundo, Fernandez-Cortes e colaboradores destacaram 35 quinases que também foram apontadas por Jones e colegas, além de outras 14 que se mostraram relevantes para o organismo somente no estudo *in vivo*, totalizando 49 quinases. Na **Tabela 8** estão apresentadas as 49 quinases destacadas por Fernandez-Cortes e colaboradores juntamente com as 7 quinases destacadas apenas por Jones *et al.*, totalizando 56 PKs. A este conjunto de proteínas segue também classificação em grupos e famílias, nomes e a qualidade da perda de *fitness*, quando descritos em um ou em ambos os estudos mencionados. A classificação ORPHAN utilizada nestes estudos para as quinases sem identidade clara com ortólogos conhecidos em *Saccharomyces cerevisiae*, *Caenorhabditis elegans*, *Drosophila melanogaster* ou *Homo sapiens*³² foi substituída por *Unique* para manter a nomenclatura utilizada durante a classificação dos conjuntos de quinases de *T. cruzi*.

Tabela 8 - Quinases de *T. brucei* destacadas como importantes para a proliferação e/ou ciclo celular do parasito segundo Jones *et al.* (2014)³¹ e Fernandez-Cortes *et al.* (2017)³²

Quinases <i>T. brucei</i> (código TriTrypDB)	Classificação			Perda de <i>Fitness</i> ^c	
	Grupo	Família	Nome	Defeito de proliferação	Defeito no ciclo celular
Tb927.3.2440	AGC	(na)	AGC1	Morte	C
Tb927.7.5770		NDR	PK53	Lenta	K/C
Tb927.10.4940		NDR	PK50	Interrompida	C
Tb927.9.11100/ 9.11030 ^b		PKA	PKAC1/PKAC2	Lenta	C
Tb927.9.4910 ^b		(na)	PDK1	Lenta	Não
Tb927.3.5400	aPK	RIO	RIO1	Interrompida	Não
Tb927.11.850 ^a		(na)	Bud32	Perda de <i>fitness</i>	
Tb927.6.2840		RIO	RIO2	Interrompida	Não
Tb927.11.14680		PIKK	ATR	Interrompida	M
Tb927.1.1930		PIKK/FRAP	TOR4	Lenta	G1/S
Tb927.10.8420		PIKK/FRAP	TOR1	Lenta	U
Tb927.4.420		PIKK/FRAP	TOR2	Morte	Não
Tb927.7.2750	CAMK	(na)	CAMK3	Lenta	Não
Tb927.10.14770 ^a		(na)	CAMK1(AKB1)	Perda de <i>fitness</i>	
Tb927.7.6220		(na)	CAMK2	Interrompida	C
Tb927.2.1820 ^a		CAMKL	CAMKL	Perda de <i>fitness</i>	
Tb927.5.790/ 5.800*	CK1	CK1	CK1.1/CK1.2	Morte	K/C
Tb927.11.5340	CMGC	(na)	CMGC1	Lenta	M
Tb927.10.1070		CDK	CRK1	Lenta	G1/S
Tb927.2.4510		CDK	CRK9	Morte	K
Tb927.7.7360		CDK	CRK2	Lenta	Não
Tb927.10.4990		CDK	CRK3	Morte	M
Tb927.11.1180 ^b		CDK	CRK6	Lenta	Não
Tb927.11.12410		CLK	CLK1/KKT10	Morte	M/C
Tb927.7.3880 ^a		DYRK	DYRK1	Perda de <i>fitness</i>	
Tb927.10.13780		GSK	GSK3-short	Interrompida	M/C
Tb927.10.7780		MAPK	KFR1	Lenta	Não
Tb927.10.5140		MAPK	MAPK6 (ERK8)	Morte	C
Tb927.3.690		RCK	RCK1	Morte	C
Tb927.7.960 ^a		SRPK	SRPK1	Perda de <i>fitness</i>	
Tb927.6.4970		SRPK	SRPK2	Interrompida	Não
Tb927.11.10520 ^a		Unique	(na)	KKT2	Perda de <i>fitness</i>
Tb927.9.10920	(na)		PK6/KKT3	Interrompida	Não

Quinases <i>T. brucei</i> (código TriTrypDB)	Classificação			Perda de <i>Fitness</i> ^c	
	Grupo	Família	Nome	Defeito de proliferação	Defeito no ciclo celular
Tb927.3.3920 ^a		AUR	AUK2	Perda de <i>fitness</i>	
Tb927.9.1670		AUR	AUK3	Lenta	M/C
Tb927.11.8220		AUR	AUK1	Morte	M
Tb927.10.15300		CAMKK	CAMKK	Lenta	Não
Tb927.2.2430 ^a		CK2	CK2A2	Perda de <i>fitness</i>	
Tb927.9.14430		CK2	CK2A1	Interrompida	K
Tb927.4.5310	OTHER	NEK	NEK12.2 /RDK2	Morte	K/C
Tb927.6.5100 ^a		PLK	PLK-like	Perda de <i>fitness</i>	
Tb927.7.6310 ^a		PLK	PLK1	Perda de <i>fitness</i>	
Tb927.4.5180/ 8.7220*		TLK	TLK1/TLK2	Morte	M/C
Tb927.11.4470		ULK	ULK	Interrompida	K/C
Tb927.11.9190 ^a		VPS15	VPS15	Perda de <i>fitness</i>	
Tb927.4.3420 ^b		WEE	WEE1	Lenta	Não
**Tb927.9.6560	OTHER	NAK	-	Interrompida	Não
**Tb927.11.9290 ^a	Unique	(na)	-	Perda de <i>fitness</i>	
**Tb927.7.3210 ^b	Unique	(na)	-	Lenta	U
Tb927.10.2040		STE11	STE11/CDC15-like	Morte	Não
Tb927.10.10350 ^a		STE11	STE11/Bck1p	Perda de <i>fitness</i>	
Tb927.11.2040	STE	STE11	STE11a	Interrompida	Não
Tb927.10.14300 ^a		STE11	STE11/MRK1	Perda de <i>fitness</i>	
Tb927.8.5730		STE20	SLK1	Interrompida	Não
Tb927.9.12880 ^b		STE20	SLK2	Interrompida	Não
Tb927.11.14070 ^b		STE11	RDK1	Lenta	Não

^a Quinases destacadas apenas por Fernandez-Cortes *et al.*³².

^b Quinases destacadas apenas por Jones *et al.*³¹ que não levaram a perda de *fitness* significativa *in vivo*.

^c Os dados "perda de *fitness*" se referem às quinases que se destacaram no estudo *in vivo* de Fernandez-Cortes *et al.* e que não apareceram nos resultados de Jones *et al.* Todos os demais dados foram obtidos de Jones *et al.*

* Silenciamento duplo

** Pseudoquinases

(na) – *not assigned to a Family*

C, citocinese; M, mitose; K, cinetoplasto (*kinetoplast*) duplicação/segregação; U, não classificado (*unclassified*)

Após realização de BLAST para determinar as quinases ortólogas em *T. cruzi*, chegou-se num total de 46 quinases (**Tabela 9**). Nesta tabela, além da identificação dos códigos TriTrypDB das quinases de *T. cruzi*, estão apresentadas as respectivas classificações em grupo, família e nome e se há ou não presença de cisteínas mapeadas. Como referência, foi mantida a descrição do impacto do silenciamento das quinases de *T. brucei* na proliferação e no ciclo celular nos respectivos ortólogos de *T. cruzi*.

Tabela 9 - Quinases de *T. cruzi* ortólogas das quinases de *T. brucei* destacadas por Jones *et al.*³¹ e Fernandez-Cortes *et al.*³² em estudos de RNAi como importantes para a proliferação e/ou ciclo celular do parasito

Quinases <i>T. cruzi</i> (código TriTrypDB)	Classificação			Perda de <i>fitness</i> em <i>T. brucei</i> ^a		Presença Cys Mapeada
	Grupo	Família	Nome	Defeito de proliferação	Defeito no ciclo celular	
TcCLB.510291.40	AGC	NDR	-	Interrompida	C	Não
TcCLB.508277.20		(na)	PDK1	Lenta	Não	Sim
TcCLB.510141.4	aPK	RIO	RIO1	Interrompida	Não	Não
TcCLB.506925.260		(Na)	Bud32	Perda de <i>fitness</i>		Não
TcCLB.511753.70		RIO	RIO2	Interrompida	Não	Não
TcCLB.506223.120		PIKK	ATR	Interrompida	M	Não
TcCLB.510525.10	CAMK	(na)	CAMKL	Lenta	Não	Sim
TcCLB.506211.220		(na)	AKB1	Perda de <i>fitness</i>		Sim
TcCLB.506513.50		(na)	-	Interrompida	C	Sim
TcCLB.510257.130		CAMKL	CAMKL	Perda de <i>fitness</i>		Sim
TcCLB.508541.220	CK1	CK1	CK1.1	Morte	K/C	Não
TcCLB.508541.230			CK1.2			Não
TcCLB.508541.240			CK1.2			Não
TcCLB.504929.10			CK1.2			Não
TcCLB.508541.225			CK1.2			Não
TcCLB.504181.40	CMGC	CDK	CRK1	Lenta	G1/S	Não
TcCLB.509099.150		CDK	CRK9	Morte	K	Sim
TcCLB.510609.70		CDK	CRK2	Lenta	Não	Não
TcCLB.506677.20		CDK	CRK3	Morte	M	Não
TcCLB.511127.320		CLK	CLK1/KKT10	Morte	M/C	Sim
TcCLB.506869.60		DYRK	PK4	Perda de <i>fitness</i>		Sim
TcCLB.507993.80		GSK	GSK3	Interrompida	M/C	Sim
TcCLB.504167.30		MAPK	KFR1	Lenta	Não	Não
TcCLB.510295.50		MAPK	MAPK6	Morte	C	Não
TcCLB.509231.20		RCK	-	Morte	C	Não
TcCLB.477079.10		SRPK	-	Perda de <i>fitness</i>		Não
TcCLB.506945.220		SRPK	SRPK	Interrompida	Não	Não

Quinases <i>T. cruzi</i> (código TriTrypDB)	Classificação			Perda de <i>fitness</i> em <i>T. brucei</i> ^a		Presença Cys Mapeada	
	Grupo	Família	Nome	Defeito de proliferação	Defeito no ciclo celular		
TcCLB.510285.70	Unique		KKT2	Perda de <i>fitness</i>		Sim	
TcCLB.504125.90			-	Lenta	M	Não	
TcCLB.503685.10	OTHER	AUR	AUK2	Perda de <i>fitness</i>		Sim	
TcCLB.510349.80		AUR	AUK3	Lenta	M/C	Sim	
TcCLB.506715.10		AUR	AUK3			Sim	
TcCLB.503799.4		AUR	AUK1	Morte	M	Sim	
TcCLB.506505.40		CAMKK	-	Lenta	Não	Não	
TcCLB.503513.10		CK2	CK2A2	Perda de <i>fitness</i>		Não	
TcCLB.510761.60		CK2	CK2A1	Interrompida	K	Não	
TcCLB.506945.350		PLK	-	Perda de <i>fitness</i>		Não	
TcCLB.506513.160		PLK	PLK	Perda de <i>fitness</i>		Sim	
TcCLB.510597.9		TLK	TLK2	Morte	M/C	Não	
TcCLB.504089.54		ULK	-	Interrompida	K/C	Não	
TcCLB.507485.90		WEE	WEE1	Lenta	Não	Sim	
**TcCLB.504883.10		OTHER	NAK	-	Interrompida	Não	Sim
**TcCLB.510533.70		Unique	-	Lenta	U	Sim	
TcCLB.506525.140		STE11	MRK1	Perda de <i>fitness</i>		Sim	
TcCLB.506743.160	STE	STE20	-	Interrompida	Não	Sim	
TcCLB.510745.40		STE20	-	Interrompida	Não	Sim	

^a Os dados "perda de *fitness*" se referem às quinases que se destacaram no estudo *in vivo* de Fernandez-Cortes *et al.*³² e que não apareceram nos resultados de Jones *et al.*³¹ Todos os demais dados foram obtidos de Jones *et al.*

(na) – *not assigned to a Family*

** Pseudoquinases

C, citocinese; M, mitose; K, cinetoplasto (*kinetoplast*) duplicação/segregação; U, não classificado (*unclassified*)

Enquanto na relação das 56 quinases de *T. brucei* havia indicação de apenas uma CK1.2, em conjunto com uma CK1.1, e apenas uma AUK3, no organismo *T. cruzi* o BLAST retornou 4 quinases ortólogas CK1.2, além de uma CK1.1, e duas AUK3. Duas das quinases CK1.2 são cópias, TcCLB.508541.230 e TcCLB.508541.240, com identidade de 100%. A CK1.2 TcCLB.504929.10 possui uma identidade de 96,4% com as duas cópias e a CK1.2 TcCLB.508541.225 de 86,2%. Entre estas duas últimas a identidade é de 82,0%. As quinases CRK3 são cópias, apresentando uma identidade de 99,6%. Desconsiderando essa multiplicidade tem-se uma variedade de 41 quinases de *T. cruzi*. Além disso, dentre estas 41, há apenas uma quinase da família TLK, a TLK2, enquanto na relação das proteínas de *T. brucei*

há também uma TLK1 em conjunto com uma TLK2. Para 15 das 56 quinases de *T. brucei* não foram identificadas proteínas ortólogas na cepa CL Brener El de *T. cruzi*.

Em relação à existência de cisteínas mapeadas, segundo filtro utilizado para a priorização das quinases, 12 proteínas ortólogas no *T. cruzi* apresentaram uma ou mais cisteínas mapeadas em posições relevantes para uma abordagem de inibição covalente (**Tabela 9**). Dentre estas, 3 quinases, a CRK9, a CLK1 e a AUK1, ao serem silenciadas em *T. brucei* levaram à morte do organismo e foram quinases destacadas nos dois estudos considerados, cumprindo os critérios adotados no processo de priorização. Ademais, duas destas três quinases mencionadas, a CRK9 e a CLK1, em *T. cruzi*, apresentaram cisteína mapeada em uma das regiões mais promissoras segundo mapeamento construído por Zhao *et al.* (2017)⁶¹, o subsítio D1 do sítio DFG. A CLK1 também possui uma cisteína no subsítio F1, próximo ao subsítio F2 que se mostrou ser uma posição relevante para inibição de algumas quinases humanas. O subsítio F1, por apresentar proporcionalmente poucas cisteínas mapeadas no quinoma humano (mapeamento de Chaikuad *et al.*⁶²) em relação ao quinoma de *T. cruzi*, pode contribuir para o alcance de uma inibição seletiva. A CLK1 ainda apresenta mais uma cisteína mapeada, no subsítio O2, e a quinase CRK9 possui também uma cisteína no subsítio H2. A quinase AUK1 apresentou cisteínas mapeadas no subsítio H2 e no subsítio O3, nenhum pertencente às 5 regiões destacadas por Zhao e colaboradores dentre as 17 de seu mapeamento. Mas, ainda assim, esta quinase segue sendo um alvo interessante visto que o silenciamento em *T. brucei* levou à perda de *fitness* e à morte do organismo. Ainda que algumas cisteínas mapeadas estejam em posições em que haja certa abundância de cisteínas também no quinoma humano, existem outras estratégias para se buscar seletividade no processo de inibição, explorando as diferenças entre o sítio ativo das quinases parasitárias e humanas. Esta priorização foi pautada primariamente pela possível relevância biológica das quinases de *T. cruzi* a partir da extrapolação das observações apresentadas nos estudo para *T. brucei*, podendo existir diferenças nas quinases tidas como essenciais entre estes dois organismos. Esta é uma primeira abordagem não esgotando a exploração de novos critérios de seleção de outros alvos. Desta forma, estas 3 quinases foram selecionadas como alvos terapêuticos promissores para a eliminação do parasito *T. cruzi* e controle da doença de Chagas.

4.4 MODELAGEM COMPARATIVA DAS QUINASES PRIORIZADAS

Conforme mencionado, foram priorizadas e modeladas três quinases do *T. cruzi*, AUK1, CLK1 e CRK9. Os resultados e discussões acerca de cada modelagem estão apresentados separadamente para cada proteína.

4.4.1 AUK1

A AUK1 da cepa CL Brener El é identificada pelo código TcCLB.503799.4, no TriTrypDB, e pelo código Q4CMJ2, no Uniprot. Conforme anotação presente no TriTrypDB e na classificação feita no servidor kinase.com, esta é uma quinase ePK, pertencente à família Aurora do grupo *Other*. Segundo Fassolari *et al.* (2019), esta proteína é expressa nas formas epimastigota, amastigota e tripomastigota. Em tripomastigotas, foi reportada como possivelmente sendo expressa por curto período de tempo e estar relacionada com o processo de diferenciação celular, visto sua localização nuclear nos primeiros dias de infecção²⁸. Fassolari e colegas reforçam esta ideia, uma vez que em fase inicial de infecção a AUK1 também foi identificada no núcleo da forma amastigota intracelular, indicando possível atuação na diferenciação de tripomastigotas para amastigotas. Em infecção celular mais avançada, na forma amastigota, forma replicativa ativa do parasito, a AUK1 foi detectada nas extremidades do cinetoplasto, apontando para um possível papel desta quinase também na divisão celular²⁸. Semelhante à forma amastigota, em epimastigota a AUK1 foi detectada em dois locais diferentes do parasito, no núcleo e em associação com o fuso mitótico, durante a mitose, e nas extremidades do cinetoplasto na interfase²⁸. A superexpressão da quinase em epimastigotas levou a um atraso na duplicação celular, aumentando o tempo de conclusão da mitose e validando o papel desta proteína na divisão celular. Além disso, a superexpressão da AUK1 durante as fases G2/M levou a uma alteração na duplicação do cinetoplasto, que foi associada à perda de localização da AUK1 nas extremidades desta organela, validando também seu papel na iniciação da duplicação do cinetoplasto²⁸. As AUK1 de *T. cruzi* e de *T. brucei* são reportadas como homólogas funcionais da quinase Aurora B humana^{28,101}. Estudos sobre esta proteína em *T. brucei* também reportaram um papel importante desta quinase na divisão celular, porém com

algumas diferenças entre as formas procíclica (inseto) e tripomastigota sanguínea (hospedeiro vertebrado). Na forma procíclica é relatado que esta quinase está envolvida na montagem do fuso mitótico, com a mitose, e também com a citocinese e com a replicação das organelas, contudo sem gerar alterações morfológicas aparentes^{101,102}. O silenciamento da AUK1 nesta forma do parasito reteve as células na fase G2/M, na qual pôde-se observar a presença de dois corpos basais, dois flagelos, dois cinetoplastos segregados, cujo movimento de afastamento é interrompido logo antes da citocinese, e um núcleo alongado¹⁰¹. Já na forma tripomastigota sanguínea, o silenciamento da AUK1 gerou defeitos na formação do fuso mitótico, na segregação cromossômica e na citocinese, porém com ocorrência de contínua progressão no ciclo do cinetoplasto associada ao bloqueio da citocinese¹⁰². Estas alterações, diferente do observado para a forma procíclica, levaram a replicação continuada de organelas, com sucessiva síntese de DNA nuclear e multiplicação de nucléolo, ocasionando alteração morfológica com células redondas alargadas¹⁰². A quinase AUK1 de *T. brucei* se associa a um complexo proteico chamado *Chromosomal Passenger Complex* (CPC), constituído de duas outras proteínas, a TbCPC1 e a TbCPC2A¹⁰³. O silenciamento individual destas proteínas também gerou comprometimento na montagem do fuso, na segregação cromossômica e na citocinese, comprometendo também a atividade da quinase Aurora¹⁰³. A AUK1 de *Leishmania major*, identificada como Lmairk, ainda não foi estudada quanto à sua função biológica, foi feita apenas sua identificação e clonagem¹⁰⁴. Mas, a presença deste gene é uma evidência de uma possível relação desta proteína com a divisão celular também neste organismo. Em conjunto, estes estudos indicam a importância de AUK1 no ciclo de vida de tripanossomatídeos.

No banco de dados TriTrypDB, a AUK1 CL Brener El está anotada como um fragmento de apenas 206 resíduos de aminoácidos, não tendo parte da porção do lóbulo C rico em α -hélices característico nas quinases. Para melhor compreender esta sequência antes de iniciar a modelagem, foi feito um BLAST no próprio TriTrypDB para identificar as proteínas correspondentes em outras cepas do *T. cruzi*. Em todas as cepas analisadas as sequências apresentavam 309 resíduos de aminoácidos e uma alta identidade com a quinase em estudo (**Tabela 10**). Além disso, as mesmas cisteínas mapeadas como relevantes para uma inibição covalente na AUK1 da cepa CL Brener El também estão presentes nas outras cepas analisadas (**Tabela 10**). Ao conferir o trecho do genoma subsequente à extremidade C terminal truncada do fragmento da AUK1 da cepa CL Brener El, verifica-se que se trata de falha no sequenciamento genômico. Considerando estas informações, é provável que o tamanho original desta sequência seja também de 309 resíduos. Sendo assim, optou-se por conduzir a modelagem

comparativa utilizando a quinase CL Brener Non-Esmeraldo-like - CL Brener NEI (TriTrypDB: TcCLB.508817.80).

Tabela 10 - Comparativo do número de resíduos, identidade e número de cisteínas mapeadas entre as sequências de quinases AUK1 de diversas cepas de *T. cruzi* com a sequência AUK1 da cepa CL Brener Esmeraldo-like (CL Brener EI)

Cepa	Código TryTripDB	Nº Resíduos	Identidade com AUK1 de CL Brener EI	Cisteínas mapeadas
CL Brener EI	TcCLB.503799.4	206	—	
CL Brener NEI	TcCLB.508817.80	309	99%	
Dm28c 2014	TCDM_04589	309	99%	
Dm28c 2018	C4B63_6g509	309	99%	
Y C6	TcYC6_0013480	309	100%	
Sylvio X10/1-2012	TCSYLVIO_002533	309	99%	
Brazil A4	TcBrA4_0079430	309	99%	C55 (Roof A3) C109 (Hinge H2) C186 (Others O3)

O primeiro passo para a modelagem da AUK1 foi a busca por moldes para esta sequência. Abaixo segue **Tabela 11** contendo os possíveis moldes considerados, juntamente com os dados empregados para a avaliação e escolha da estrutura molde.

Tabela 11 - Dados dos candidatos a moldes para a modelagem comparativa da quinase AUK1 da cepa CL Brener NEI de *T. cruzi*

BLAST NCBI						
Posição	Código PDB e Resolução /Å	Identidade e Cobertura /%	Nº Resíduos Faltantes e localização ^d		Presença de	
					Fosforilação	Peptídeo ^e
1	3D14 - 1,90 ^a	44,1 - 85	6	Alça de ativação	Não	Não
2	3DAJ - 2,00 ^a	44,1 - 85	1 3	Alça de ativação	Não	Não
3	2WTW - 3,30 ^b	42,9 - 88	6	Alça de ativação	Não	Não
4	6C83 - 2,55 ^b	42,4 - 89	1 5	Alça de ativação	Não	Não
5	1MUO - 2,90 ^b	42,4 - 90	1 0	Alça de ativação	Não	Não

BLAST NCBI						
Posição	Código PDB e Resolução /Å	Identidade e Cobertura /%	Nº Resíduos Faltantes e localização ^d		Presença de	
					Fosforilação	Peptídeo ^e
6	2X6D - 2,80 ^b	42,9 - 88	4	Entre fita β 3 e α C	Não	Não
			4	Alça de ativação		
7	2XNG - 2,60 ^b	42,9 - 88	9	Alça de ativação	Não	Não
8	4J8M - 1,85 ^b	43,0 - 88	-	-	Não	Não
9	4C3P - 2,69 ^b	42,9 - 88	-	-	Não	Sim
10	2J4Z - 2,00 ^b	42,4 - 90	-	-	Não	Não
11	5ZAN - 2,85 ^b	43,0 - 88	5	Alça de ativação	Não	Não
12	3UNZ - 2,80 ^b	43,0 - 88	-	-	Não	Não
13	2J50 - 3,00 ^b	43,8 - 88	8	Alça de ativação	Não	Não
			4	Logo antes da α F		
14	4JAI - 3,20 ^b	42,9 - 88	1 6	Alça de ativação	Não	Não
15	3W16 - 2,80 ^b	43,7 - 86	1	Alça de ativação	Não	Não
			1	Logo antes da α F		
34	1OL5 - 2,50 ^b	42,5 - 88	-	-	Sim	Sim
71	2BFY - 1,80 ^c	41,5 - 85	-	-	Sim	Sim

Swiss-Model						
Posição	PDB Resolução /Å	Identidade e Cobertura /%	Nº Resíduos Faltantes e localização ^d		Presença de	
					Fosforilação	Peptídeo ^e
1	5K3Y - 1,60 ^c	40,9 - 87	1	Alça de ativação	Sim	Sim
2	2BFY - 1,80 ^c	40,5 - 87	-	-	Sim	Sim
3	6MM5 - 1,95 ^a	27,7 - 94	-	-	Sim	Sim
4	4DFZ - 2,00 ^a	27,8 - 93	-	-	Sim	Sim
5	4C2W - 1,70 ^c	40,7 - 87	-	-	Sim	Sim

^a Organismo *Mus musculus*

^b Organismo *Homo sapiens*

^c Organismo *Xenopus laevis*

^d Os dados indicados correspondem ao número de resíduos ausentes na estrutura cristalográfica dentro do intervalo do domínio característico das quinases.

^e Trecho de proteína ativadora associada à quinase

Para todos os candidatos a molde havia a presença de inibidor no sítio do ATP. Desta forma, esta informação foi omitida na tabela acima. O interesse por moldes com a presença de molécula inibidora se deve ao fato destas moléculas gerarem ajustes induzidos no sítio ativo, deixando a estrutura da proteína candidata a molde numa conformações adaptada importante para as etapas futuras de realização de *docking* e busca por potenciais inibidores. É notória a dificuldade de se encontrar um molde que contemple todos os aspectos avaliados de forma satisfatória. Praticamente todos os 15 primeiros resultados retornados no BLASTp do NCBI apresentam resíduos faltantes em região importante da quinase, a alça de ativação. Esta é uma alça longa com alto grau de flexibilidade⁴⁷, justificando uma maior dificuldade para observar a densidade eletrônica corresponde a esta região em estruturas cristalográficas. Nos poucos casos em que a sequência estava completa, ou não se tinha uma boa resolução, ou não se tinha a presença de fosforilação ou peptídeo associado. Tendo em vista esta dificuldade em se encontrar um molde adequado, fez-se uma busca no servidor Swiss-Model. Dentre os 5 primeiros resultados, 4 apresentaram diversos dos requisitos de interesse, como não possuir resíduos faltantes, possui inibidor, fosforilação e peptídeo associado atuando na ativação da quinase. A estrutura 2BFY foi selecionada como molde, por possuir boa resolução (1,8 Å) e estar ranqueada em segundo lugar. Nesta estrutura, a proteína está na forma semiativada, mas ainda assim passível de se ligar a inibidores no sítio de ligação do ATP¹⁰⁵. Além deste molde, considerou-se também a estrutura de outro molde na forma ativa, de código PDB 1OL5, reportada por Sessa *et al.* (2005)¹⁰⁵. Sendo assim, optou-se pela realização de duas modelagens, uma utilizando o molde na forma semiativa, denominada AUK1-2BFY, e outra utilizando o molde na forma ativa, chamada de AUK1-1OL5.

4.4.1.1 AUK1-2BFY

A estrutura do molde de código PDB 2BFY¹⁰⁵ corresponde a uma Aurora B do organismo *Xenopus laevis*, contendo também um peptídeo referente a um trecho da proteína ativadora *Inner centromere protein* (INCENP). A sequência proteica empregada na cristalografia da Aurora molde não se referia à sequência completa, do aminoácido 1 ao 361, mas sim apenas do resíduo 78 ao 361. Além disso, trechos das porções N e C terminais deste intervalo não foram elucidadas e, no fim, a estrutura obtida correspondia do resíduo 86 ao 356, mas ainda sim cobrindo todo o domínio de quinases. O molde 2BFY apresenta uma fosforilação na treonina

248, localizada na alça de ativação, e possui uma mutação accidental no resíduo 96, onde houve a troca de uma glicina por uma valina. Segundo Sessa *et al.* (2005), esta mutação não interferiu nas características bioquímicas e biofísicas da quinase e não comprometeu a sua atividade catalítica¹⁰⁵. A identidade entre a sequência desta estrutura molde, de resolução 1,80 Å, com a sequência da AUK1 CL Brener NEI foi de 41,5% e a cobertura entre ambas foi de 85%, de acordo com BLASTp. A partir destas informações fez-se o alinhamento da sequência do molde selecionado com a sequência da AUK1 no ClustalΩ (**Figura 15**). Neste alinhamento o trecho inicial da AUK1 de *T. cruzi*, do resíduo 1 ao 21, e o trecho final, do resíduo 294 ao 309, ficaram sem cobertura. Contudo, este fato não comprometeu o domínio característico de quinases. A proteína ficou alinhada com o molde do resíduo 22 ao 293 e este foi o trecho utilizado na modelagem comparativa conduzida no programa Modeller.

```

TcCLB.508817.80(AUK1)  MSAAEQQVWASYIALPKAPRSQKASDFEMLHKLGGGNYGDVYLASVRDCNFVCAIKKL 60
2BFY                    -----RKFTIDDFDIVRPLGKGFVYLAAREKQNKFIMALKVL 47
                        .. .**::: ** *:::**** :: :*: ** *

TcCLB.508817.80(AUK1)  SIKKLAEFDIVAQLRREIEIAFHTRHRYLLRITYGYFFDDTDIYLILEPCSNGLYTELNR 120
2BFY                    FKSQLEKEGVEHQLRREIEIQSHLRHPNILRMVNYFHDRKRIYLMLEFAPRGELYKELQK 107
                        .:* : .: ***** * ** :** *.*.* . ***:** . * **.*:::

TcCLB.508817.80(AUK1)  VKCFPPPTAARYVAQLAEALLYLHQHILHRDIKPENILLDHNQNIKLADFGWSVHDPHN 180
2BFY                    HGRFDEQRSATFMEELADALHYCHERKVIHRDIKPENLLMGYKGGELKIADFGWSVHAPSL 167
                        *   :* :: :**:* * *:::*****:*::: :*:***** *

TcCLB.508817.80(AUK1)  RRKTSCGTPEYFPPEIVGRQPYDTSADLWCLGIFCYELLVGKTPFVSKDTENICKKIHAM 240
2BFY                    RRRTMCGTLDYLPPEMIEGKTHDEKVDLWCAGVLCYFLVGMPPFDSPSHTETHRRIVNV 227
                        **.* ** .*:***: : :* ..*** *::***:** ** * . : :*: :

TcCLB.508817.80(AUK1)  QYTIPDTPPEAKELISSLLLRDGGKRLALHRVWVNHFPFLKYYYPNGIQPPRGKRMGA 300
2BFY                    DLKFPFLSDGSKDLISKLLRYHPPQRLPLKGVMEHPWVKANSRRVLPVYQS----- 284
                        : :* : :*:***:** . :** * :*:***:

TcCLB.508817.80(AUK1)  ADFSGGKEN 309
2BFY          -----| 284

```

Figura 15 - Alinhamento produzido no servidor ClustalΩ entre a sequência da quinase AUK1 de *T. cruzi* da cepa CL Brener NEI e a sequência do molde (código PDB 2BFY) para execução da modelagem comparativa.

Durante a modelagem foram gerados 100 modelos dos quais foram pré-selecionados os 3 melhores pontuados em relação ao DOPE e ao Molpdf. Estes 3 modelos e suas estruturas refinadas foram validados quanto a seus aspectos estéricos e energéticos (**Tabela 12**). Após análises dos dados apresentados, o modelo selecionado foi o AUK1.26 refinado no servidor LocPREFMD. Esta estrutura apresentou resultados de validação satisfatórios e coerentes com

a validação do molde 2BFY, tendo 96,3% dos resíduos nas regiões mais favoráveis e 100,0% em regiões permitidas no diagrama de Ramachandran do servidor MolProbity (**Figura 16**). Além de melhorar os valores retornados pelo Errat, de 62,9 para 94,2, o processo de refinamento reduziu de forma significativa os *clashes* presentes. No gráfico gerado no ProSA (**Figura 16**) nota-se que o Z-score está dentro da faixa de valores correspondentes a estruturas experimentais de proteínas de tamanho semelhante. Além disso, o modelo AUK1.26 apresentou alguns rotâmeros *outliers*, mas nenhum dentro do sítio de ligação do ATP.

Tabela 12 - Dados de validação dos 3 modelos pré-selecionados da modelagem da quinase AUK1 de *T. cruzi* da cepa CL Brener NEI (AUK1-2BFY) gerados no programa Modeller

Modelos	Modeller		PyMOL/ Å	ProSA	Ramachandran MolProbity /%			Saves / %	
	Dope Score	Molpdf	RMSD	Z-score	Favorável	Permitido	Não permitido	Errat	Verivy 3D
AUK1.72	-31589,5	1696,5	0,104 ^b	-7,26	96,3	99,6	0,4	62,9	94,9
Refinado ^a			0,201 ^c	-7,59	94,8	99,3	0,7	87,9	92,7
AUK1.36	-31861,5	1700,3	0,096 ^b	-7,13	95,2	100,0	0,0	54,2	93,0
Refinado ^a			0,206 ^c	-7,33	97,4	99,6	0,4	90,4	92,3
AUK1.26	-31533,3	1720,2	0,107 ^b	-7,25	95,6	98,9	1,1	62,9	96,3
Refinado ^a			0,331 ^c	-7,24	96,3	100,0	0,0	94,2	91,5
Molde 2BFY				-6,90	96,3	99,6	0,0	93,1	97,0

^aRefinamento realizado do servidor LocPREFMD (<http://feig.bch.msu.edu/web/services/locprefmd/>)

^bRMSD calculado em relação ao molde 2BFY

^cRMSD calculado em relação ao modelo inicial não refinado

DOPE - *Discrete Optimized Protein Energy*. Função que permite avaliar a energia dos modelos gerados, cujos resultados com os menores valores correspondem aos modelos hipoteticamente mais próximos à estrutura nativa.

Molpdf - *Molecular probability density function*. Parâmetro que corresponde à soma das contribuições de todas as violações ocorridas. Quanto menor o valor, melhor é a qualidade do modelo gerado.

RMSD - *Root Mean Square Deviation*. Valores de RMSD menores ou iguais a 2Å correspondem a uma boa sobreposição estrutural entre modelo e molde.

Z-score. Indica a qualidade geral da estrutura da proteína analisada ao se comparar com o valor de Z-score de proteínas de tamanho semelhante resolvidas experimentalmente.

Diagrama de Ramachandran. Avalia os ângulos ϕ (phi) ψ (psi). Pelo servidor MolProbity, uma estrutura de qualidade possui 98,0% ou mais dos resíduos em regiões favoráveis, 99,8% ou mais em regiões permitidas e até 0,5% em regiões não permitidas.

Errat. Fator de qualidade geral que corresponde à porcentagem da proteína para a qual o valor do erro fica abaixo do limite de rejeição de 95%. Estruturas com resoluções altas costumam apresentar um fator de qualidade de 95% ou superior e estruturas de resoluções mais baixas apresentam um fator de qualidade geral em torno de 91%.

Verify 3D. Verifica a compatibilidade da estrutura tridimensional das proteínas com as informações presentes nas suas sequências de aminoácidos. Quando pelo menos 80% dos resíduos em análise apresentam uma pontuação maior ou igual a 0,2 há um indício de que a estrutura terciária da proteína está adequada.

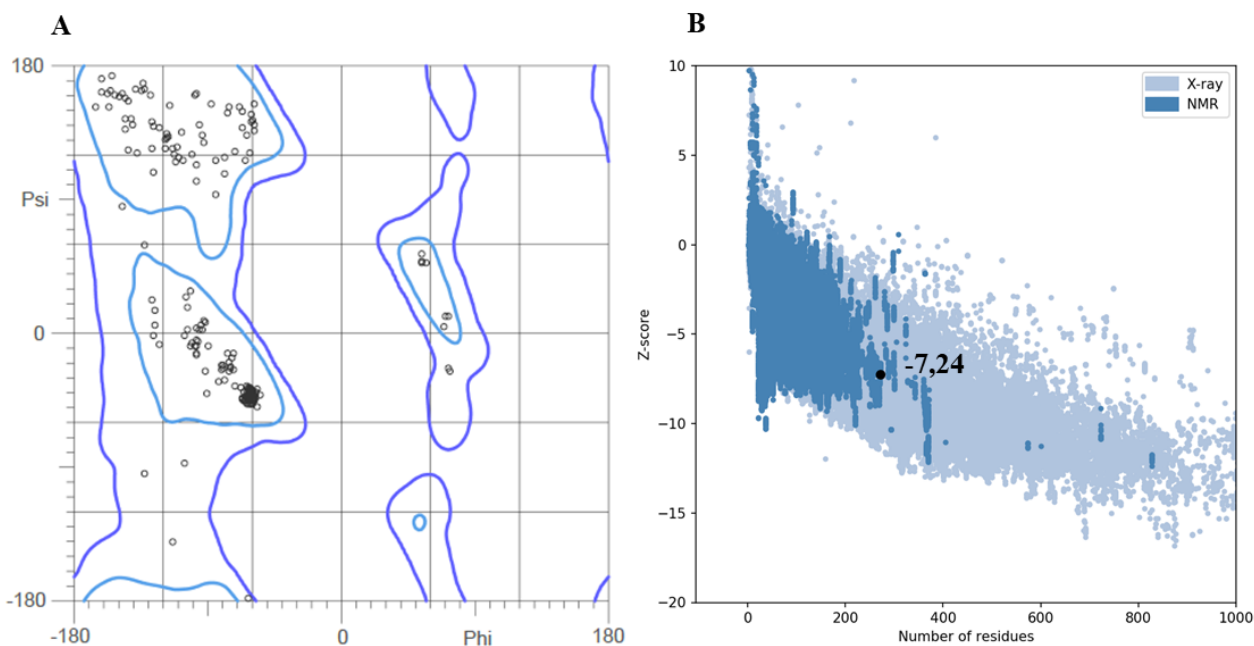


Figura 16 - Validação do modelo AUK1.26 refinado construído por modelagem comparativa para a proteína AUK1 de *T. cruzi* da cepa CL Brener NE1, utilizando o molde 2BFY. **A.** Diagrama de Ramachandran gerado no servidor MolProbity para o modelo AUK1.26 refinado. Contornos em azul escuro indicam regiões permitidas, enquanto contornos em azul claro indicam as regiões favoráveis. **B.** Gráfico gerado no servidor ProSA, destacando o valor de Z-score para o modelo AUK1.26 refinado em relação a estruturas determinadas experimentalmente com número de resíduos semelhantes.

A estrutura do modelo selecionado está ilustrada na **Figura 17**, na qual se destaca a sobreposição com a estrutura do molde utilizado, além dos resíduos catalíticos (AUK1: K58, D152 e D170) e cisteínas mapeadas (AUK1: C109 (Hinge H2) e C186 (Other O3)). Observa-se uma boa sobreposição entre a estrutura do modelo e do molde e entre os resíduos catalíticos das duas estruturas. Quanto à cavidade que comporta o ATP, esta se encontra preservada e coerente com a estrutura do molde 2BFY.

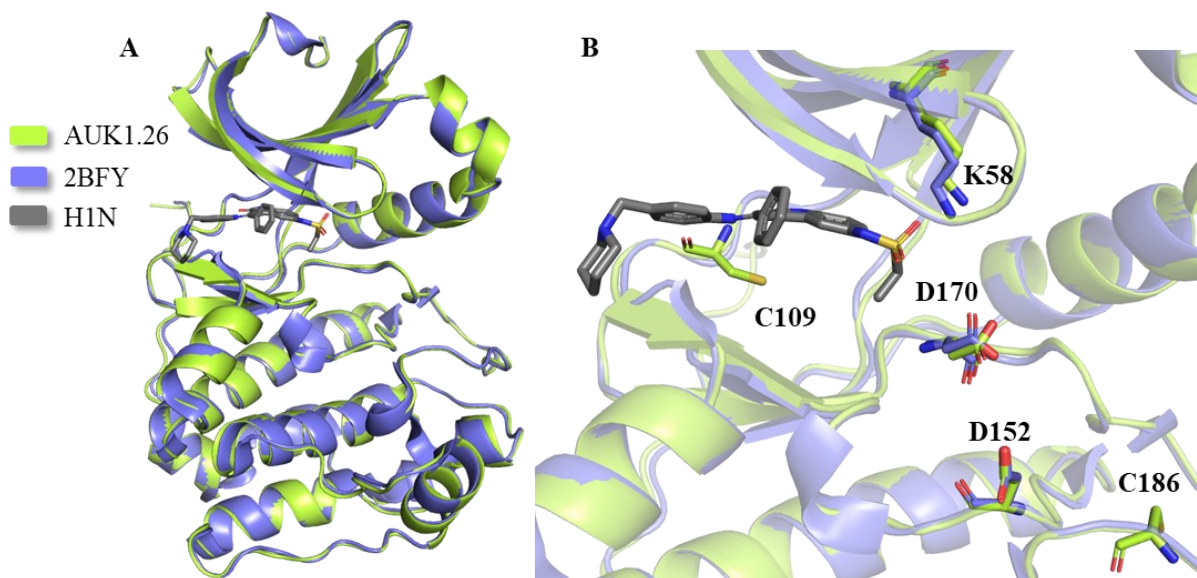


Figura 17 - Estrutura tridimensional do modelo AUK1.26 refinado da quinase AUK1 de *T. cruzi* da cepa CL Brener NE1 gerado por modelagem comparativa em sobreposição com o molde 2BFY, com destaque para os resíduos catalíticos e cisteínas mapeadas. **A.** Sobreposição do modelo AUK1.26 com o molde 2BFY, com a presença do inibidor H1N. **B.** Destaque para os 3 resíduos catalíticos da AUK1, K58, D152 e D170, e para as 2 cisteínas mapeadas C109 (Hinge H2) e C186 (Other O3).

Com o intuito de melhor compreender as diferenças na região do sítio do ATP entre o modelo e o molde utilizado, foram feitas algumas análises comparativas. Dos 30 resíduos que se encontram a até 6 Å de distância do inibidor H1N, tanto no modelo quanto no molde, 19 são iguais, ou seja, tem-se uma identidade de 63,3% nesta região. Grande parte dos resíduos diferentes (**Figura 18**) estão em áreas expostas ao solvente e não comprometem estericamente o sítio. No modelo estes resíduos diferentes são em sua maioria de tamanhos semelhantes ou menores que os resíduos correspondentes no molde, e são apolares ou polares não carregados, com exceção de uma lisina, enquanto que no molde existe uma maior proporção de resíduos carregados, principalmente positivamente, como pode ser observado na **Figura 18**, painel B. Diante do que foi exposto e discutido sobre a modelagem da quinase Aurora com o molde 2BFY, conclui-se que a estrutura do modelo AUK1.26 refinado está adequada e apresenta qualidade para ser utilizada em estudos futuros de busca por inibidores.

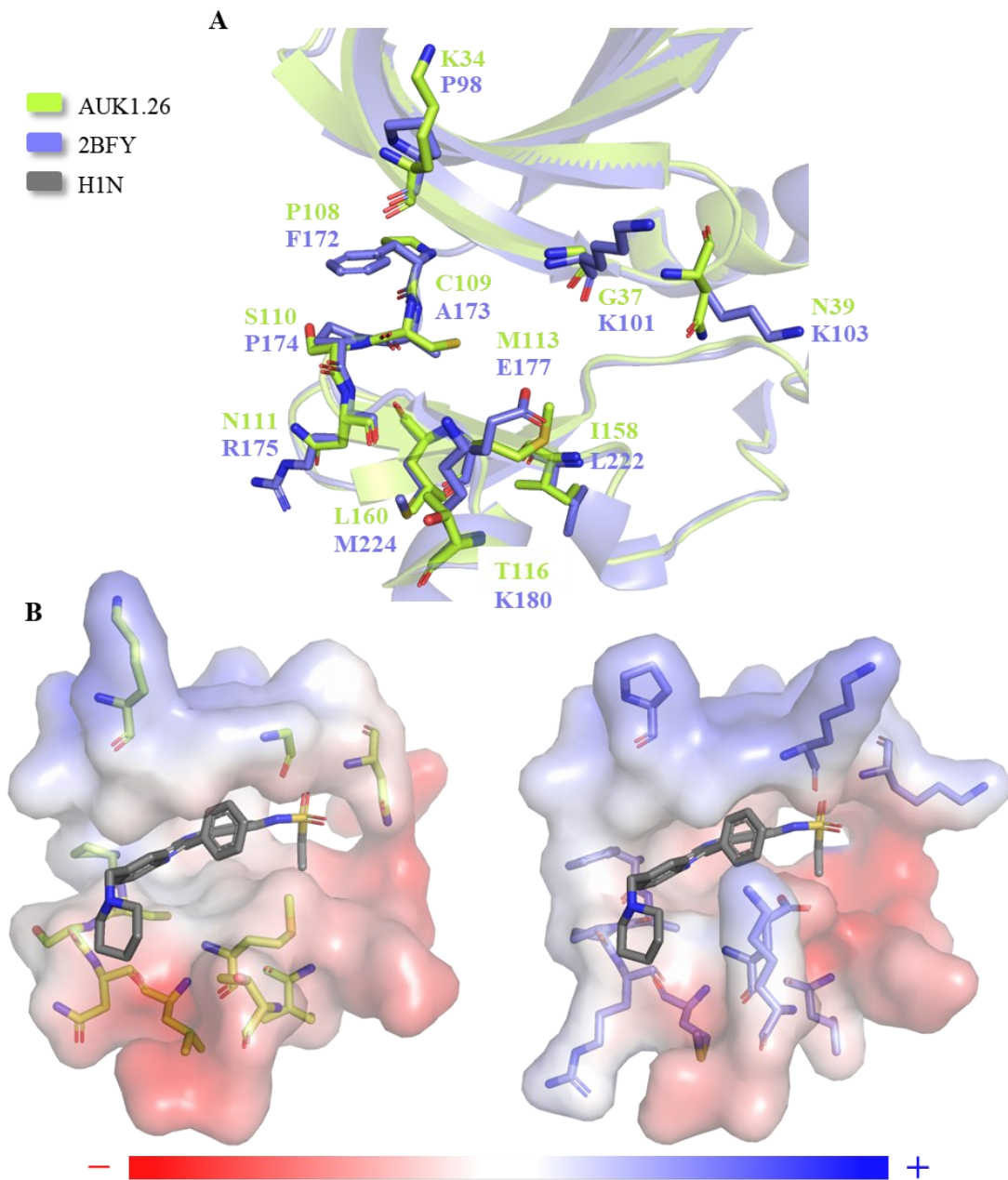


Figura 18 - Destaque das diferenças no sítio de ligação do ATP entre a estrutura do modelo AUK1.26 refinado da quinase AUK1 de *T. cruzi* da cepa CL Brener NE1 e a estrutura do molde 2BFY. **A.** Destaque em bastões dos resíduos diferentes entre o modelo AUK1.26 refinado e o molde 2BFY que se encontram a até 6 Å do inibidor H1N. A estrutura da molécula H1N e parte da estrutura tridimensional foram omitidas para melhor visualização da região do sítio. **B.** Superfície eletrostática da região analisada para o modelo (à esquerda) e o molde (à direita) gerada pelo *plugin* APBS do PyMOL^{106,107}. A coloração em vermelho indica regiões de potencial eletrostático negativo e em azul regiões de potencial positivo.

4.4.1.2 AUK1-1OL5

Para modelagem da AUK1 da cepa CL Brener NEI na forma ativa, foi empregado o molde de código PDB 1OL5¹⁰⁸, uma Aurora A de *Homo sapiens* no estado ativo que está associada a um trecho da proteína ativadora TPX2. Esta Aurora A é uma proteína de 403 aminoácidos. Porém, a construção utilizada para a cristalografia continha apenas o domínio característico de quinases (aminoácido 122 ao 403), e a estrutura obtida a partir do mapa de densidade eletrônica continha do resíduo 123 ao 388. Este molde possui duas fosforilações na alça de ativação, nas treoninas 287 e 288, e não apresenta nenhuma mutação na sequência proteica. A estrutura cristalográfica PDB 1OL5, de resolução 2,50 Å, apresentou uma identidade e cobertura em relação à sequência a ser modelada de 42,5% e 88%, respectivamente, conforme BLASTp. O alinhamento da sequência da estrutura da 1OL5 com a sequência da quinase AUK1 de *T. cruzi* no ClustalΩ se encontra na **Figura 19**. Em relação a este molde, a AUK1 ficou sem cobertura nos 18 primeiros resíduos e nos 25 últimos. Contudo, a proteína ficou coberta em toda a extensão do domínio de quinase, do resíduo 19 ao 284.

```

TcCLB.508817.80(AUK1)      MSAAEGGQVVASYIALPKAPRSDWKASDFEMLHKLGGGNYGDVYLASVRDCNFVCAIKKL 60
1OL5                      -----SKKRQWALEDFEIGRPLGKGFVNYLAREKQSKFILALKVL 42
                          : : :* .***: : ** *::*:*** :::*: ** *

TcCLB.508817.80(AUK1)      SIKKLAEFDIVAQLRREIEIAFHTRHRYLLRITYGYFFDDTDIYLILEPCSNGLYTELNR 120
1OL5                      FKAQLEKAGVEHQLRREVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQK 102
                          :* : .: *****:* * ** :** ***,* * :***** . * :* **::

TcCLB.508817.80(AUK1)      VKCFPPPTAARYVAQLAEALLYLHQHILHRDIKPENILLDHNQNIKLADFGWSVHDPHN 180
1OL5                      LSKFDEQRTATYITELANALSYCHSKRVIHRDIKPENLLGSAGELKIADFGWSVHAPSS 162
                          :. * :* *::*:** * * :::*:*****:**. :*:***** * .

TcCLB.508817.80(AUK1)      RRKTSCGTPPEYFPPEIVGRQPYDTSADLWCLGIFCYELLVGKTPFVSKDTENICKKIHAM 240
1OL5                      RR--LCGTLDYLPPEMIEGRMHDEKVDLWSLGVLCYEFVVGKPPFEANTYQETYKRISRV 220
                          ** *** :*:***: : :* ..***.*:***:*** ** : : :*:* :

TcCLB.508817.80(AUK1)      QYTIPDTPPEAKELISSLLLRDGGKRLALHRVNHFPFLKYYYPNGIQPPRGKRMARGA 300
1OL5                      EFTFPDFVTEGARDLISRLKHNPSQRPMLREVLEHPWITANSS----- 264
                          :*:** * *:**** ** : : :* *:.*:***:

TcCLB.508817.80(AUK1)      ADFSGGKEN 309
1OL5                      ----- 264

```

Figura 19 - Alinhamento produzido no servidor ClustalΩ entre a sequência da quinase AUK1 de *T. cruzi* da cepa CL Brener NEI e a sequência do molde (código PDB 1OL5) para execução da modelagem comparativa.

Após o alinhamento foi feita a modelagem e a validação dos modelos melhores pontuados, assim como das respectivas estruturas refinadas (**Tabela 13**). Selecionou-se o modelo AUK1.56 refinado como o mais adequado. Este modelo apresentou uma boa qualidade estrutural, com 97,0% dos resíduos em regiões favoráveis e 100,0% em regiões permitidas, pelo diagrama de Ramachandran gerado no MolProbity (**Figura 20**). O refinamento conduzido no servidor LocPREFMD trouxe uma melhora nos valores de Errat, 89,5, e Verify 3D, 92,9, assim como nos *clashes*. O Z-score obtido no ProSA ficou coerente com as estruturas experimentais de proteínas de tamanho semelhante (**Figura 20**). Este modelo apresentou uma melhor qualidade quanto aos rotâmeros em comparação com o modelo selecionado na modelagem AUK1-2BFY, e também não apresentou rotâmeros *outliers* no sítio de ligação do ATP.

Tabela 13 - Dados de validação dos 3 modelos pré-selecionados da modelagem da quinase AUK1 de *T. cruzi* da cepa CL Brener NE1 (AUK1-1OL5) gerados no programa Modeller

Modeller		PyMOL/ Å	ProSA	Ramachandran MolProbity /%			Saves / %		
Modelos	Dope Score	Molpdf	RMSD	Z-score	Favorável	Permitido	Não permitido	Errat	Verivy 3D
AUK1.62	-33602,3	1285,2	0,108 ^b	-7,27	96,6	100,0	0,0	83,9	92,1
Refinado ^a			0,202 ^c	-7,34	97,7	100,0	0,0	96,4	84,2
AUK1.83	-33586,3	1377,5	0,128 ^b	-7,41	96,2	100,0	0,0	84,9	100,0
Refinado ^a			0,190 ^c	-7,39	97,0	100,0	0,0	87,6	92,9
AUK1.56	-33542,3	1415,2	0,136 ^b	-7,30	97,0	99,6	0,4	83,2	91,7
Refinado ^a			0,272 ^c	-7,45	97,0	100,0	0,0	89,5	92,9
Molde 1OL5				-6,76	95,8	100,0	0,0	97,2	94,7

^aRefinamento realizado do servidor LocPREFMD (<http://feig.bch.msu.edu/web/services/locprefmd/>)

^bRMSD calculado em relação ao molde 1OL5

^cRMSD calculado em relação ao modelo inicial não refinado

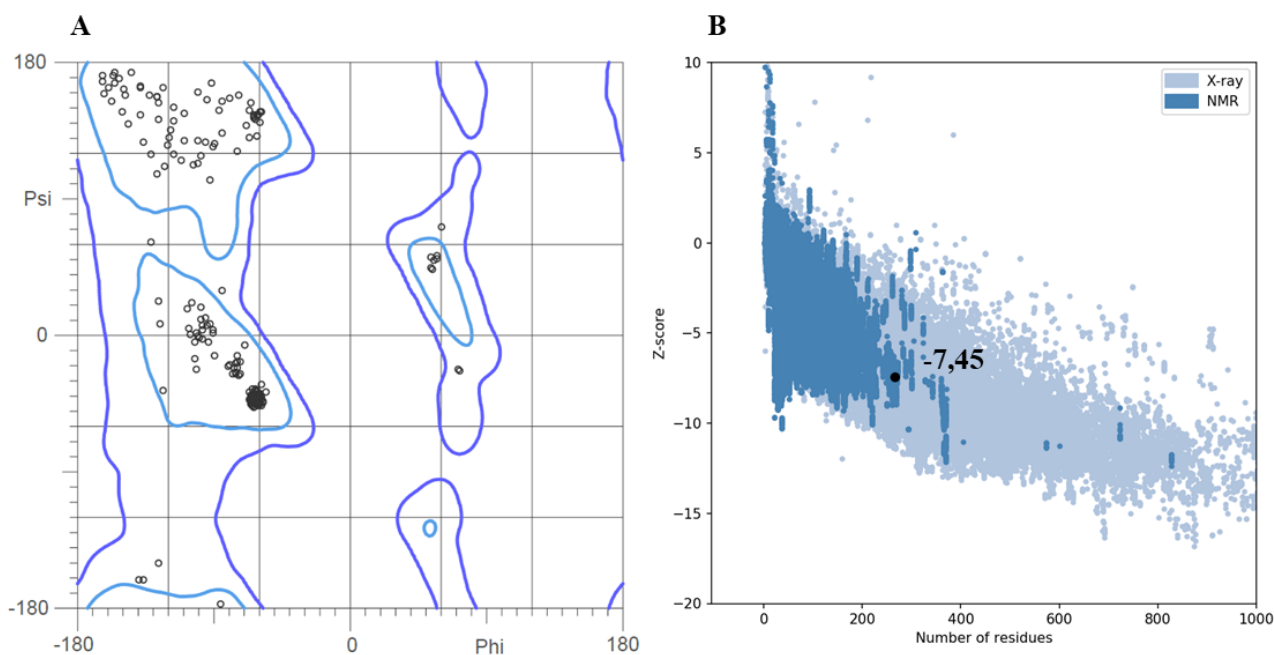


Figura 20 - Validação do modelo AUK1.56 refinado construído por modelagem comparativa para a proteína AUK1 de *T. cruzi* da cepa CL Brener NE1, utilizando o molde PDB 1OL5. **A.** Diagrama de Ramachandran gerado no servidor MolProbity para o modelo AUK1.56 refinado. Contornos em azul escuro indicam regiões permitidas, enquanto contornos em azul claro indicam as regiões favoráveis. **B.** Gráfico gerado no servidor ProSA, destacando o valor de Z-score para o modelo AUK1.56 refinado em relação a estruturas determinadas experimentalmente, com número de resíduos semelhantes.

Na **Figura 21** está representada a sobreposição das estruturas do modelo da quinase AUK1 e do molde 1OL5, além dos resíduos catalíticos (AUK1: K58, D152 e D170) e cisteínas mapeadas (AUK1: C109 (Hinge H2) e C186 (Other O3)). A sobreposição destas estruturas no PyMOL retornou um RMSD de 0,335 Å, indicando alta similaridade entre as estruturas terciárias. Nesta modelagem também foram transferidos para o modelo dois íons magnésio presentes na região do sítio do ATP do molde.

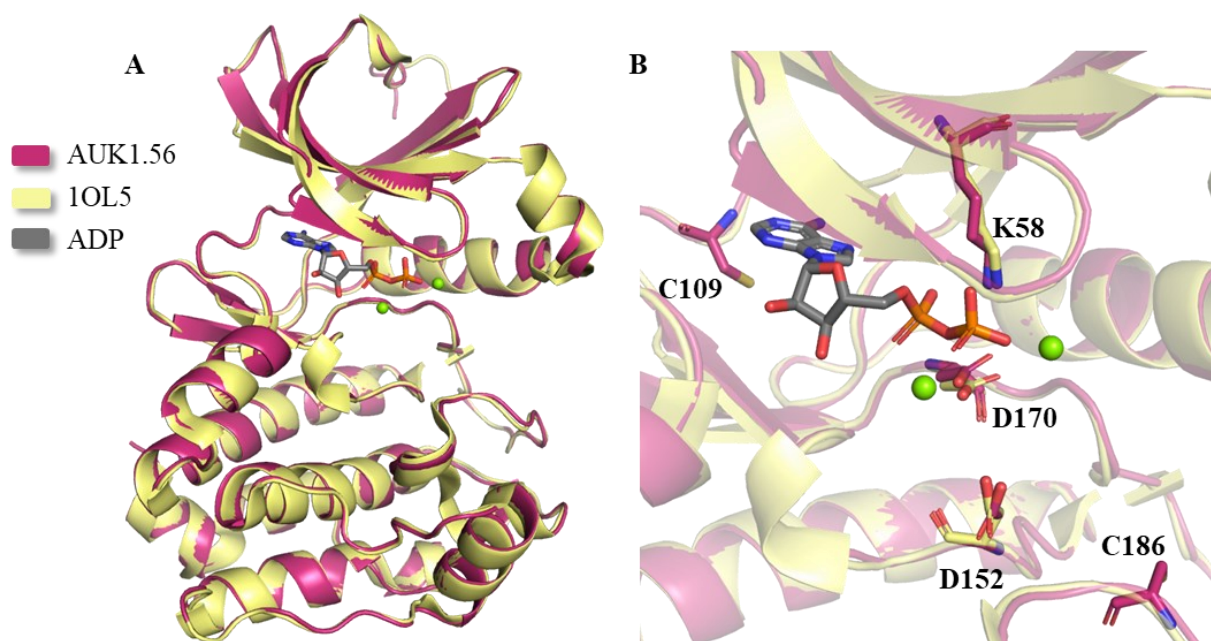


Figura 21 - Estrutura tridimensional do modelo AUK1.56 refinado da quinase AUK1 de *T. cruzi* da cepa CL Brener NE1 gerado por modelagem comparativa em sobreposição com o molde 1OL5 com destaque para os resíduos catalíticos e cisteínas mapeadas. **A.** Sobreposição do modelo AUK1.56 com o molde 1OL5, com a presença do inibidor ADP e de 2 íons magnésio (esferas em verde). **B.** Destaque para os 3 resíduos catalíticos da AUK1, K58, D152 e D170, e para as 2 cisteína mapeadas C109 (Hinge H2) e C186 (Other O3).

Em relação aos resíduos do sítio ativo do modelo e do molde que se localizam a até 6 Å de distância do ligante ADP, tem-se uma identidade de 71,9%, com 23 resíduos iguais dentre um total de 32. Semelhante ao observado na modelagem AUK1-2BFY, grande parte dos resíduos diferentes (**Figura 22**) estão em áreas expostas ao solvente e não comprometem espacialmente o sítio. Igualmente, no modelo AUK1.56 refinado a maior parcela dos resíduos diferentes é de tamanho semelhante ou menor que os resíduos correspondentes no molde, além de apolares ou polares não carregados. Dentre os 9 resíduos diferentes, no molde 1OL5 se observam 3 resíduos carregados positivamente enquanto que no modelo os resíduos em mesma posição não apresentam carga, conforme pode ser observado qualitativamente na coloração da superfície no painel B da **Figura 22**. A diferença observada entre os potenciais eletrostáticos da proteína molde (humana) e do modelo pode reduzir a probabilidade de que os inibidores de uma destas proteínas sejam efetivos contra ambas.

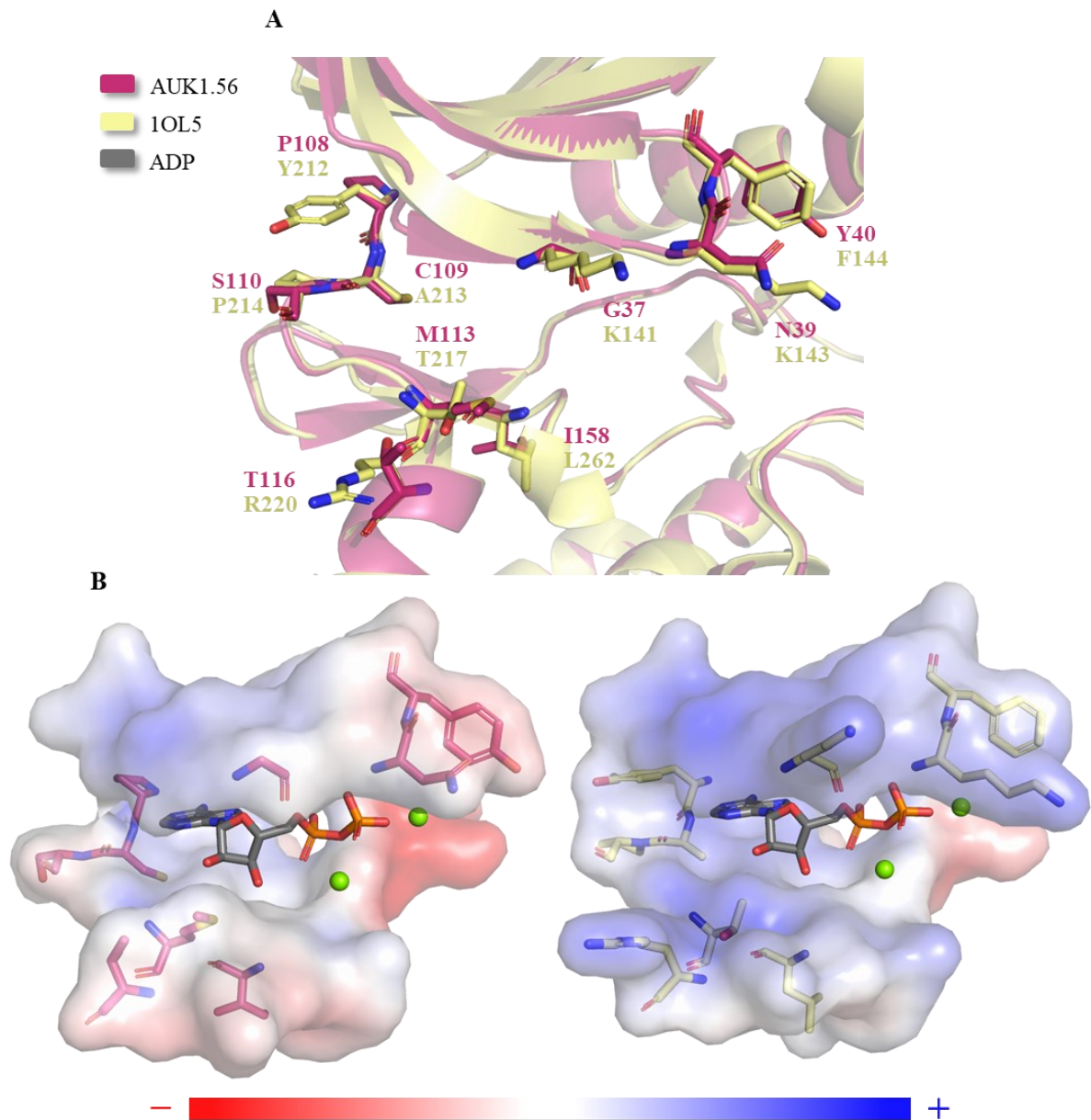


Figura 22 - Destaque das diferenças no sítio de ligação do ATP entre a estrutura do modelo AUK1.56 refinado da quinase AUK1 de *T. cruzi* da cepa CL Brener NE1 e a estrutura do molde 1OL5. **A.** Destaque em bastões dos resíduos diferentes entre o modelo AUK1.56 refinado e molde 1OL5 que se encontram a até 6 Å do ligante ADP. A estrutura da molécula ADP e parte da estrutura tridimensional foram omitidas para melhor visualização da região do sítio. **B.** Superfície eletrostática da região analisada para o modelo (à esquerda) e o molde (à direita), gerada pelo *plugin* APBS do PyMOL^{106,107}. A coloração em vermelho indica regiões de potencial eletrostático negativo e em azul regiões de potencial positivo.

Assim como o modelo AUK1.26 refinado da modelagem AUK1-2BFY, o modelo AUK1.56 refinado apresentou uma boa qualidade geral, podendo ser utilizado em etapas futuras do trabalho. Para compreender se há uma diferença significativa entre as estruturas dos dois modelos gerados para a quinase AUK1, algumas análises foram conduzidas. Em relação à estrutura do molde 2BFY, Sessa *et al.* (2005) associam a semiatividade da Aurora B a uma supressão da interação entre a lisina catalítica e o glutamato presente na α -hélice C no sítio ativo em conformação ativa. Estes resíduos estão mais distantes nesta conformação semiativa. Um outro aspecto apontado é uma rotação de 15° entre o lóbulo N e o lóbulo C em comparação com a estrutura de código PDB 1OL5 em conformação ativa¹⁰⁵. Este mesmo padrão é observado nos modelos selecionados. A distância entre a lisina catalítica e o glutamato da α -hélice C no modelo AUK1.26 refinado (4,9 Å) é maior do que a do modelo construído com o molde na forma ativa (2,9 Å). Com o auxílio do servidor *Protein Domain Movement Analysis*⁹⁸ calculou-se o valor do deslocamento entre os lóbulos N e C comparando os dois modelos da AUK1. O servidor alinhou o lóbulo C das duas estruturas e apontou uma rotação de $14,3^\circ$ entre os lóbulos N das duas estruturas. Estas observações estão ilustradas na **Figura 23**. Além disso, foi avaliado o volume do sítio de ligação do ATP dos dois modelos, utilizando a ferramenta DoGSite⁹⁹. Os volumes encontrados foram semelhantes, $542,3 \text{ \AA}^3$ para o modelo AUK1.26 refinado (molde 2BFY) e $515,8 \text{ \AA}^3$ para o modelo AUK1.56 refinado (molde 1OL5) (**Figura 23**). Apesar do deslocamento dos lóbulos entre a estrutura semiativa e a ativa, e da supressão da ponte salina na estrutura semiativa, houve uma certa preservação do volume do sítio ativo. No entanto, diferenças quanto ao formato da cavidade podem impactar no reconhecimento de inibidores, sendo interessante considerar ambas as estruturas na busca por inibidores.

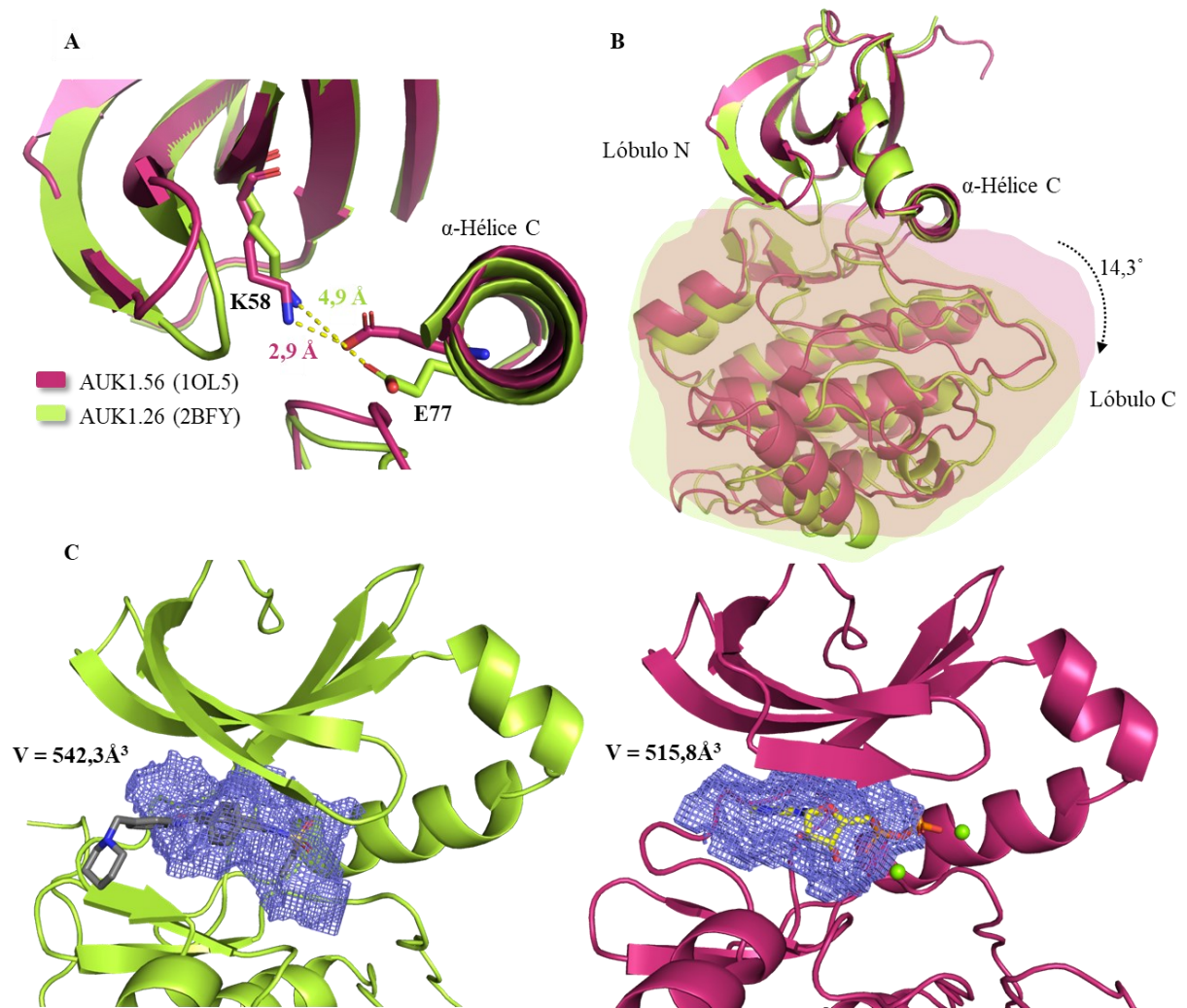


Figura 23 - Análise comparativa entre os modelos gerados para a quinase AUK1 de *T. cruzi* da cepa CL Brener NEI, AUK1.26 (molde 2BFY), estrutura semiativa, e AUK1.56 (molde 1OL5), estrutura ativa. **A.** Destaque para a distância entre a lisina catalítica K58 e o Glutamato E77 que estabelecem uma ponte salina na quinase em conformação ativa. O lóbulo N das duas estruturas foram alinhados para melhor evidenciar a variação de distância (parte da estrutura tridimensional foi omitida para despoluir a região analisada). **B.** Ilustração do deslocamento do lóbulo C entre as estruturas dos dois modelos a partir do alinhamento do lóbulo N. **C.** Volume do sítio do ATP calculado com a ferramenta DoGSite destacado em azul. À esquerda para a estrutura AUK1.26 e à direita para a estrutura AUK1.56.

Um outro aspecto interessante de analisar é a conservação dos resíduos na região do sítio ativo em outras cepas do *T. cruzi*. Quanto maior a conservação destes resíduos, maior é a probabilidade de inibidores da AUK1 de CL Brener NEI serem eficientes também nas demais cepas. Após o alinhamento no Clustal Ω das sequências consideradas na **Tabela 10**, de CL Brener EI e NEI, Dm28c 2014 e 2018, YC6, Sylvio X10/1-2012 e Brazil A4, é possível verificar que 100% dos resíduos que se encontram a até 6 Å de distância dos ligantes presentes nos dois moldes, H1N e ADP, são conservados (**Figura 24**). Diante deste resultado e do que foi

apresentado sobre a AUK1, sua modelagem comparativa e discussões, tem-se as evidências e ferramentas necessárias para explorar esta quinase como um alvo terapêutico contra este parasito.

<i>Esmeraldo</i>	1	MSAAEGGQVVASYIALPKAPRSDWKASDFEMLH	K	L	G	G	N	Y	G	D	V	L	A	S	V	R	D	C	N	F	V	C	A	I	K	K	L	60																																	
<i>Non-Esmeraldo</i>	1	MSAAEGGQVVASYIALPKAPRSDWKASDFEMLH	K	L	G	G	N	Y	G	D	V	L	A	S	V	R	D	C	N	F	V	C	A	I	K	K	L	60																																	
<i>Dm28c-2014</i>	1	MSAAEGGQVVASYIALPKAPRSDWKASDFEMLH	K	L	G	G	N	Y	G	D	V	L	A	S	V	R	D	C	N	F	V	C	A	I	K	K	L	60																																	
<i>Dm28c-2018</i>	1	MSAAEGGQVVASYIALPKAPRSDWKASDFEMLH	K	L	G	G	N	Y	G	D	V	L	A	S	V	R	D	C	N	F	V	C	A	I	K	K	L	60																																	
<i>YC6</i>	1	MSAAEGGQVVASYIALPKAPRSDWKASDFEMLH	K	L	G	G	N	Y	G	D	V	L	A	S	V	R	D	C	N	F	V	C	A	I	K	K	L	60																																	
<i>Sylvio</i>	1	MSAAEGGQVVASYIALPKAPRSDWKASDFEMLH	K	L	G	G	N	Y	G	D	V	L	A	S	V	R	D	C	N	F	V	C	A	I	K	K	L	60																																	
<i>Brazil</i>	1	MSAAEGGQVVASYIALPKAPRSDWKASDFEMLH	K	L	G	G	N	Y	G	D	V	L	A	S	V	R	D	C	N	F	V	C	A	I	K	K	L	60																																	
<i>Esmeraldo</i>	61	S	I	K	K	L	A	E	F	D	I	V	A	Q	L	R	R	E	I	E	I	A	F	H	T	R	H	R	Y	L	L	R	T	Y	G	Y	F	F	D	D	T	I	Y	L	I	L	E	P	C	S	N	G	M	L	Y	T	E	L	N	R	120
<i>Non-Esmeraldo</i>	61	S	I	K	K	L	A	E	F	D	I	V	A	Q	L	R	R	E	I	E	I	A	F	H	T	R	H	R	Y	L	L	R	T	Y	G	Y	F	F	D	D	T	I	Y	L	I	L	E	P	C	S	N	G	M	L	Y	T	E	L	N	R	120
<i>Dm28c-2014</i>	61	S	I	K	K	L	A	E	F	D	I	V	A	Q	L	R	R	E	I	E	I	A	F	H	T	R	H	R	Y	L	L	R	T	Y	G	Y	F	F	D	D	T	I	Y	L	I	L	E	P	C	S	N	G	M	L	Y	T	E	L	N	R	120
<i>Dm28c-2018</i>	61	S	I	K	K	L	A	E	F	D	I	V	A	Q	L	R	R	E	I	E	I	A	F	H	T	R	H	R	Y	L	L	R	T	Y	G	Y	F	F	D	D	T	I	Y	L	I	L	E	P	C	S	N	G	M	L	Y	T	E	L	N	R	120
<i>YC6</i>	61	S	I	K	K	L	A	E	F	D	I	V	A	Q	L	R	R	E	I	E	I	A	F	H	T	R	H	R	Y	L	L	R	T	Y	G	Y	F	F	D	D	T	I	Y	L	I	L	E	P	C	S	N	G	M	L	Y	T	E	L	N	R	120
<i>Sylvio</i>	61	S	I	K	K	L	A	E	F	D	I	V	A	Q	L	R	R	E	I	E	I	A	F	H	T	R	H	R	Y	L	L	R	T	Y	G	Y	F	F	D	D	T	I	Y	L	I	L	E	P	C	S	N	G	M	L	Y	T	E	L	N	R	120
<i>Brazil</i>	61	S	I	K	K	L	A	E	F	D	I	V	A	Q	L	R	R	E	I	E	I	A	F	H	T	R	H	R	Y	L	L	R	T	Y	G	Y	F	F	D	D	T	I	Y	L	I	L	E	P	C	S	N	G	M	L	Y	T	E	L	N	R	120
<i>Esmeraldo</i>	121	V	K	F	F	P	P	T	A	A	R	Y	V	A	Q	L	A	E	A	L	L	Y	L	H	Q	H	H	I	L	H	R	D	I	K	P	E	N	I	L	L	D	H	N	Q	N	I	K	L	A	D	F	G	W	S	V	H	D	P	H	N	180
<i>Non-Esmeraldo</i>	121	V	K	F	F	P	P	T	A	A	R	Y	V	A	Q	L	A	E	A	L	L	Y	L	H	Q	H	H	I	L	H	R	D	I	K	P	E	N	I	L	L	D	H	N	Q	N	I	K	L	A	D	F	G	W	S	V	H	D	P	H	N	180
<i>Dm28c-2014</i>	121	V	K	F	F	P	P	T	A	A	R	Y	V	A	Q	L	A	E	A	L	L	Y	L	H	Q	H	H	I	L	H	R	D	I	K	P	E	N	I	L	L	D	H	N	Q	N	I	K	L	A	D	F	G	W	S	V	H	D	P	Y	N	180
<i>Dm28c-2018</i>	121	V	K	F	F	P	P	T	A	A	R	Y	V	A	Q	L	A	E	A	L	L	Y	L	H	Q	H	H	I	L	H	R	D	I	K	P	E	N	I	L	L	D	H	N	Q	N	I	K	L	A	D	F	G	W	S	V	H	D	P	Y	N	180
<i>YC6</i>	121	V	K	F	F	P	P	T	A	A	R	Y	V	A	Q	L	A	E	A	L	L	Y	L	H	Q	H	H	I	L	H	R	D	I	K	P	E	N	I	L	L	D	H	N	Q	N	I	K	L	A	D	F	G	W	S	V	H	D	P	H	N	180
<i>Sylvio</i>	121	V	K	F	F	P	P	T	A	A	R	Y	V	A	Q	L	A	E	A	L	L	Y	L	H	Q	H	H	I	L	H	R	D	I	K	P	E	N	I	L	L	D	H	N	Q	N	I	K	L	A	D	F	G	W	S	V	H	D	P	Y	N	180
<i>Brazil</i>	121	V	K	F	F	P	P	T	A	A	R	Y	V	A	Q	L	A	E	A	L	L	Y	L	H	Q	H	H	I	L	H	R	D	I	K	P	E	N	I	L	L	D	H	N	Q	N	I	K	L	A	D	F	G	W	S	V	H	D	P	Y	N	180

Figura 24 - Alinhamento das sequências das quinases AUK1 correspondentes em outras cepas de *T. cruzi* com destaque para os resíduos que se encontram no sítio do ATP a até 6 Å de distância dos inibidores H1N (molde 2BFY) e ADP (Molde 1OL5). As sequências foram alinhadas considerando o intervalo do aminoácido 1 ao 180 da quinase AUK1 de CL Brener NEI.

4.4.2 CLK1

A modelagem e respectivas análises referentes à segunda quinase priorizada, a CLK1 de *T. cruzi* CL Brener haplótipo El (código TriTrypDB: TcCLB.511127.320; código Uniprot: Q4E3Z1), são apresentadas a seguir. Assim como a quinase Aurora, a CLK1 é uma ePK, porém pertencente ao grupo CMGC e família CLK, conforme anotação em bancos de dados e classificação no kinase.com. Esta quinase também está identificada com o nome KKT10 no TriTrypDB. A CLK1 de *T. brucei* na forma sanguínea é uma das 24 proteínas que compõem o cinetócoro, um complexo proteico envolvido na mitose que conecta os cromossomos ao fuso de microtúbulos, facilitando o alinhamento e posterior separação dos cromossomos duplicados

de forma controlada e coordenada¹⁰⁹. Um dos substratos da CLK1 é a quinase KKT2, uma outra proteína constituinte do cinetócoro cuja fosforilação pela CLK1 é crucial para a montagem desse complexo proteico¹⁰⁹. A inibição ou o silenciamento da quinase CLK1 prejudica o recrutamento do cinetócoro interno, comprometendo a progressão do ciclo celular, que fica retido irreversivelmente na fase M, causando a morte celular^{109,110}. Saldivia *at al.* (2021) ainda ponderam que em virtude da conservação da CLK1 entre os organismos *T. brucei*, *T. cruzi* e *Leishmania mexicana*, assim como a inibição do crescimento destes outros dois organismo por um composto que inibiu a CLK1 de *T. brucei* (dados de Saldivia *at al.* (2020)¹¹⁰), e a conservação do sítio de fosforilação na quinase KKT2 destes dois outros organismos, é provável que essa via de sinalização seja conservada nos tripanosomatídeos.¹⁰⁹ Até o presente momento ainda não existem estudos focados na quinase CLK1 de *T. cruzi*, restando, em um primeiro momento, extrapolar o conhecimento sobre a quinase de *T. brucei*.

A sequência da CLK1 de *T. cruzi* CL Brener El no TriTrypDB é constituída por 482 resíduos de aminoácidos. Avaliando as quinases mais próximas da CLK1 em outras cepas deste organismo, por meio do BLAST no TriTrypDB, constata-se que são bem semelhantes em identidade e tamanho, apresentando uma identidade mínima de 86% com a enzima de CL Brener El (**Tabela 14**). No mapeamento das cisteínas relevantes para inibição covalente nestas cepas, foram encontradas cisteínas nas mesmas posições mapeadas na quinase de CL Brener El (**Tabela 14**). Como para a CLK1 a sequência da cepa CL Brener El não apresentou alterações ou divergências em comparação com as demais cepas analisadas, a modelagem foi conduzida utilizando a sequência da quinase CL Brener El, conforme planejado.

Tabela 14 - Comparativo do número de resíduos, identidade e número de cisteínas mapeadas entre as sequências de quinases CLK1 de diversas cepas de *T. cruzi* com a cepa CL Brener Esmeraldo-like (CL Brener El)

Cepa	Código TryTripDB	Nº Resíduos	Identidade com CLK1 de CL Brener El	Cisteínas mapeadas
CL Brener El	TcCLB.511127.320	482	———	C172 (Roof A3) C232 (Front F1) C305 (DFG D1) C310 (Others O2)

Cepa	Código TryTripDB	Nº Resíduos	Identidade com CLK1 de CL Brener EI	Cisteínas mapeadas
CL Brener NEI	TcCLB.509027.60	488	98%	C178 (Roof A3) C238 (Front F1) C311 (DFG D1) C316 (Others O2)
Dm28c 2014	TCDM_03791	493	86%	C183 (Roof A3) C243 (Front F1) C316 (DFG D1) C321 (Others O2)
Dm28c 2018	C4B63_27g204	485	98%	C175 (Roof A3) C235 (Front F1) C308 (DFG D1) C313 (Others O2)
Y C6	TcYC6_0036270	486	99%	C176 (Roof A3) C236 (Front F1) C309 (DFG D1) C314 (Others O2)
Sylvio X10/1-2012	TCSYLVIO_008709	334	99%	C22 (Roof A3) C82 (Front F1) C155 (DFG D1) C160 (Others O2)
Brazil A4	TcBrA4_0109090	490	87%	C180 (Roof A3) C240 (Front F1) C313 (DFG D1) C318 (Others O2)

A busca por molde para conduzir a modelagem comparativa foi feita tanto por BLASTp no NCBI quanto no servidor do Swiss-Model (**Tabela 15**). Visto que todos os moldes considerados na tabela apresentavam inibidores, novamente esta informação foi omitida. Como no caso da modelagem da AUK1, aqui também se observam muitos potenciais moldes com resíduos faltantes. Contudo, para esta quinase foi detectada uma estrutura cristalográfica de um organismo próximo, a estrutura de código PDB 6Q2A, de *T. brucei*, com uma alta identidade em relação à CLK1 do *T. cruzi*, 82,5%, com uma cobertura de 72% e ausência de resíduos faltantes dentro do domínio de quinase. Dada a qualidade estrutural e a proximidade com o organismo foco deste estudo, esta estrutura foi selecionada como molde. Um outro ponto interessante sobre a estrutura 6Q2A é a presença de um inibidor ligado covalentemente a uma

cisteína no sítio do ATP, o que é bastante relevante visto que é a abordagem de inibição de interesse deste trabalho.

Tabela 15 - Dados dos candidatos a moldes para a modelagem comparativa da quinase CLK1 da cepa CL Brener El de *T. cruzi*

BLAST NCBI					
Posição	PDB Resolução /Å	Identidade e Cobertura /%	Nº Resíduos Faltantes e localização ^c		Presença de Fosforilação
1	6Q2A - 2,60 ^a	82,5 - 72	-	-	Não
2	2EXE - 2,35 ^b	35,7 - 70	70	Fitas β de 1 a 3	Não
			5	Logo após αC	
			10	Fitas β4 e 5	
			5	Alça entre fitas β7 e 8	
3	2WU6 - 1,92 ^b	35,7 - 70	-	-	Não
4	6YTY - 1,76 ^b	35,7 - 70	2	Alça entre fitas β7 e 8	Não
5	6FT7 - 2,02 ^b	35,7 - 70	-	-	Não

Swiss-Model					
Posição	PDB Resolução /Å	Identidade e Cobertura /%	Nº Resíduos Faltantes e localização ^c		Presença de Fosforilação
1	6Q2A - 2,60 ^a	82,5 - 72	-	-	Não
2	6Z50 - 1,60 ^b	33,1 - 66	-	-	Não
3	6R3D - 1,85 ^b	33,1 - 66	-	-	Não
4	6I5I - 1,60 ^b	33,1 - 66	3	Após αI	Não
5	5X8I - 1,90 ^b	33,1 - 66	-	-	Não

^a Organismo *Trypanosoma brucei*

^b Organismo *Homo sapiens*

^c Os dados indicados correspondem ao número de resíduos ausentes na estrutura cristalográfica dentro do intervalo do domínio característico das quinases.

A quinase molde de *T. brucei* também está classificada como uma CLK1.¹¹⁰ Assim como nos demais moldes analisados, a estrutura cristalográfica 6Q2A não está associada a outras proteínas ou peptídeos e não há fosforilação na alça de ativação e nem em outras regiões da quinase. Contudo, Saldivia *et al.* (2020) pontuam que a estrutura está coerente com quinases em conformação ativa, apresentando uma alça de ativação bem ordenada, Asp-Leu-Gly-in

(DLG-in) e uma ponte salina intacta com a α -hélice C¹¹⁰. A CLK1 de *T. brucei* é uma proteína de 465 resíduos. Durante o estudo cristalográfico focou-se no trecho do aminoácido 117 ao 465, chegando, no fim, a uma estrutura com o intervalo de resíduos do 127 ao 465. O alinhamento no Clustal Ω da sequência da estrutura da 6Q2A com a sequência da quinase CLK1 de *T. cruzi* está representado na **Figura 25**. Apesar da ausência de cobertura nos primeiros 143 aminoácidos da sequência da quinase de *T. cruzi*, todo o domínio característico das quinases ficou dentro do intervalo coberto, do resíduo 144 ao 482.

TcCLB.511127.320(CLK1) 6Q2A	MDTLAGNLGAARTGIQYTVGQANVASGNGGKALNSARSTMTTSQSTNSARRSGSKRDRE -----	60 0
TcCLB.511127.320(CLK1) 6Q2A	TATSTDITRTRKSQYGEVTKKQATATTTTNTTTTTTATAGGANRIVEVLPKPKKKVTYA -----	120 0
TcCLB.511127.320(CLK1) 6Q2A	LPNQSREEGHFYVVLGEDIDVSTQRFKILSMLGEGTFGKVVEAWDRKRKEYCAVKIVRNV -----GRFKILSLLGEGTFGKVVEAWDRKRKEYCAVKIVRNV *****:*****	180 37
TcCLB.511127.320(CLK1) 6Q2A	PKYTRDAKIEIQFMEKVRQADDRFSLMKIQRYFQNETGHMCIVMPKYGPCLLDWIMKH PKYTRDAKIEIQFMEVRVLSDEVDRFPLMKIQRYFQNETGHMCIVMPKYGPCLLDWIMKH *****:.* :.*:*** *****	240 97
TcCLB.511127.320(CLK1) 6Q2A	GPLSHRHQAQIIFQVGTALDYFHTELRLMHTDLKPENMLLETSETVIDPVTRKATPPDPC GPFNHRHLAQIIFQVGAALDYFHTELRLMHTDLKPENILMESGDTSDPMPHRLPPEPC *:*:*****:*****:*****:.*:.*:.* :*:*:.*:.* :*:*:.*:.*	300 157
TcCLB.511127.320(CLK1) 6Q2A	RIRICDLGGCCDERHSRTAIVSTRHYRSPEVVLGLGWMYSTDMWSMGCIIYELYTGKLLY RVRICDLGGCCDERHSRTAIVSTRHYRSPEVLSLGLWYSTDLMWSMGCIIYELYTGKLLY *:*****:*****:*****:*****:*****:*****	360 217
TcCLB.511127.320(CLK1) 6Q2A	DTHDNLEHLHLEKTLGRLPQEWAGRCGTEEARQLYNSVAQLRPCTDPKHLARIARARPV DTHDNLEHLHLEKTLGRLPADWSVRCGTQEARDLFTAAGTLQPCDKPKHARIARARPV *****:.* :.*:***:***:.*:.*:.* :*:*:.*:.*:*****:*****	420 277
TcCLB.511127.320(CLK1) 6Q2A	REVISDKLLCDLIYLLHYDRQKRLTARQMTTHQYVLKYYPECRQHPNYPDNRPTLRPTP REVITEPLLCDLILNLLHYDRQRRLNARQMMSHAYVHKYFPECRQHPNVDNRSKLPPTP ***:.* :***** :*****:***:*** :* ** **:*:*****:*** :.* **	480 337
TcCLB.511127.320(CLK1) 6Q2A	LM 482 VM 339 :*	

Figura 25 - Alinhamento produzido no servidor Clustal Ω entre a sequência da quinase CLK1 de *T. cruzi* da cepa CL Brener El e a sequência do molde (código PDB 6Q2A) para execução da modelagem comparativa.

Após conduzir a modelagem e o processo de refinamento e de validação dos modelos (**Tabela 16**), o modelo CLK1.35 refinado foi o selecionado para representar a quinase CLK1 de *T. cruzi*. No diagrama de Ramachandran sua estrutura apresentou 97,3% dos resíduos em regiões favoráveis e 100% em regiões permitidas (**Figura 26**). Os valores do Errat e Verify 3D, 90,0 e

94,7, respectivamente, ficaram melhores na estrutura refinada em comparação com a não refinada e houve melhora na quantidade de *clashes* presentes. O valor de Z-score retornado pelo ProSA está semelhante aos valores das estruturas experimentais, reforçando que se obteve um modelo com boa qualidade (**Figura 26**).

Tabela 16 - Dados de validação dos 3 modelos pré-selecionados da modelagem da quinase CLK1 de *T. cruzi* da cepa CL Brener El gerados no programa Modeller

Modelos	Modeller		PyMOL/ Å	ProSA	Ramachandran MolProbit /%			Saves / %	
	Dope Score	Molpdf	RMSD	Z-score	Favorável	Permitido	Não permitido	Errat	Verivy 3D
CLK1.24 Refinado ^a	-42422,2	2665,2	0,097 ^b	-9,11	97,3	99,4	0,6	84,9	96,2
			0,189 ^c	-9,22	97,0	99,7	0,3	93,4	95,3
CLK1.78 Refinado ^a	-42278,6	2673,5	0,098 ^b	-9,16	97,6	99,7	0,3	76,7	95,9
			0,297 ^c	-9,14	95,8	99,4	0,6	99,4	93,2
CLK1.35 Refinado ^a	-42316,0	2681,8	0,099 ^b	-9,13	97,9	99,4	0,6	86,4	91,7
			0,337 ^c	-9,12	97,3	100,0	0,0	90,1	94,7
Molde 6Q2A				-9,32	97,9	100,0	0,0	94,2	90,6

^aRefinamento realizado do servidor LocPREFMD (<http://feig.bch.msu.edu/web/services/locprefmd/>)

^bRMSD calculado em relação ao molde 6Q2A

^cRMSD calculado em relação ao modelo inicial não refinado

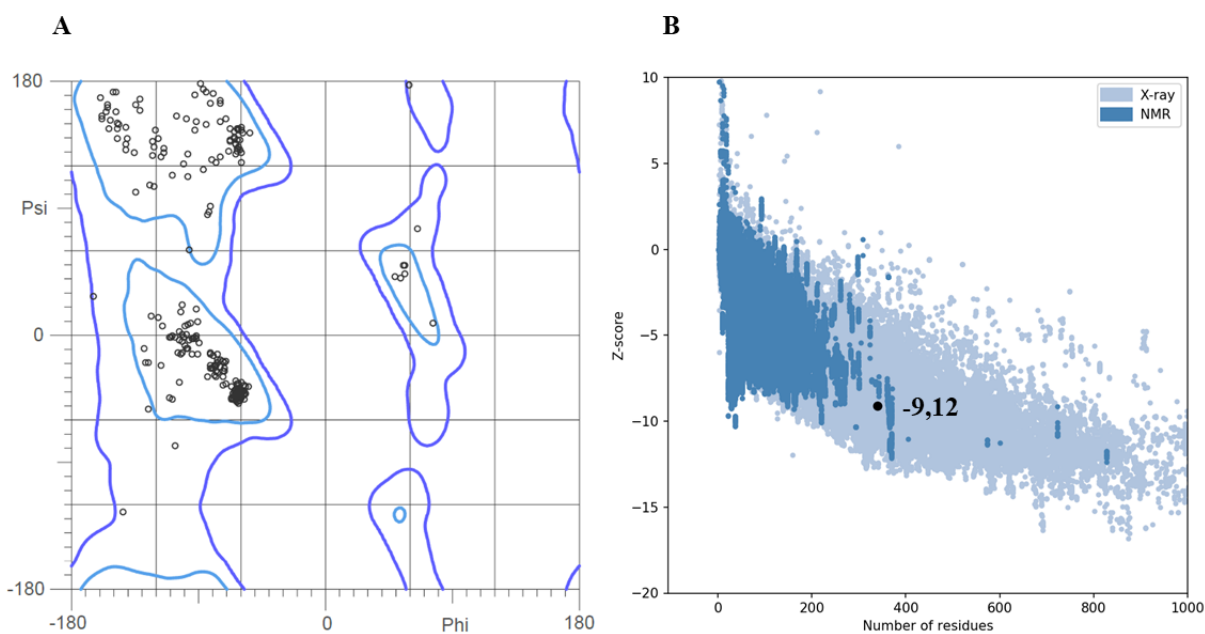


Figura 26 - Validação do modelo CLK1.35 refinado construído por modelagem comparativa para a proteína CLK1 de *T. cruzi* da cepa CL Brener El, utilizando o molde 6Q2A. **A.** Diagrama de Ramachandran gerado no servidor MolProbit para o modelo CLK1.35 refinado. Contornos em azul

escuro indicam regiões permitidas, enquanto contornos em azul claro indicam as regiões favoráveis. **B.** Gráfico gerado no servidor ProSA, destacando o valor de Z-score para o modelo CLK1.35 refinado em relação a estruturas determinadas experimentalmente, com número de resíduos semelhantes.

A sobreposição das estruturas do modelo da CLK1 e do molde (**Figura 27**) apresentaram um RMSD de 0,361 Å, reforçando a qualidade do modelo obtido. Observou-se que os resíduos catalíticos (CLK1: K175; D172 e D306) e as cisteínas mapeadas (CLK1: C232 (Front F1); C305 (DFG D1) e C310 (Others O2)) ficaram bem sobrepostos com os resíduos correspondentes do molde. Ao observar mais pontualmente o sítio do ATP, apesar de uma boa correlação entre as estruturas nesta região, existem dois resíduos no modelo classificados como rotâmeros *outliers* pelo servidor MolProbity, a lisina 228 e o glutamato 276. Embora estes resíduos estejam no sítio, suas cadeias laterais estão apontando para fora da cavidade, não gerando nenhum comprometimento estérico, e apresentam conformações semelhantes aos resíduos correspondentes no molde, que não estão classificados como rotâmeros *outliers* pelo MolProbity (**Figura 28**).

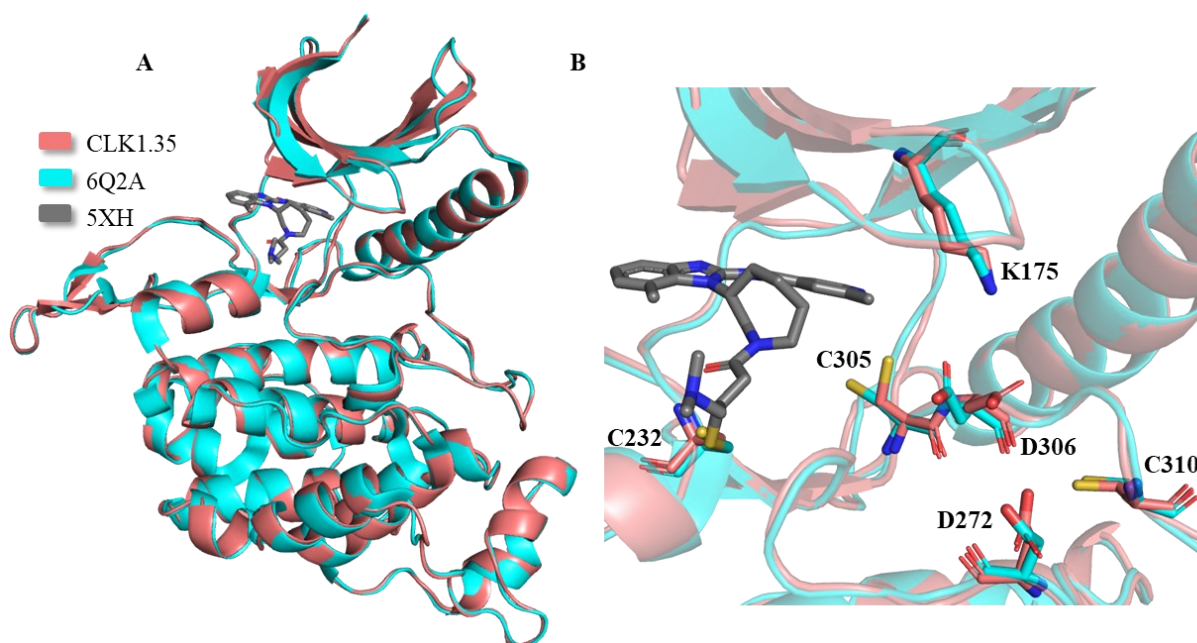


Figura 27 - Estrutura tridimensional do modelo CLK1.35 refinado da quinase CLK1 de *T. cruzi* da cepa CL Brener El gerado por modelagem comparativa em sobreposição com o molde 6Q2A com destaque para os resíduos catalíticos e cisteínas mapeadas. **A.** Sobreposição do modelo CLK1.35 com o molde 6Q2A, com a presença do inibidor 5XH. **B.** Destaque para os 3 resíduos catalíticos da CLK1, K175; D172 e D306, e para as 3 cisteína mapeadas C232 (Front F1); C305 (DFG D1) e C310 (Others O2).

Ainda em relação às características do sítio do ATP, entre as duas estruturas há uma alta conservação dos resíduos, com 25 de 26 resíduos conservados, conferindo uma identidade de 96,15%, considerando uma distância de até 6 Å do inibidor 5XH presente na estrutura cristalográfica do molde 6Q2A. A única diferença observada foi a presença de uma glicina no molde na mesma posição onde há uma serina na quinase de *T. cruzi*, um resíduo maior e polar, numa região exposta ao solvente e já mais externa no sítio (**Figura 28**).

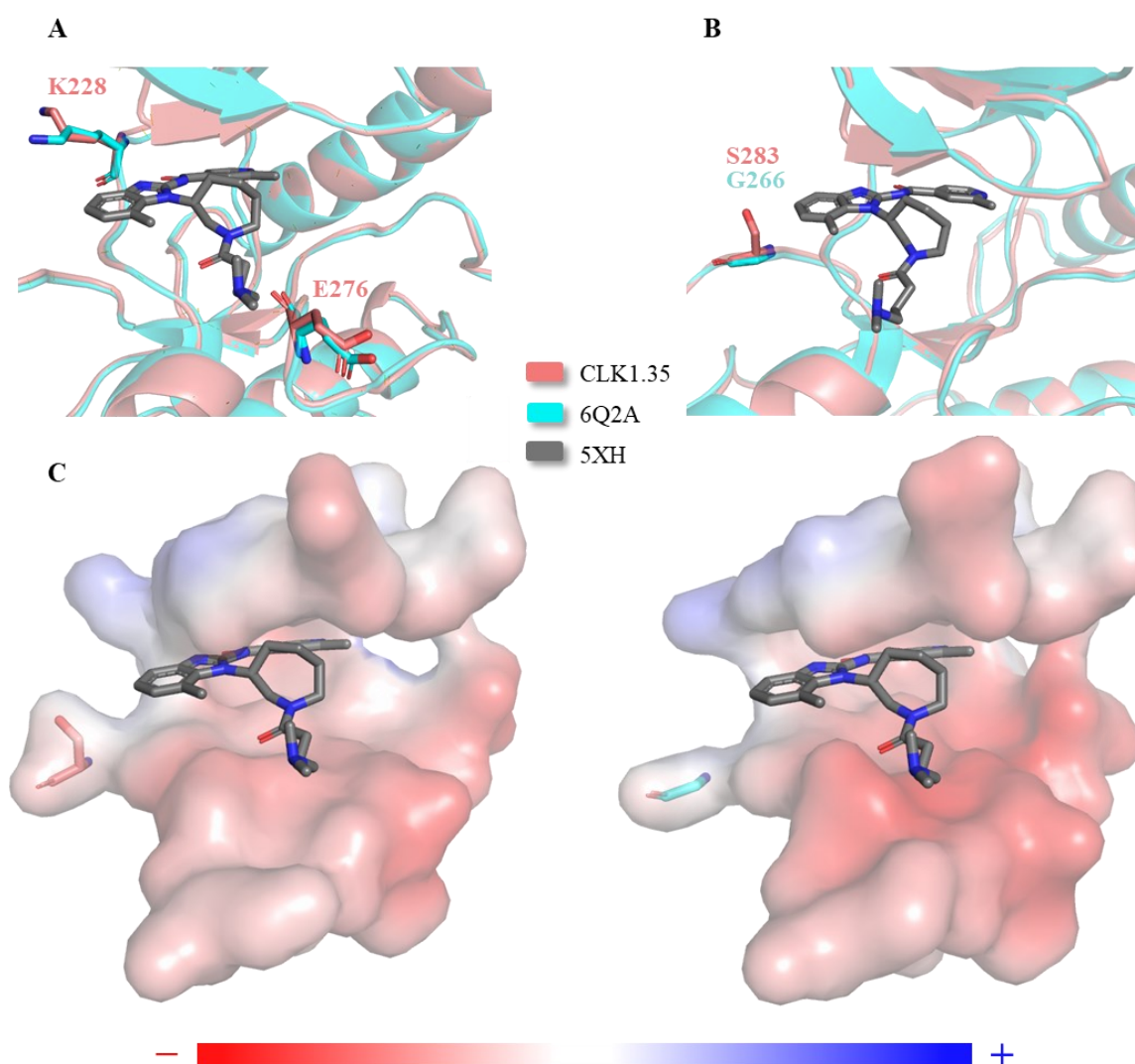


Figura 28 - Destaque das diferenças no sítio de ligação do ATP entre a estrutura do modelo CLK1.35 refinado da quinase CLK1 de *T. cruzi* da cepa CL Brener El e a estrutura do molde 6Q2A. **A.** Rotâmeros *outliers* (K228 e E276) na estrutura do modelo CLK1.35 refinado na região do sítio de ligação do ATP em sobreposição com o molde. **B.** Destaque em bastões dos resíduos diferentes entre o modelo CLK1.35 refinado e o molde 6Q2A que se encontram a até 6 Å do inibidor 5XH. **C.** Superfície eletrostática da região analisada para o modelo (à esquerda) e o molde (à direita), gerada pelo *plugin* APBS do PyMOL^{106,107}. A coloração em vermelho indica regiões de potencial eletrostático negativo e em azul regiões de potencial positivo.

Tendo em vista a significativa semelhança entre as CLK1 dos dois organismos, é provável que a quinase de *T. cruzi* também seja covalentemente inibida pelo mesmo inibidor do molde, um amidobenzimidazol apelidado de AB1 (5XH no PDB 6Q2A), originalmente estudado como inibidor covalente da quinase mutante do receptor de fator de crescimento epidermal humano - *human Epidermal Growth Factor Receptor* (hEGFR) no tratamento de câncer pulmonar de células não pequenas¹¹⁰. De fato, Saldivia *et al.* (2020) apresentam evidências experimentais da inibição do crescimento do parasito *T. cruzi* na forma amastigota intracelular em ensaios *in vitro* com este inibidor, reportando um EC₅₀ de 44nM¹¹⁰. Um aspecto que sugere que provavelmente a CLK1 seria a quinase envolvida na inibição do crescimento do *T. cruzi* é a conservação da cisteína que estabelece a ligação covalente com o inibidor no sítio do ATP, cisteína 232 (subsítio F1 do sítio *Front region*) na CLK1 de *T. cruzi*. Em relação às demais cepas de *T. cruzi* (Tabela 14), ao analisar os resíduos do sítio ativo, percebe-se uma conservação de 100% dos resíduos situados a até 6 Å do inibidor do molde (Figura 29). Isto implica que inibidores planejados para a cepa CL Brener El provavelmente poderão ser extrapolados para as demais cepas. Portanto, assim como a quinase AUK1, a CLK1 se mostra como um alvo terapêutico promissor.

Esmeraldo	123	NQSREEGHFYVVLGEDIDVSTQRFKILSML	EGTFGK	VVEAWDRKRKEYCAVK	IVRNVPKY	183
Non-Esmeraldo	129	NQSREEGHFYVVLGEDIDVSTQRFKILSML	EGTFGK	VVEAWDRKRKEYCAVK	IVRNVPKY	189
Dm28c-2014	134	NQSREEGHFYVVLGEDIDVSTQRFKILSML	EGTFGK	VVEAWDRKRKEYCAVK	IVRNVPKY	194
Dm28c-2018	126	NQSREEGHFYVVLGEDIDVSTQRFKILSML	EGTFGK	VVEAWDRKRKEYCAVK	IVRNVPKY	186
YC6	127	NQSREEGHFYVVLGEDIDVSTQRFKILSML	EGTFGK	VVEAWDRKRKEYCAVK	IVRNVPKY	187
Sylvio	1	-----ML	EGTFGK	VVEAWDRKRKEYCAVK	IVRNVPKY	33
Brazil	131	NQSREEGHFYVVLGEDIDVSTQRFKILSML	EGTFGK	VVEAWDRKRKEYCAVK	IVRNVPKY	191
Esmeraldo	184	TRDAKIEIQFMEKVRQADADDRFSLM	IKQRYFQNETGHMC	VMPKYGPCLLDWIMKHGALS	244	
Non-Esmeraldo	190	TRDAKIEIQFMEKVRQADADDRFSLM	IKQRYFQNETGHMC	VMPKYGPCLLDWIMKHGALS	250	
Dm28c-2014	195	TRDAKIEIQFMEKVRQADADDRFSLM	IKQRYFQNETGHMC	VMPKYGPCLLDWIMKHGALS	255	
Dm28c-2018	187	TRDAKIEIQFMEKVRQADADDRFSLM	IKQRYFQNETGHMC	VMPKYGPCLLDWIMKHGALS	247	
YC6	188	TRDAKIEIQFMEKVRQADADDRFSLM	IKQRYFQNETGHMC	VMPKYGPCLLDWIMKHGALS	248	
Sylvio	34	TRDAKIEIQFMEKVRQADADDRFSLM	IKQRYFQNETGHMC	VMPKYGPCLLDWIMKHGALS	94	
Brazil	192	TRDAKIEIQFMEKVRQADADDRFSLM	IKQRYFQNETGHMC	VMPKYGPCLLDWIMKHGALS	252	
Esmeraldo	245	HRHLAQIIFQAGTALDYFHTELRLMHTDLK	PENMLLET	SETVIDPVTRKATPPDPPCRIRIC	305	
Non-Esmeraldo	251	HRHLAQIIFQAGTALDYFHTELRLMHTDLK	PENMLLET	SETVIDPVTRKATPPDPPCRIRIC	311	
Dm28c-2014	256	HRHLAQIIFQAGTALDYFHTELRLMHTDLK	PENMLLET	SETVIDPVTRKATPPDPPCRIRIC	316	
Dm28c-2018	248	HRHLAQIIFQAGTALDYFHTELRLMHTDLK	PENMLLET	SETVIDPVTRKATPPDPPCRIRIC	308	
YC6	249	HRHLAQIIFQAGTALDYFHTELRLMHTDLK	PENMLLET	SETVIDPVTRKATPPDPPCRIRIC	309	
Sylvio	95	HRHLAQIIFQAGTALDYFHTELRLMHTDLK	PENMLLET	SETVIDPVTRKATPPDPPCRIRIC	155	
Brazil	253	HRHLAQIIFQAGTALDYFHTELRLMHTDLK	PENMLLET	SETVIDPVTRKATPPDPPCRIRIC	313	
Esmeraldo	306	DLGGCCDERHSRTAIVSTRHYRSPEVVLGLGWMYS			482	
Non-Esmeraldo	312	DLGGCCDERHSRTAIVSTRHYRSPEVVLGLGWMYS			488	
Dm28c-2014	317	DLGGCCDERHSRTAIVSTRHYRSPEVVLGLGWMYS			493	
Dm28c-2018	309	DLGGCCDERHSRTAIVSTRHYRSPEVVLGLGWMYS			485	
YC6	310	DLGGCCDERHSRTAIVSTRHYRSPEVVLGLGWMYS			486	
Sylvio	156	DLGGCCDERHSRTAIVSTRHYRSPEVVLGLGWMYS			334	
Brazil	314	DLGGCCDERHSRTAIVSTRHYRSPEVVLGLGWMYS			490	

Figura 29 - Alinhamento das sequências das quinases CLK1 correspondentes em outras cepas de *T. cruzi* com destaque para os resíduos que se encontram no sítio do ATP a até 6 Å de distância do inibidor 5XH (molde 6Q2A). As sequências foram alinhadas considerando o intervalo do aminoácido 123 ao 340 da quinase CLK1 de CL Brener El.

4.4.3 CRK9

A terceira quinase de *T. cruzi* priorizada, a CRK9, código TriTrypDB TcCLB.509099.150 e código Uniprot Q4DNX1, é uma ePK do grupo CMGC e da família CDK. No organismo *T. brucei*, na forma procíclica (inseto), a CRK9 se mostrou essencial para o crescimento celular. Seu silenciamento levou a um acúmulo de células nas fases G2/M com dois pares de cinetoplasto/corpo basal incapazes de se separarem o suficiente para iniciar a citocinese, além de levar a alterações morfológicas nas células que adquiriram um formato arredondado¹¹¹. A CRK9 na forma procíclica estaria, então, envolvida no controle tanto da mitose quanto da citocinese. Gourguechon *et al.* (2009) também apontam que na forma sanguínea do parasito não foram observadas essas alterações, indicando um possível papel diferenciado da CRK9 nestas duas formas de *T. brucei*.¹¹¹ Outro trabalho, de Badjatia *et al.* (2013), focado neste mesmo organismo, discorreu sobre o envolvimento e importância da CRK9 na expressão gênica. Foi visto que o silenciamento da CRK9 direta ou indiretamente interfere na fosforilação do domínio carboxi-terminal (CTD) da maior subunidade da DNA polimerase II, a RPB1. Curiosamente, este comprometimento da fosforilação da DNA polimerase II não impediu a ocorrência da transcrição, mas comprometeu etapas de maturação dos RNAs precursores produzidos e, por conseguinte, a sobrevivência do parasito¹¹². Estas constatações foram feitas para as formas procíclica e sanguínea do *T. brucei*. As quinases da família das CDKs requerem a associação com as ciclinas para sua atividade enzimática. Em *T. brucei* foi visto que a quinase CRK9 atua em conjunto com outras duas proteínas, a ciclina tipo-L CYC12, e uma proteína específica de cinetoplastídeos denominada proteína associada-CRK9 - CRK9-associated protein, ou, simplificada CRK9AP¹¹³. É formado um complexo com estas 3 proteínas, de maneira que a atividade da CRK9 é dependente da presença e associação destas outras duas proteínas. O silenciamento tanto da CYC12 quanto da CRK9AP levaram aos mesmos fenótipos do silenciamento da CRK9, demonstrando que são parceiras funcionais essenciais desta quinase¹¹³. Em relação à CRK9 de *T. cruzi*, ainda não existem trabalhos publicados sobre esta quinase.

A proteína CRK9 de *T. cruzi* CL Brener El possui uma sequência de 806 resíduos de aminoácidos, conforme dados do TriTrypDB. As quinases correspondentes em outras cepas do *T. cruzi* apresentam sequências com número de resíduos semelhantes e alta identidade quando comparadas à CRK9 de CL Brener El (**Tabela 17**), além de possuírem cisteínas nas mesmas

posições relevantes para inibição covalente. Conforme planejado, então, a modelagem foi conduzida utilizando a sequência da cepa CL Brener EI. O levantamento dos possíveis moldes para a CRK9 está apresentado na **Tabela 18**.

Tabela 17 - Comparativo do número de resíduos, identidade e número de cisteínas mapeadas entre as sequências de quinases CRK9 de diversas cepas de *T. cruzi* com a sequência CLK1 da cepa CL Brener Esmeraldo-like (CL Brener EI)

Cepa	Quinases Código TryTripDB	Nº Resíduos	Identidade com CRK9 de CL Brener EI	Cisteínas mapeadas
CL Brener EI	TcCLB.509099.150	806	—	C425 (Hinge H2) C502 (DFG D1)
CL Brener NEI	TcCLB.507209.40	804	98%	C424 (Hinge H2) C501 (DFG D1)
Dm28c 2014	TCDM_05476	805	97%	C424 (Hinge H2) C501 (DFG D1)
Dm28c 2018	C4B63_9g251	805	97%	C424 (Hinge H2) C501 (DFG D1)
Y C6	TcYC6_0076360	806	99%	C425 (Hinge H2) C502 (DFG D1)
Sylvio X10/1-2012	TCSYLVIO_000434	805	97%	C424 (Hinge H2) C501 (DFG D1)
Brazil A4	TcBrA4_0100410	805	97%	C424 (Hinge H2) C501 (DFG D1)

Tabela 18 - Dados dos candidatos a molde(s) para a modelagem comparativa da quinase CRK9 da cepa CL Brener EI de *T. cruzi*

BLAST NCBI						
Posição	PDB Resolução /Å	Identidade e Cobertura /%	Nº Resíduos Faltantes e localização ^f		Presença de	
					Fosforilação	Ciclina
1	6Z3U - 2,60 ^{a,e}	37,0 - 25	-	-	Não	Sim
2	7JV7 - 1,85 ^{b,e}	36,2 - 31	4	Entre fitas β 1 e β 2	Não	Sim
			3	Entre fita β 3 e α C		
3	6OQL - 2,71 ^{c,d}	37,7 - 26	7	Entre fita β 3 e α C	Não	Não
			8	Entre fitas β 4 e β 5		
			13	Alça de ativação		
4	1JOW - 3,10 ^{c,e}	37,6 - 26	7	Entre fitas β 1 e β 2	Não	Sim
			7	Alça entre α G e H		
5	1BI7 - 3,40 ^{c,e}	35,8 - 31	23	Entre fitas β 3 e β 4	Não	Não

BLAST NCBI						
Posição	PDB Resolução /Å	Identidade e Cobertura /%	Nº Resíduos Faltantes e localização ^f		Presença de	
					Fosforilação	Ciclina
6	3NUP - 2,60 ^{c,d}	37,7 - 26	4	Entre fitas $\beta 1$ e $\beta 2$	Não	Não
			10	Alça entre fita $\beta 3$ e αC		
			7	Entre fitas $\beta 4$ e $\beta 5$		
			14	Alça de ativação		
			2	Alça entre αG e H		
7	1VYW - 2,30 ^{c,d}	34,7 - 31	-	-	Não	Sim
8	4EC8 - 3,60 ^{c,e}	32,6 - 31	7	Entre fitas $\beta 4$ e $\beta 5$	Sim	Sim
9	4BCF - 3,01 ^{c,d}	32,6 - 31	8	Entre fitas $\beta 4$ e $\beta 5$	Sim	Sim
			5	Alça de ativação		
10	3BLH - 2,48 ^{c,e}	32,6 - 31	4	Alça entre fita $\beta 3$ e αC	Sim	Sim
			10	Entre fitas $\beta 4$ e $\beta 5$		
			5	Alça de ativação		
			7	Alça entre αG e H		
11	6GZH - 3,17 ^{c,d}	32,8 - 30	6	Entre fitas $\beta 4$ e $\beta 5$	Sim	Sim
			3	Alça de ativação		
12	5L1Z - 5,90 ^{c,e}	32,6 - 31	4	Entre fitas $\beta 1$ e $\beta 2$	Sim	Sim
			8	Entre fitas $\beta 4$ e $\beta 5$		
13	4IMY - 2,94 ^{c,d}	32,6 - 31	8	Entre fitas $\beta 4$ e $\beta 5$	Sim	Sim
14	4OR5 - 2,90 ^{c,e}	32,6 - 31	4	Entre fitas $\beta 4$ e $\beta 5$	Sim	Não
15	3MI9 - 2,10 ^{c,e}	32,6 - 31	8	Entre fitas $\beta 4$ e $\beta 5$	Sim	Sim

Swiss-Model

Posição	PDB Resolução /Å	Identidade e Cobertura /%	Nº Resíduos Faltantes e localização ^f		Presença de	
					Fosforilação	Ciclina
1	4BCF - 3,01 ^{c,d}	34,3 - 39	8	Entre fitas $\beta 4$ e $\beta 5$	Sim	Sim
			5	Alça de ativação		
2	3MIA - 3,00 ^{c,d}	34,0 - 40	8	Entre fitas $\beta 4$ e $\beta 5$	Sim	Sim
3	3BLH - 2,48 ^{c,e}	34,4 - 39	4	Alça entre fita $\beta 3$ e αC	Sim	Sim
			10	Entre fitas $\beta 4$ e $\beta 5$		
			5	Alça de ativação		
4	3TN8 - 2,95 ^{c,d}	34,4 - 39	7	Alça entre αG e H	Sim	Sim
			9	Entre fitas $\beta 4$ e $\beta 5$		
5	3MY1 - 2,80 ^{c,d}	34,39 - 39	4	Alça de ativação	Sim	Sim
			7	Entre fitas $\beta 4$ e $\beta 5$		

^a Organismo *Chaetomium thermophilum*

^b Organismo *Saccharomyces cerevisiae*

^c Organismo *Homo sapiens*

^d Possui inibidor no sítio do ATP

^e Não possui inibidor no sítio do ATP

^f Os dados indicados correspondem ao número de resíduos ausentes na estrutura cristalográfica dentro do intervalo do domínio característico das quinases.

A definição dos moldes para a modelagem da CRK9 foi um processo desafiador. Diversos candidatos apresentaram resíduos faltantes em suas estruturas, ou não tinham a ciclina associada, ou mesmo não apresentavam fosforilação, elementos importantes para a atividade da quinase. Em muitos dos candidatos também não havia a presença de um inibidor no sítio do ATP. Dentre as opções disponíveis (**Tabela 18**) foram pré-selecionadas as estruturas de código PDB 1VYW, uma CDK2, 4IMY e 3MI9, ambas CDK9. A terminologia CRK9 (*CDC2-related kinases 9*) e não CDK9 para a quinase de *T. cruzi* e *T. brucei*, dentre outras, se deve a uma alta divergência das sequências com as demais CDKs (*cyclin-dependent kinases*)¹¹³. Para este caso em particular, a escolha final dos moldes foi feita analisando em conjunto o alinhamento múltiplo das sequências de alguns dos candidatos a moldes e analisando com mais detalhes os resíduos do sítio do ATP. Para conseguir um alinhamento satisfatório, principalmente em relação ao lóbulo N, foi preciso utilizar as sequências dos moldes pré-selecionados juntamente com as sequências dos PDBs 6Z3U e 7JV7, conforme **Figura 30**. Neste alinhamento, devido à ausência de cobertura nas porções N (resíduos 1 ao 243) e C terminais (resíduos 688 ao 806) pelos moldes pré-selecionados, estas porções foram omitidas na figura. Além destas regiões, é visível no alinhamento que existem grandes trechos da CRK9 dentro do domínio de quinases que não foram cobertos pelos moldes. A CRK9 de *T. cruzi* parece ser uma quinase bem particular deste organismo, com diversas inserções na sua sequência. Esta característica também foi descrita para a CRK9 de *T. brucei*, sendo reportada a existências de longos trechos de inserções nas extremidades N e C terminais e dentro do domínio de quinases, inclusive na ciclina CYC12 que se associa a esta quinase, indicando que provavelmente esta seja uma característica das CRK9 dos tripanosomatídeos e das ciclinas correspondentes^{111,113}.

```

CRK9          DKKRCILVTGKQCPPCRSVNEYIPKARIAQGVYGVVISAIRATSEVPHHDQEVKRMVKKY 298
3MI9          -----VE-----PFCDEVSKYEKLAKIGQGTGFEVFKARHRKT-----GQKV 38
4IMY          -----VE-----PFCDEVSKYEKLAKIGQGTGFEVFKARHRKT-----GQKV 38
1VYW_1|Chains -----G-----P-LVDMENFQKVEKIGEGTYGVVYKARNKLT-----GEV 35
6Z3U_2|Chains SRPAITPDPVEQ-MNEAEKRKYIKGKLLGEGTYANVYLGHSDDP-----NFKV 119
7JV7_1|Chain  -----KKVPVSVLTQQRSTSVYLRIMQVGEQTYGKVKAKNTNT-----EKL 56
              . : :*:.* * .

CRK9          ALKQVKEQWLVDVQVGFPPYLLREFDLLRM-RHPNIVRAREVVILDKRLGGTKHLRGDA 357
3MI9          ALKKVLM--N-EKEGFPITALREIKILQLL-KHENVVNLEIC----- 78
4IMY          ALKKVLM--N-EKEGFPITALREIKILQLL-KHENVVNLEIC----- 78
1VYW_1|Chains ALKKIRLD--T-ETEGVPSTAIRESLLKEL-NHPNIVKLLDVI----- 75
6Z3U_2|Chains AIKKIKVQ--AQYKDGMAPDAVRELKYLRELGRHPNIIGLISVFS----- 162
7JV7_1|Chain  ALKKLRLQ--G-EREGFPITSIREIKLLQSF-DHPNVSTIKEIMV----- 97
              *:*: : * :*:.* * : * *: .:

CRK9          GGDGVAENNIKRSHPDVAGEFLQGVKGVAPNGNVATQNQEKSGADGVLRGERSRSHVKD 417
3MI9          -----RTKGS 83
4IMY          -----RTKGS 83
1VYW_1|Chains -----HTENK 80
6Z3U_2|Chains -----SKDQN 167
7JV7_1|Chain  -----ESQKT 102

CRK9          VYLVMDYCPY-DLKSFLSLHNEKGVYMHLSYNTHPDAPNNFLSRVKCIMQQLFRALNFL 476
3MI9          IYLVDFDFCEH-DLAGLLSN-VLVK-----FTLSEIKRVMQMLLNGLYYI 125
4IMY          IYLVDFDFCEH-DLAGLLSN-VLVK-----FTLSEIKRVMQMLLNGLYYI 125
1VYW_1|Chains LYLVEFELHQ-DLKKFMDASALTG-----IPLPLIKSYLFQLLQGLAFC 123
6Z3U_2|Chains LNLVLEYLPLGDLEMLIRDVERVR-----YGAADIKAWMGMLTRAVWVC 211
7JV7_1|Chain  VYMIFEYADN-DLSGLLLN-KEVQ-----ISHSQCKHLFKQLLGM EYL 144
              : :*:.* * : * *: .:

CRK9          HDHRILHRDIKTSNILISHEGVVVKLDFGLGRLYREG-----EELTANVVTLMYRAPELH 531
3MI9          HRNKILHRDMKAANVLITRDGVLKLADFGLARAFSLAKNSQPNRY-NRVVTLWYRPEL 184
4IMY          HRNKILHRDMKAANVLITRDGVLKLADFGLARAFSLAKNSQPNRY-NRVVTLWYRPEL 184
1VYW_1|Chains HSHRVLHRDLKPQNLINTEGAIKLADFGLARAFQVVPV---RTYTHEVVTWYRAPEIL 179
6Z3U_2|Chains HENFILHRDIKPNLLIAADGEVVKLADFGLARAFADPG---RRMTANVITRWRPEL 267
7JV7_1|Chain  HDNKILHRDVKGSNILIDNQGKLTDFGLARKMNSR---ADYTNRVITLWYRPEL 199
              * : :*:.* * : * *: .:

CRK9          FGVRDYSHKMDVWSLGCIMAELFLKMPFRAEVVKHFAAICDIIGIPTTEETFSGLYKMP 591
3MI9          LGERDYGPPIDLWGAGCIMAEMWTRSPIMQGNTEQHQLALISQLCGSITPEVWPNVDNYE 244
4IMY          LGERDYGPPIDLWGAGCIMAEMWTRSPIMQGNTEQHQLALISQLCGSITPEVWPNVDNYE 244
1VYW_1|Chains LGCKYYSTAVDIWSLGCIFAEMVTRRALFPGDSEIDQLFRIFRTLGTDPDEVVWPGVTSMP 239
6Z3U_2|Chains FGARHYGGAVDIWSVGMVFAELIIRSPFLPGNTEMEQITLICKHIGTPTEENWPGVSKLP 327
7JV7_1|Chain  LGTTNYGTEVDMWGCGLLVELFNKTAIFQGSNELQIESIFKIMGTPTINSWPTLYDMP 259
              :* * . :*:.* * :*:.* : : . . * : : * * : : .

CRK9          ELLRIMRSLTKYNR-E--NNLSRVFAKSHHAGASTLPASGLDLLQCFIRWNPDLRLSARD 648
3MI9          LYEKL----ELVKGQK--RKVKD-----RLKAYVRDPYALDLIDKLLVLDPAQRIDSDD 292
4IMY          LYEKL----ELVKGQK--RKVKD-----RLKAYVRDPYALDLIDKLLVLDPAQRIDSDD 292
1VYW_1|Chains DYKPS----FPKWA-R--QDFSK-----VVP--PLDEDGRSLLSQMLHYDPNKRISAKA 284
6Z3U_2|Chains EWWDPMEEPIPVWGD---AYMA-----RFG--AVGSEGVLLWRTLQLDPPKKRITARE 376
7JV7_1|Chain  WFFMI----MPQQTTKYVNNFSE-----KFKSVLPSSKCLQLAINLLCYDQTKRFSATE 309
              .* : : .*:.*

CRK9          ALNHVFFHEAPLCEPAELLRPIPTSNVQSRRYDGKATAKRNEDGVGGTPPPCKRPRVRT 708
3MI9          ALNHDFFWSDPMPSDLK--GM-LSTHLSMFEYLAPRR----- 328
4IMY          ALNHDFFWSDPMPSDLK--GM-LST----- 314
1VYW_1|Chains ALAHPFFQDVTKPVPHL--RL-ERPHRD----- 309
6Z3U_2|Chains MLEHRWRTDPKPTRKEDLPK-KSGGEDKMG-----ADLKRPRGMVEDENGPRGS---- 425
7JV7_1|Chain  ALQSDYFKEEPKPEPLVLDGL-VSCHE-----YEVKLARKQKRPNILSTNTN----- 355

```

Figura 30 - Alinhamento produzido no servidor ClustalΩ entre a sequência da quinase CRK9 de *T. cruzi* da cepa CL Brener El e as sequências dos candidatos a moldes para execução de modelagem comparativa. Neste alinhamento foi mostrado apenas o trecho do resíduo 239 ao 708 da CRK9, englobando o domínio de quinases.

Ao analisar a conservação dos resíduos do sítio do ATP dos moldes pré-selecionados em relação à sequência da CRK9 de *T. cruzi*, considerando uma distância de até 6 Å dos inibidores, percebe-se que há uma maior identidade com as sequências dos moldes 4IMY e 3MI9, ambas estruturas da CDK9 humana. A identidade com estes moldes foi de 70,4%, correspondente a 19 resíduos iguais de um total de 27, enquanto a identidade com a estrutura 1VYW nesta região foi de apenas 37,5%, 9 resíduos iguais de 24. Os moldes 4IMY e 3MI9 apresentaram resíduos faltantes entre as fitas β 4 e β 5, região onde o alinhamento indica existir um longo trecho de inserção na CRK9 de *T. cruzi*. Considerando que este trecho não é coberto por nenhuma das potenciais sequências molde, estas estruturas foram definidas como moldes finais para a modelagem comparativa. Vale destacar que a estrutura 4IMY possui inibidor e uma resolução de 2,94 Å, enquanto que a estrutura 3MI9 não possui inibidor, mas apresenta uma resolução mais alta, 2,10 Å, e ambas apresentaram fosforilação na alça de ativação e estão associadas a ciclinas.

A CDK9 humana molde é uma quinase de 372 resíduos de aminoácido. A estrutura cristalográfica 4IMY, já desconsiderando os resíduos das porções N e C terminais que não apresentaram densidade eletrônica suficiente para elucidação estrutural, compreende do resíduos 8 ao 330¹¹⁴. De maneira semelhante, a estrutura 3MI9 inclui do resíduo 8 ao 344¹¹⁵. A identidade e a cobertura tanto do molde 4IMY quanto 3MI9 com a CRK9 de *T. cruzi* foram de 32,6% e 31%, respectivamente. O baixo valor de cobertura se deve em parte aos longos trechos de inserções da CRK9. Ambas as estruturas apresentaram uma fosforilação na treonina 186, na alça de ativação. No alinhamento, a CRK9 ficou coberta por estes moldes do resíduo 244 ao 687, desconsiderando, naturalmente, os trechos de inserções presentes nas extremidades N e C terminais, e também as inserções dentro do domínio de quinase, que não ficaram cobertas.

Após a definição das estruturas molde e a obtenção de um alinhamento adequado, foi conduzida a modelagem comparativa, utilizando o programa Modeller. A validação dos melhores modelos e de suas estruturas refinadas se encontram na **Tabela 19**. O modelo selecionado após análise dos dados foi o CRK9.57 refinado. Esta estrutura apresentou grande parte dos seus resíduos em regiões permitidas no diagrama de Ramachandran do servidor MolProbity, 99,3%, e 93,9% dos resíduos em regiões favoráveis. Neste modelo dois resíduos ficaram classificados como *outliers* no diagrama de Ramachandran, a serina 256 e a valina 521, correspondendo aos 0,7% de resíduos em regiões não permitidas (**Figura 31**). Nenhum destes dois resíduos estão no sítio de ligação do ATP. A estrutura refinada, além de retornar melhores resultados para o Errat e Verify 3D, também melhorou a qualidade quanto aos *clashes*. Os valores de Errat e Verify 3D, 88,7 e

77,3, respectivamente, apesar de não estarem dentro dos valores de corte padrões, são satisfatórios considerando as particularidades da sequência da CRK9, e também a qualidade geral dos moldes (**Figura 31**). O valor de Z-score obtido no servidor ProSA está coerente com os valores das estruturas experimentais de tamanho semelhante.

Tabela 19 - Dados de validação dos 3 modelos pré-selecionados da modelagem da quinase CRK9 de *T. cruzi* da cepa CL Brener El gerados no programa Modeller

Modelos	Modeller		PyMOL/ Å	ProSA	Ramachandran MolProbity /%			Saves / %	
	Dope Score	Molpdf	RMSD	Z-score	Favorável	Permitido	Não permitido	Errat	Verivy 3D
CRK9.34	-50091,3	10960,0	0,231 ^b	-5,30	92,9	98,3	2,3	58,2	74,5
Refinado ^a			0,653 ^c	-4,56	92,9	98,7	1,4	82,1	71,1
CRK9.99	-50585,5	11031,7	0,226 ^b	-4,48	93,6	98,3	2,3	57,0	70,8
Refinado ^a			0,586 ^c	-4,04	92,9	98,7	1,4	83,6	72,4
CRK9.57	-50238,5	11038,9	0,207 ^b	-4,81	94,9	99,0	1,3	60,3	77,0
Refinado ^a			0,608 ^c	-4,48	93,9	99,3	0,7	88,7	77,3
Molde 4IMY				-7,38	93,9	98,7	1,3	95,2	91,1
Molde 3MI9				-7,11	92,3	96,9	3,1	93,6	86,0

^aRefinamento realizado do servidor LocPREFMD (<http://feig.bch.msu.edu/web/services/locprefmd/>)

^bRMSD calculado em relação ao molde 4IMY

^cRMSD calculado em relação ao modelo inicial não refinado

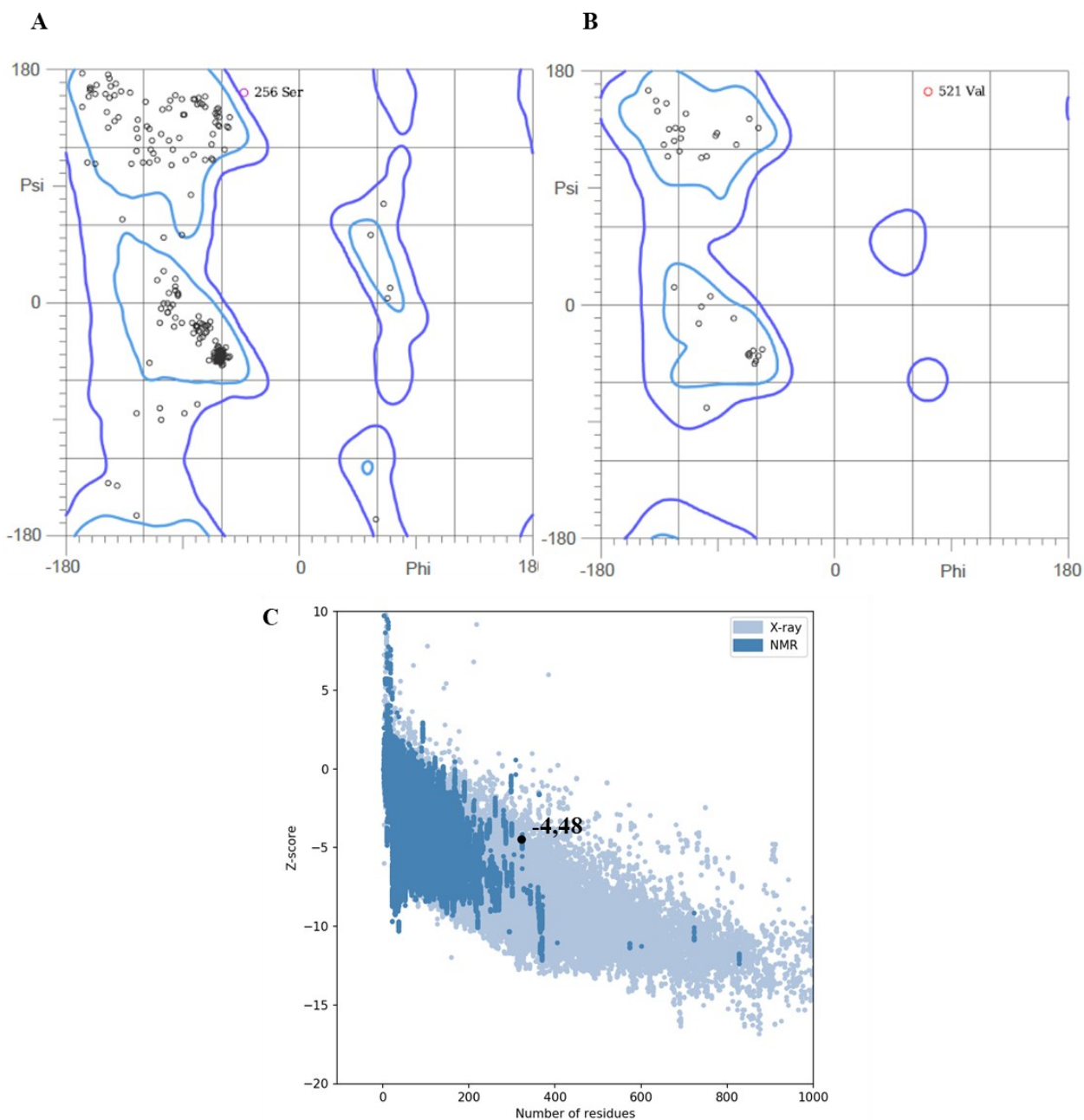


Figura 31 - Validação do modelo CRK9.57 refinado construído por modelagem comparativa para a proteína CRK9 de *T. cruzi* da cepa CL Brener EI, utilizando os moldes 4IMY e 3MI9. **A.** Diagrama de Ramachandran geral gerado no servidor MolProbity para o modelo CRK9.57 refinado. Contornos em azul escuro indicam regiões permitidas, enquanto contornos em azul claro indicam as regiões favoráveis. Os rotâmeros outliers são indicados por círculos vermelhos. **B.** Diagrama de Ramachandran para isoleucinas e valinas gerado no servidor MolProbity para o modelo CRK9.57 refinado. **C.** Gráfico gerado no servidor ProSA, destacando o valor de Z-score para o modelo CRK9.57 refinado em relação a estruturas determinadas experimentalmente, com número de resíduos semelhantes.

A sobreposição das estruturas dos moldes com a do modelo selecionado está ilustrada na **Figura 32**. No painel A, a estrutura do modelo CRK9.57 refinado contém os trechos de inserções

localizados dentro do domínio de quinases que não ficaram cobertos no alinhamento e não foram modelados. Estes trechos foram destacados em azul para uma visão qualitativa de suas localizações. É possível notar uma grande inserção mais à direita, acima do lóbulo N. Nas estruturas cristalográficas das CDKs/CRKs é comum haver ciclina associadas próximas à esta região, indicando que talvez esta longa inserção na CRK9 estabeleça interações com a ciclina correspondente. No painel B foi feita a sobreposição entre as estruturas dos moldes e a estrutura do modelo CRK9.57 refinado sem as inserções, para facilitar a visualização das regiões efetivamente modeladas. No painel C foram destacados os resíduos catalíticos da CRK9 (K301, D485, D503) e as cisteínas mapeadas (C425 (Hinge H2) e C502 (DFG D1)) em sobreposição com os respectivos resíduos catalíticos e cisteína dos moldes. Interessante destacar que o molde apresenta uma cisteína conservada na região do sítio do ATP em relação à cisteína 425 mapeada na quinase CRK9 de *T. cruzi*. Visto que os moldes 4IMY e 3MI9 são estruturas de uma quinase humana, este é um ponto a se ter atenção durante a busca ou o desenvolvimento de inibidores covalentes para a quinase de *T. cruzi* para se conseguir seletividade. O RMSD da sobreposição das estruturas dos moldes e do modelo retornou um valor de 0,617 Å, e houve uma boa sobreposição dos resíduos catalíticos entre as três estruturas, mostrando que, apesar das características da CRK9, o modelo selecionado possui uma boa qualidade na região de interesse. Ao analisar mais detalhadamente o sítio do ATP do modelo CRK9.57 refinado, observa-se a presença de um resíduo classificado como rotâmero *outlier* pelo servidor MolProbity, a isoleucina 266, na região do *P-Loop*. Em relação às estruturas dos moldes, apenas no molde 4IMY este mesmo resíduo, uma isoleucina, que se alinha estruturalmente com a isoleucina 266 da CRK9, também está classificado como rotâmero *outlier* pelo MolProbity (**Figura 32**). Este resíduo não estabelece interações com o composto AMP presente no molde, contudo, é possível que outros compostos, potenciais inibidores da CRK9 de *T. cruzi*, interajam com este resíduo.

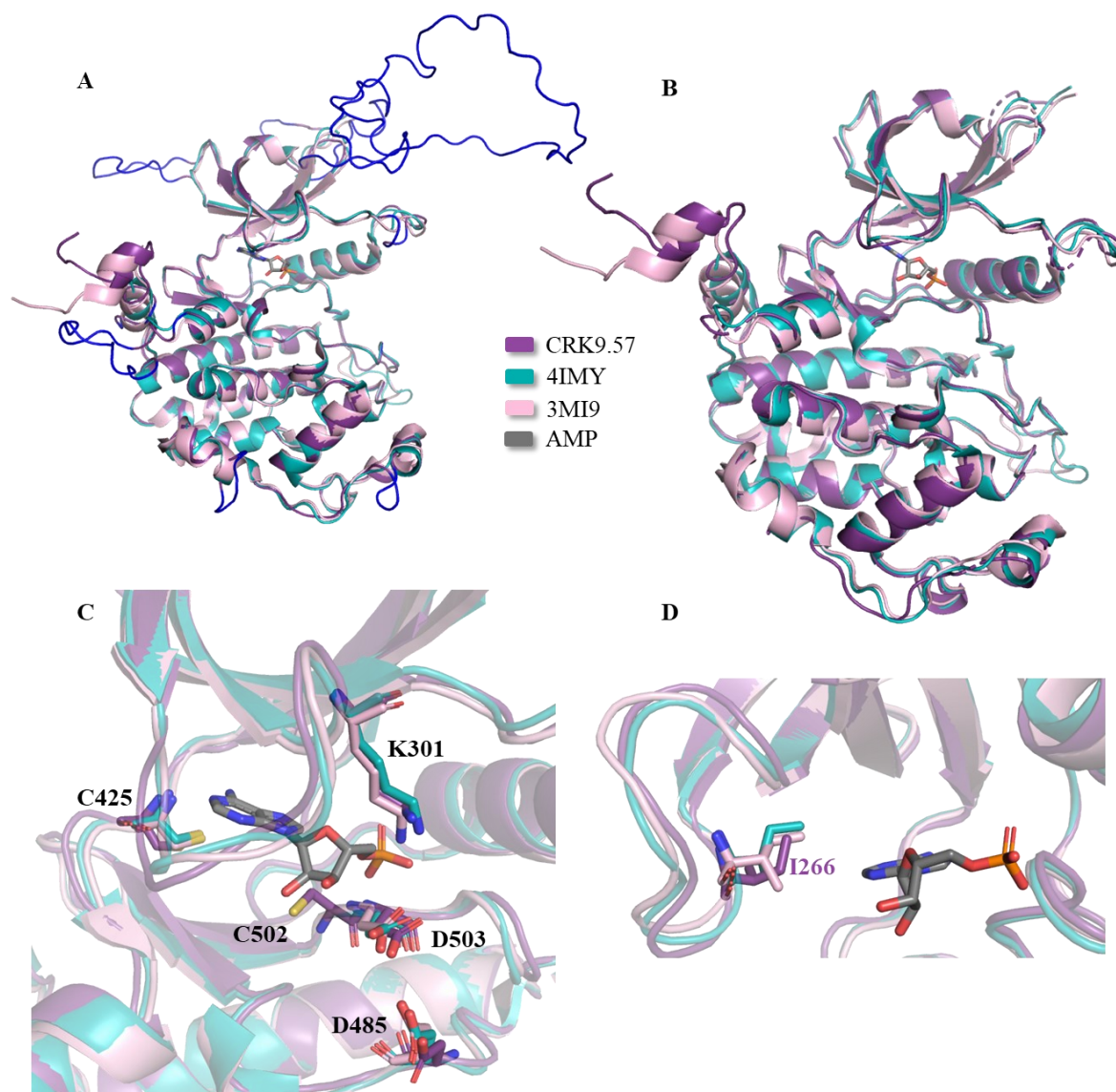


Figura 32 - Estrutura tridimensional do modelo CRK9.57 refinado da quinase CRK9 de *T. cruzi* da cepa CL Brener El gerado por modelagem comparativa em sobreposição com os moldes 4IMY e 3MI9, com destaque para as inserções e para os resíduos catalíticos e cisteínas mapeadas. **A.** Estrutura tridimensional geral do modelo CRK9.57 refinado, **com inserções** dentro do domínio de quinases (destacadas em azul), sobreposta às estruturas dos moldes 4IMY e 3MI9 com a presença do ligante AMP. **B.** Sobreposição do modelo CRK9.57 refinado **sem inserções** com os moldes 4IMY e 3MI9, com a presença do ligante AMP. **C.** Destaque para os 3 resíduos catalíticos da CRK9, K301, D485, D503, e para as 2 cisteína mapeadas C425 (Hinge H2) e C502 (DFG D1) em sobreposição com os respectivos resíduos correspondentes nas estruturas moldes. **D.** Rotâmero *outlier* (I266) na estrutura do modelo CRK9.57 refinado na região do sítio de ligação do ATP em sobreposição com os moldes 4IMY e 3MI9.

Conforme já mencionado, 70,4% dos resíduos que se encontram a até 6 Å de distância do ligante AMP são iguais nos moldes e no modelo (**Figura 33**). Diferentemente do que foi observado nas outras quinases priorizadas, em que a maioria dos resíduos diferentes se encontravam na região de entrada do sítio, em área exposta ao solvente, na CRK9 estes resíduos se encontram

mais distribuídos, muitos em regiões mais internas da proteína. Grande parte dos resíduos diferentes mantiveram as mesmas características de polaridade, porém o molde é ligeiramente mais carregado negativamente (**Figura 33**). Em relação ao tamanho dos resíduos houve certa alternância, ora maiores na CRK9, ora nos moldes, principalmente nas regiões menos enterradas.

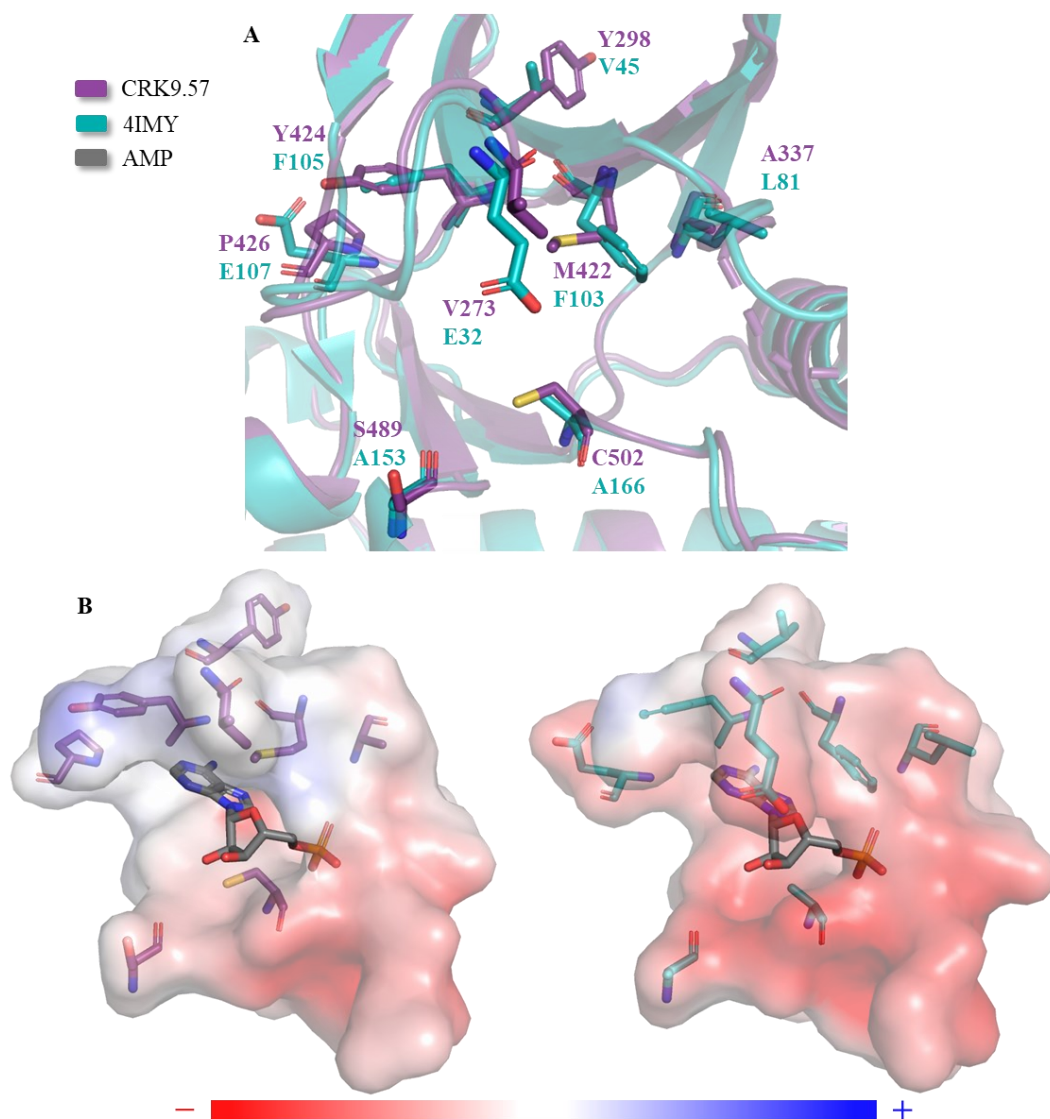


Figura 33 - Destaque das diferenças no sítio de ligação do ATP entre a estrutura do modelo CRK9.57 refinado da quinase CRK9 de *T. cruzi* da cepa CL Brener E1 e a estrutura do molde 4IMY. A estrutura do molde 3MI9 não foi representada nesta figura. **A**. Destaque em bastões dos resíduos diferentes entre o modelo CRK9.57 refinado e o molde 4IMY que se encontram a até 6 Å do ligante AMP. A estrutura da molécula AMP foi omitida para melhor visualização do sítio. **B**. Superfície eletrostática da região analisada para o modelo (à esquerda) e o molde (à direita) gerada pelo *plugin* APBS do PyMOL^{106,107}. A coloração em vermelho indica regiões de potencial eletrostático negativo e em azul regiões de potencial positivo.

Em função da particularidade da sequência da CRK9, a título de comparação e para validar o modelo construído por modelagem comparativa, fez-se uma comparação com o modelo da CRK9 de *T. cruzi* construído pelo preditor de estruturas de proteínas AlphaFold¹⁰⁰. Na **Figura 34**, painel A, está representada a estrutura predita pelo AlphaFold. As inserções, tanto nas porções N e C terminais quanto dentro do domínio de quinase, aparecem majoritariamente como longas alças com alguns pontos de α -hélices e são regiões classificadas principalmente como predição de baixa confiança (em amarelo) ou muito baixa confiança (em alaranjado). A região que representa o domínio de quinases, por sua vez, está classificada em grande parte como predição de confiança (azul claro) ou de alta confiança (azul escuro). A sobreposição entre o modelo CRK9.57 refinado, sem os trechos de inserções, e o modelo gerado pelo AlphaFold, também com remoção das inserções, está ilustrada no painel B. Esta sobreposição retornou um RMSD de 0,880 Å, demonstrando uma boa correlação estrutural entre os dois modelos e mais uma vez sinalizando que o modelo CRK9.57 refinado possui boa qualidade.

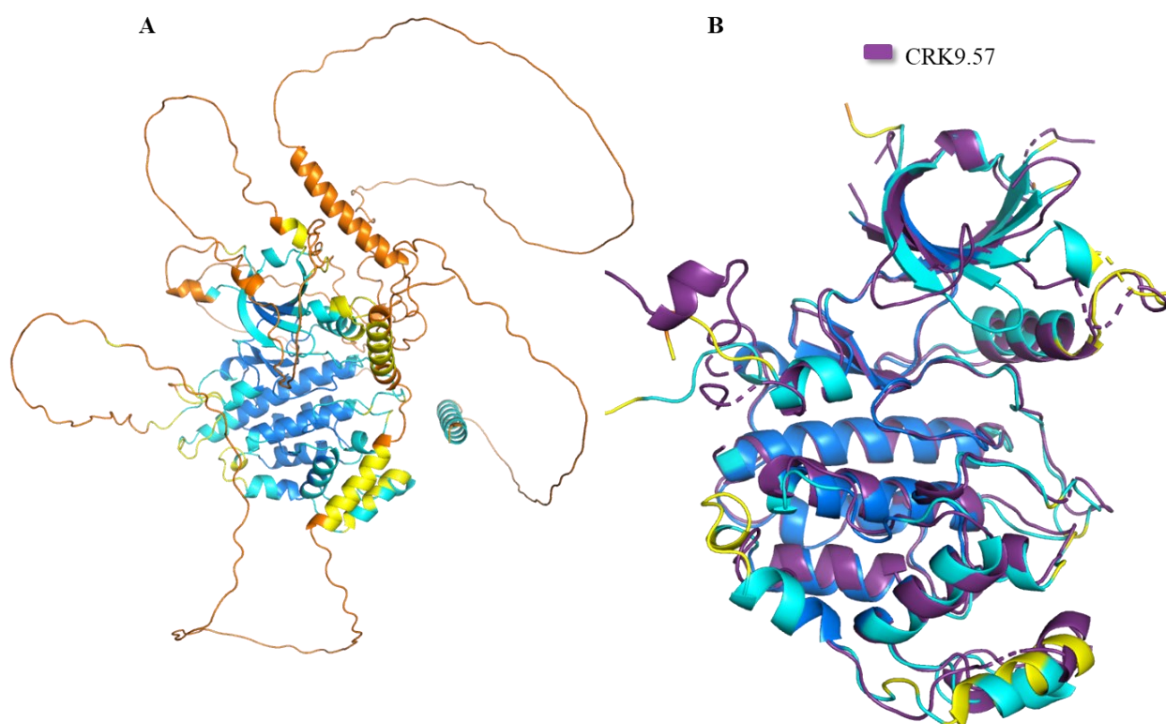


Figura 34 - Comparativo entre o modelo CRK9.57 gerado por modelagem comparativa com o modelo gerado pelo AlphaFold. **A.** Modelo completo gerado pelo AlphaFold. As cores representam o grau de confiabilidade na predição da estrutura terciária, sendo: laranja confiança muito baixa; amarelo confiança baixa; azul claro boa confiança e azul escuro alta confiança. **B.** Sobreposição do modelo CRK9.57 gerado por modelagem comparativa com o modelo gerado pelo AlphaFold sem os trechos de inserção.

A seguir, foi analisada a conservação dos resíduos da CRK9 na região do sítio do ATP em outras cepas do *T. cruzi*. Em relação às cepas apresentadas na **Tabela 17**, há uma conservação de 100% dos resíduos situados a até 6 Å do ligante presente na estrutura molde (**Figura 35**). Como para as modelagens anteriores, esta conservação dos resíduos é um ponto a favor no esforço de encontrar um bom alvo terapêutico para a eliminação do parasito. Contudo, devido à particularidade desta quinase do *T. cruzi*, e de outros tripanossomatídeos, para melhor explorar a CRK9 como alvo de inibidores covalentes, seria interessante tentar obter sua estrutura cristalográfica, o que parece desafiador, ou mesmo outras técnicas que permitam enriquecer as informações sobre a estrutura desta quinase.

<i>Esmeraldo</i>	250	QCPPCRSVNEY	IPKAR	I	AQGVYGVV	I	SA	I	RATSEVP	PHHDQEV	KRMVKK	YAL	K	QVKEQWL	VD	310																																
<i>Non-Esmeraldo</i>	249	QCLPCRSVNEY	IPKAR	I	AQGVYGVV	I	SA	I	RAASEVP	PHHDQEV	KRMVKK	YAL	K	QVKEQWL	VD	309																																
<i>Dm28c-2014</i>	249	QCLPCRSVNEY	IPKAR	I	AQGVYGVV	I	SA	I	RATSEVP	PHHYQEV	KRMVKK	YAL	K	QVKEQWL	VD	309																																
<i>Dm28c-2018</i>	249	QCLPCRSVNEY	IPKAR	I	AQGVYGVV	I	SA	I	RATSEVP	PHHYQEV	KRMVKK	YAL	K	QVKEQWL	VD	309																																
<i>YC6</i>	250	QCPPCRSVNEY	IPKAR	I	AQGVYGVV	I	SA	I	RATSEVP	PHHDQEV	KRMVKK	YAL	K	QVKEQWL	VD	310																																
<i>Sylvio</i>	249	QCLPCRSVNEY	IPKAR	I	AQGVYGVV	I	SA	I	RATSEVP	PHHYQEV	KRMVKK	YAL	K	QVKEQWL	VD	309																																
<i>Brazil</i>	249	QCLPCRSVNEY	IPKAR	I	AQGVYGVV	I	SA	I	RATSEVP	PHHYQEV	KRMVKK	YAL	K	QVKEQWL	VD	309																																
▼																																																
<i>Esmeraldo</i>	311	SQVGFPPYLLR	FDL	L	LRMRHPN	I	VR	A	REVV	SHVKDVY	L	V	M	D	Y	C	P	Y	D	L	K	S	F	L	S	L	H	N	E	K	G	V	Y	442														
<i>Non-Esmeraldo</i>	310	SQVGFPPYLLR	FDL	L	LRMRHPN	I	VR	A	REVV	SHVKDVY	L	V	M	D	Y	C	P	Y	D	L	K	S	F	L	S	L	H	N	E	K	G	V	Y	441														
<i>Dm28c-2014</i>	310	SQVGFPPYLLR	FDL	L	LRMRHPN	I	VR	A	REVV	SHVKDVY	L	V	M	D	Y	C	P	Y	D	L	K	S	F	L	S	L	H	N	E	K	G	V	Y	441														
<i>Dm28c-2018</i>	310	SQVGFPPYLLR	FDL	L	LRMRHPN	I	VR	A	REVV	SHVKDVY	L	V	M	D	Y	C	P	Y	D	L	K	S	F	L	S	L	H	N	E	K	G	V	Y	441														
<i>YC6</i>	311	SQVGFPPYLLR	FDL	L	LRMRHPN	I	VR	A	REVV	SHVKDVY	L	V	M	D	Y	C	P	Y	D	L	K	S	F	L	S	L	H	N	E	K	G	V	Y	442														
<i>Sylvio</i>	310	SQVGFPPYLLR	FDL	L	LRMRHPN	I	VR	A	REVV	SHVKDVY	L	V	M	D	Y	C	P	Y	D	L	K	S	F	L	S	L	H	N	E	K	G	V	Y	441														
<i>Brazil</i>	310	SQVGFPPYLLR	FDL	L	LRMRHPN	I	VR	A	REVV	SHVKDVY	L	V	M	D	Y	C	P	Y	D	L	K	S	F	L	S	L	H	N	E	K	G	V	Y	441														
▼																																																
<i>Esmeraldo</i>	443	MHLSSYNTHPDAPNNF	LSRVKC	I	M	Q	L	F	R	A	L	N	F	L	H	D	H	R	I	L	H	R	I	K	T	S	N	I	L	I	S	H	E	G	V	K	L	C	D	503								
<i>Non-Esmeraldo</i>	442	MHLSSYNTHPDAPNNF	LSRVKC	I	M	Q	L	F	R	A	L	N	F	L	H	D	H	R	I	L	H	R	I	K	T	S	N	I	L	I	S	H	E	G	V	K	L	C	D	502								
<i>Dm28c-2014</i>	442	MHLSSYNTHPDAPNNF	LSRVKC	I	M	Q	L	F	R	A	L	N	F	L	H	D	H	R	I	L	H	R	I	K	T	S	N	I	L	I	S	H	E	G	V	K	L	C	D	502								
<i>Dm28c-2018</i>	442	MHLSSYNTHPDAPNNF	LSRVKC	I	M	Q	L	F	R	A	L	N	F	L	H	D	H	R	I	L	H	R	I	K	T	S	N	I	L	I	S	H	E	G	V	K	L	C	D	502								
<i>YC6</i>	443	MHLSSYNTHPDAPNNF	LSRVKC	I	M	Q	L	F	R	A	L	N	F	L	H	D	H	R	I	L	H	R	I	K	T	S	N	I	L	I	S	H	E	G	V	K	L	C	D	503								
<i>Sylvio</i>	442	MHLSSYNTHPDAPNNF	LSRVKC	I	M	Q	L	F	R	A	L	N	F	L	H	D	H	R	I	L	H	R	I	K	T	S	N	I	L	I	S	H	E	G	V	K	L	C	D	502								
<i>Brazil</i>	442	MHLSSYNTHPDAPNNF	LSRVKC	I	M	Q	L	F	R	A	L	N	F	L	H	D	H	R	I	L	H	R	I	K	T	S	N	I	L	I	S	H	E	G	V	K	L	C	D	502								
▼																																																
<i>Esmeraldo</i>	504	F	G	L	G	R	L	Y	R	E	G	E	L	T	A	N	V	V	T	L	M	Y	R	A	P	E	L	H	F	G	V	R	D	Y	S	H	K	M	D	V	W	S	L	G	C	I	M	806
<i>Non-Esmeraldo</i>	503	F	G	L	G	R	L	Y	R	E	G	E	L	T	A	N	V	V	T	L	M	Y	R	A	P	E	L	H	F	G	V	R	D	Y	S	H	K	M	D	V	W	S	L	G	C	I	M	804
<i>Dm28c-2014</i>	503	F	G	L	G	R	L	Y	R	E	G	E	L	T	A	N	V	V	T	L	M	Y	R	A	P	E	L	H	F	G	V	R	D	Y	S	H	K	M	D	V	W	S	L	G	C	I	M	805
<i>Dm28c-2018</i>	503	F	G	L	G	R	L	Y	R	E	G	E	L	T	A	N	V	V	T	L	M	Y	R	A	P	E	L	H	F	G	V	R	D	Y	S	H	K	M	D	V	W	S	L	G	C	I	M	805
<i>YC6</i>	504	F	G	L	G	R	L	Y	R	E	G	E	L	T	A	N	V	V	T	L	M	Y	R	A	P	E	L	H	F	G	V	R	D	Y	S	H	K	M	D	V	W	S	L	G	C	I	M	806
<i>Sylvio</i>	503	F	G	L	G	R	L	Y	R	E	G	E	L	T	A	N	V	V	T	L	M	Y	R	A	P	E	L	H	F	G	V	R	D	Y	S	H	K	M	D	V	W	S	L	G	C	I	M	805
<i>Brazil</i>	503	F	G	L	G	R	L	Y	R	E	G	E	L	T	A	N	V	V	T	L	M	Y	R	A	P	E	L	H	F	G	V	R	D	Y	S	H	K	M	D	V	W	S	L	G	C	I	M	805

Figura 35 - Alinhamento das sequências das quinases CRK9 correspondentes em outras cepas de *T. cruzi* com destaque para os resíduos que se encontram no sítio do ATP a até 6 Å de distância do ligante AMP (molde 4IMY). As sequências foram alinhadas considerando o intervalo do aminoácido 250 ao 550 da quinase CRK9 cepa CL Brener El. O trecho de grande inserção entre as fitas β 4 e β 5, do aminoácido 342 ao 412, foi omitido. (ponto de corte marcado pela linha e seta em azul).

Uma síntese das 4 modelagens comparativas discutidas nesta seção e que serão úteis para iniciar os estudos de inibição covalentes em quinases de *T. cruzi* está apresentada na **Tabela 20** abaixo.

Tabela 20 - Tabela síntese sobre as 3 quinases de *T. cruzi* priorizadas e de suas respectivas modelagens comparativas

Informações	Quinases Priorizadas			
	AUK1 (Semiativa)	AUK1 (Ativa)	CLK1	CRK9
Grupo	Other		CMGC	CMGC
Cisteínas mapeadas	C109 (Hinge H2) C186 (Other O3)		C232 (Front F1) C305 (DFG D1) C310 (Others O2)	C425 (Hinge H2) C502 (DFG D1)
Código PDB molde	2BFY	1OL5	6Q2A	4IMY e 3MI9
Nome modelo	AUK1.26 refinado	AUK1.56 refinado	CLK1.35 refinado	CRK9.57 refinado
Resíduos em regiões favoráveis (Ramachandran)	96,3%	97,0%	97,3%	93,90%
Identidade no sítio do ATP com as estruturas moldes ^a	63,3%	71,9%	96,2%	70,4%
Função biológica ^b	Divisão celular e possivelmente na diferenciação celular		Divisão celular	Divisão celular e expressão gênica

Identidade de quinases de outras cepas/haplótipo em relação às quinases de CL Brener EI

CL Brener NEI	99%	98%	98%
Dm28c 2014	99%	86%	97%
Dm28c 2018	99%	98%	97%
YC6	100%	99%	99%
Sylvio X10/1-2012	99%	99%	97%
Brazil A4	99%	87%	97%

^a Identidade calculada considerando os resíduos que se encontravam a até 6 Å de distância dos inibidores presentes nas estruturas dos moldes.

^b Apenas a função biológica referente à quinase AUK1 foi obtida de artigos científicos direcionados para *T. cruzi*. Nos demais casos, para as quinases CLK1 e CRK9, não existem publicações direcionadas para *T. cruzi*, os dados são uma extrapolações das informações sobre as quinases do organismo *T. brucei* disponíveis na literatura.

5 CONCLUSÕES E PERSPECTIVAS

Ao longo deste trabalho foi realizado o levantamento das quinases de *T. cruzi* das cepas Dm28c 2014 e CL Brener El em bancos de dados e feita a classificação destas proteínas em grupos e famílias. De acordo com este levantamento e com informações da literatura, este organismo possui de 171 a 175 quinases ePKs e 19 aPKs. Durante a classificação foram obtidos resultados condizentes com o da literatura, demonstrando que *T. cruzi* não possui quinases pertencentes aos grupos TK e TKL. Além disso, em comparação com o quinoma humano, os tripanossomatídeos são pobremente representados pelos grupos CAMK e AGC e apresentam boa representatividade de quinases dos grupos CMGC e STE e da família NEK do grupo Other.

Dado o objetivo primário deste projeto, selecionar alvos terapêuticos que potencialmente sejam inibidos covalentemente, foi feito o mapeamento das cisteínas em posições estratégicas no sítio ativo das quinases das duas cepas de *T. cruzi* a partir de mapeamentos descritos na literatura. Para a cepa Dm28c 2014 foram mapeadas 172 cisteínas distribuídas em 112 quinases, restando 59 sem cisteínas em posições de interesse. Já para a cepa CL Brener El foram identificadas 164 cisteínas distribuídas em 100 quinases, restando 75 proteínas sem cisteínas em posições promissoras para inibição covalente. Estes mapeamentos das cisteínas das duas cepas apresentaram uma boa convergência entre si, ponto importante ao se pensar que o estudo e o desenvolvimento de inibidores para uma quinase de uma determinada cepa provavelmente será capaz de inibir quinases correspondentes em outras cepas. Este mapeamento também permitiu observar que, mesmo se tendo uma distribuição semelhantes de cisteínas nos sítios e subsítios em relação às quinases humanas, em termos proporcionais, algumas posições parecem ser promissoras ao se pensar em seletividade pelas quinases de *T. cruzi*, como as posições P3, H1, F1, F3 e O1.

A partir destes levantamentos, mapeamentos e estudos prévios, juntamente com o conhecimento na literatura acerca do impacto do silenciamento das quinases de *T. brucei* na proliferação, no ciclo celular e na sobrevivência deste organismo, foi possível priorizar quinases de *T. cruzi* para serem trabalhadas como possíveis alvos terapêuticos. Neste processo, foram destacadas três quinases, a AUK1, a CLK1 e a CRK9, cujos silenciamentos levaram à perda de *fitness* no *T. brucei*, culminando na morte do organismo, e que possuíam cisteínas mapeadas nas quinases homólogas em *T. cruzi* em posições promissoras para uma inibição covalente. Apesar de não se poder afirmar que a relevância biológica destas quinases é a mesma nos dois

organismos, a escassez de estudos focados em quinases de *T. cruzi* e a semelhança entre estes organismos são pontos significativos para validar esta estratégia. Os critérios adotados para a definição das quinases priorizadas neste estudo não esgotam a possibilidade de se explorar novos alvos terapêuticos a partir de outros critérios. Durante o silenciamento de outras quinases de *T. brucei* foram observados resultados interessantes, que também levaram à morte do organismo ou ao comprometimento do ciclo celular e da proliferação. A exploração de cisteínas em subsítios que possam conferir maior seletividade é um outro ponto relevante. Todas as informações e análises construídas ao longo deste estudo sobre as proteínas quinases de *T. cruzi* poderão ser utilizadas para embasar novas propostas de alvos terapêuticos.

Por fim, foram construídos por meio de modelagem comparativa os modelos das estruturas tridimensionais das três quinases priorizadas para posteriormente iniciar os estudos com inibidores. As quatro modelagens feitas apresentaram uma qualidade satisfatória para serem usadas em etapas posteriores. Em relação aos dois modelos gerados para a quinase AUK1, um utilizando uma estrutura molde em conformação semiativa e outro em conformação ativa, ambos foram considerados uma vez que as pequenas diferenças no sítio ativo das estruturas obtidas podem interferir nos resultados de *docking*. Sobre a quinase CLK1, as evidências experimentais de inibição do crescimento de *T. cruzi* descritas no trabalho que mostrou a inibição covalente da CLK1 em *T. brucei*, somadas à conservação da cisteína envolvida na inibição covalente na CLK1 de *T. cruzi* e o fato da inibição acontecer no subsítio F1, promissor na busca por seletividade, validaram a CLK1 como um alvo de grande potencial. Quanto à quinase CRK9, cuja modelagem se mostrou desafiadora na fase de seleção das estruturas moldes e de alinhamento, devido às suas particularidades, tem-se exatamente que estas características a torna um alvo terapêutico interessante. Seus diversos trechos de inserções somados com uma variação de identidade nos resíduos do sítio ativo de cerca de 30% em relação aos dos moldes podem contribuir para o alcance de seletividade nas inibições. Esta variação de identidade dos resíduos do sítio do ATP em relação às quinases humanas moldes também irá contribuir para a busca de seletividade para as quinases AUK1 e CLK1, além de outro alvos que também apresentem certa diferença.

Este estudo se encerra apresentando as quinases priorizadas e suas estruturas tridimensionais preditas, juntamente com análises e reflexões importantes para o direcionamento de ações para desenvolver inibidores covalentes com boa seletividade e conseguir eliminar ou controlar o crescimento do parasito causador da doença de Chagas. Neste sentido, uma perspectiva deste trabalho é otimizar as estruturas tridimensionais construídas por modelagem comparativa

através de dinâmica molecular e buscar potenciais inibidores covalentes destas quinases priorizadas. Além disso, a realização de *docking* para a seleção de inibidores promissores e a condução de experimentos para validar os resultados teóricos fechariam o trabalho, contudo, sem esgotar outros desdobramentos que este estudo possa gerar.

REFERÊNCIAS BIBLIOGRÁFICAS

1. Chagas, C. Nova tripanozomíase humana: estudos sobre a morfologia e o ciclo evolutivo do *Schizotrypanum cruzi* n. gen., n. sp., agente etiológico de nova entidade morbida do homem. *Mem. Inst. Oswaldo Cruz* **1**, 159–218 (1909).
2. Coura, J. R., Viñas, P. A. & Junqueira, A. C. V. Ecoepidemiology, short history and control of Chagas disease in the endemic countries and the new challenge for non-endemic countries. *Mem. Inst. Oswaldo Cruz* **109**, 856–862 (2014).
3. Lidani, K. C. F. *et al.* Chagas Disease: From Discovery to a Worldwide Health Problem. *Front. Public Heal.* **7**, 1–13 (2019).
4. Coura, J. R. & Viñas, P. A. Chagas disease: a new worldwide challenge. *Nature* **465**, S6–S7 (2010).
5. Pérez-Molina, J. A. & Molina, I. Chagas disease. *Lancet* **391**, 82–94 (2018).
6. Pan American Health Organization. *Guidelines for the Diagnosis and Treatment of Chagas Disease*. Pan American Health Organization (2019).
7. Echeverria, L. E. & Morillo, C. A. American Trypanosomiasis (Chagas Disease). *Infect. Dis. Clin. North Am.* **33**, 119–134 (2019).
8. Lee, B. Y., Bacon, K. M., Bottazzi, M. E. & Hotez, P. J. Global economic burden of Chagas disease: a computational simulation model. *Lancet Infect. Dis.* **13**, 342–348 (2013).
9. Ministério da Saúde. Doença de Chagas Aguda e distribuição espacial dos triatomíneos de importância epidemiológica, Brasil 2012 a 2016. *Ministério da Saúde. Bol. epidemiológico* **50**, 10 (2019).
10. Schmunis, G. A. & Yadon, Z. E. Chagas disease: A Latin American health problem becoming a world health problem. *Acta Trop.* **115**, 14–21 (2010).
11. Dias, L. C. *et al.* Quimioterapia da doença de Chagas: estado da arte e perspectivas no desenvolvimento de novos fármacos. *Quim. Nova* **32**, 2444–2457 (2009).
12. Field, M. C. *et al.* Erratum: Anti-trypanosomatid drug discovery: an ongoing challenge and a continuing need (Nature reviews. Microbiology (2017) 15 4 (217–231)). *Nat. Rev. Microbiol.* **15**, 447 (2017).
13. Abdeldayem, A., Raouf, Y. S., Constantinescu, S. N., Moriggl, R. & Gunning, P. T. Advances in covalent kinase inhibitors. *Chem. Soc. Rev.* **49**, 2617–2687 (2020).
14. Tomas, A. M. & Kelly, J. M. Transformation as an approach to functional analysis of the major cysteine protease of *Trypanosoma cruzi*. *Biochem. Soc. Trans.* **22**, 3969 (1994).
15. Parsons, M., Worthey, E. A., Ward, P. N. & Mottram, J. C. Comparative analysis of the kinomes of three pathogenic trypanosomatids: *Leishmania major*, *Trypanosoma brucei* and *Trypanosoma cruzi*. *BMC Genomics* **6**, 1–19 (2005).
16. Santori, M. I. *et al.* Evidence for CRK3 participation in the cell division cycle of *Trypanosoma cruzi*. *Mol. Biochem. Parasitol.* **121**, 225–232 (2002).

17. Da Cunha, J. P. C. *et al.* Trypanosoma cruzi histone H1 is phosphorylated in a typical cyclin dependent kinase site accordingly to the cell cycle. *Mol. Biochem. Parasitol.* **140**, 75–86 (2005).
18. Schoijet, A. C. *et al.* A Trypanosoma cruzi phosphatidylinositol 3-kinase (TcVps34) is involved in osmoregulation and receptor-mediated endocytosis. *J. Biol. Chem.* **283**, 31541–31550 (2008).
19. Schoijet, A. C., Sternlieb, T. & Alonso, G. D. The Phosphatidylinositol 3-kinase Class III Complex Containing TcVps15 and TcVps34 Participates in Autophagy in Trypanosoma cruzi. *J. Eukaryot. Microbiol.* **64**, 308–321 (2017).
20. Bao, Y., Weiss, L. M., Braunstein, V. L. & Huang, H. Role of Protein Kinase A in Trypanosoma cruzi. *Infect. Immun.* **76**, 4757–4763 (2008).
21. Bao, Y., Weiss, L. M., Fen, Y., Kahn, S. & Huang, H. Protein kinase A catalytic subunit interacts and phosphorylates members of trans-sialidase super-family in Trypanosoma cruzi. *Microbes Infect.* **12**, 716–726 (2010).
22. Bao, Y. *et al.* Molecular cloning and characterization of mitogen-activated protein kinase 2 in Trypanosoma cruzi. *Cell Cycle* **9**, 2960–2968 (2010).
23. Malvezzi, A. M. *et al.* GCN2-Like Kinase Modulates Stress Granule Formation During Nutritional Stress in Trypanosoma cruzi. *Front. Cell. Infect. Microbiol.* **10**, 1–15 (2020).
24. da Silva Augusto, L. *et al.* A Membrane-bound eIF2 Alpha Kinase Located in Endosomes Is Regulated by Heme and Controls Differentiation and ROS Levels in Trypanosoma cruzi. *PLoS Pathog.* **11**, 1–27 (2015).
25. Souza, C. F. *et al.* Heme-induced Trypanosoma cruzi proliferation is mediated by CaM kinase II. *Biochem. Biophys. Res. Commun.* **390**, 541–546 (2009).
26. de Almeida Nogueira, N. P. *et al.* Heme-Induced ROS in Trypanosoma cruzi activates Camkii-Like that triggers epimastigote proliferation. One helpful effect of ROS. *PLoS One* **6**, (2011).
27. Belaunzarán, M. L., Lammel, E. M., Giménez, G., Wainszelbaum, M. J. & de Isola, E. L. D. Involvement of protein kinase C isoenzymes in Trypanosoma cruzi metacyclogenesis induced by oleic acid. *Parasitol. Res.* **105**, 47–55 (2009).
28. Fassolari, M. & Alonso, G. D. Aurora kinase protein family in Trypanosoma cruzi: Novel role of an AUK-B homologue in kinetoplast replication. *PLoS Negl. Trop. Dis.* **13**, e0007256 (2019).
29. Portal, D. *et al.* Trypanosoma cruzi TcSRPK, the first protozoan member of the SRPK family, is biochemically and functionally conserved with metazoan SR protein-specific kinases. *Mol. Biochem. Parasitol.* **127**, 9–21 (2003).
30. Peng, D., Kurup, S. P., Yao, P. Y., Minning, T. A. & Tarleton, R. L. CRISPR-Cas9-mediated single-gene and gene family disruption in Trypanosoma cruzi. *MBio* **6**, 1–11 (2015).
31. Jones, N. G. *et al.* Regulators of Trypanosoma brucei Cell Cycle Progression and Differentiation Identified Using a Kinome-Wide RNAi Screen. *PLoS Pathog.* **10**, (2014).
32. Fernandez-Cortes, F. *et al.* RNAi screening identifies Trypanosoma brucei stress response protein kinases required for survival in the mouse. *Sci. Rep.* **7**, 1–10 (2017).

33. Endicott, J. A., Noble, M. E. M. & Johnson, L. N. The Structural Basis for Control of Eukaryotic Protein Kinases. *Annu. Rev. Biochem.* **81**, 587–613 (2012).
34. Kanev, G. K. *et al.* The Landscape of Atypical and Eukaryotic Protein Kinases. *Trends Pharmacol. Sci.* **40**, 818–832 (2019).
35. Hanks, S. K., Quinn, A. M. & Hunter, T. The protein kinase family: Conserved features and deduced phylogeny of the catalytic domains. *Science (80-.)*. **241**, 42–52 (1988).
36. Hunter, T. Protein kinase classification. *Methods Enzymol.* **200**, 3–37 (1991).
37. Roskoski, R. Orally effective FDA-approved protein kinase targeted covalent inhibitors (TCIs). *Pharmacol. Res.* **165**, 105422 (2021).
38. Hanks, S. K. & Hunter, T. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J.* **9**, 576–96 (1995).
39. Hanks, S. K. & Quinn, A. M. [2] Protein kinase catalytic domain sequence database: Identification of conserved features of primary structure and classification of family members. *Methods Enzymol.* **200**, 38–62 (1991).
40. Kornev, A. P., Haste, N. M., Taylor, S. S. & Ten Eyck, L. F. Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. *Proc. Natl. Acad. Sci.* **103**, 17783–17788 (2006).
41. Manning, G., Whyte, D. B., Martinez, R., Hunter, T. & Sudarsanam, S. The Protein Kinase Complement of the Human Genome. *Science (80-.)*. **298**, 1912–1934 (2002).
42. Wikikinome. Kinase classification. Disponível em: <http://kinase.com/wiki/index.php/Kinase_classification>. Acesso em: 07 de março de 2022.
43. Gerard Manning's lab at Genentech Inc. Kinase.com. Disponível em: <<http://kinase.com/web/current/>>. Acesso em 07 de março de 2022
44. Cell Signaling Technology. Protein Kinases. Disponível em: <<https://www.cellsignal.com/learn-and-support/protein-kinases>>. Acesso em 07 de março de 2022.
45. Kornev, A. P. & Taylor, S. S. Defining the conserved internal architecture of a protein kinase. *Biochim. Biophys. Acta - Proteins Proteomics* **1804**, 440–444 (2010).
46. Roskoski, R. Classification of small molecule protein kinase inhibitors based upon the structures of their drug-enzyme complexes. *Pharmacol. Res.* **103**, 26–48 (2016).
47. Huse, M. & Kuriyan, J. The Conformational Plasticity of Protein Kinases. *Cell* **109**, 275–282 (2002).
48. Yu, L. *et al.* Role of Mg²⁺ ions in protein kinase phosphorylation: Insights from molecular dynamics simulations of ATP-kinase complexes. *Mol. Simul.* **37**, 1143–1150 (2011).
49. Recabarren, R., Zinovjev, K., Tuñón, I. & Alzate-Morales, J. How a Second Mg²⁺ Ion Affects the Phosphoryl-Transfer Mechanism in a Protein Kinase: A Computational Study. *ACS Catal.* **11**, 169–183 (2021).
50. Roskoski, R. Properties of FDA-approved small molecule protein kinase inhibitors: A

- 2020 update. *Pharmacol. Res.* **152**, 104609 (2020).
51. Nolen, B., Taylor, S. & Ghosh, G. Regulation of protein kinases: Controlling activity through activation segment conformation. *Mol. Cell* **15**, 661–675 (2004).
 52. Adams, J. A. Kinetic and catalytic mechanisms of protein kinases. *Chem. Rev.* **101**, 2271–2290 (2001).
 53. Xie, T., Saleh, T., Rossi, P. & Kalodimos, C. G. Conformational states dynamically populated by a kinase determine its function. *Science (80-.).* **370**, 1–37 (2020).
 54. Wang, Z. & Cole, P. A. Catalytic mechanisms and regulation of protein kinases. *Methods Enzymol.* **548**, 1–21 (2014).
 55. Rakshambikai, R., Manoharan, M., Gnanavel, M. & Srinivasan, N. Typical and atypical domain combinations in human protein kinases: Functions, disease causing mutations and conservation in other primates. *RSC Adv.* **5**, 25132–25148 (2015).
 56. Lahiry, P., Torkamani, A., Schork, N. J. & Hegele, R. A. Kinase mutations in human disease: Interpreting genotype-phenotype relationships. *Nat. Rev. Genet.* **11**, 60–74 (2010).
 57. Zhao, Z. & Bourne, P. E. Progress with covalent small-molecule kinase inhibitors. *Drug Discov. Today* **23**, 727–735 (2018).
 58. Carles, F., Bourg, S., Meyer, C. & Bonnet, P. PKIDB: A curated, annotated and updated database of protein kinase inhibitors in clinical trials. *Molecules* **23**, 1–18 (2018).
 59. Schwöbel, J. A. H., Madden, J. C. & Cronin, M. T. D. Examination of Michael addition reactivity towards glutathione by transition-state calculations. *SAR QSAR Environ. Res.* **21**, 693–710 (2010).
 60. Mah, R., Thomas, J. R. & Shafer, C. M. Drug discovery considerations in the development of covalent inhibitors. *Bioorganic Med. Chem. Lett.* **24**, 33–39 (2014).
 61. Zhao, Z., Liu, Q., Bliven, S., Xie, L. & Bourne, P. E. Determining Cysteines Available for Covalent Inhibition Across the Human Kinome. *J. Med. Chem.* **60**, 2879–2889 (2017).
 62. Chaikuad, A., Koch, P., Laufer, S. A. & Knapp, S. The Cysteinome of Protein Kinases as a Target in Drug Development. *Angew. Chemie Int. Ed.* **57**, 4372–4385 (2018).
 63. Muhammed, M. T. & Aki-Yalcin, E. Homology modeling in drug discovery: Overview, current applications, and future perspectives. *Chem. Biol. Drug Des.* **93**, 12–20 (2019).
 64. Bitar, M. & Franco, G. R. A basic protein comparative three-dimensional modeling methodological workflow theory and practice. *IEEE/ACM Trans. Comput. Biol. Bioinforma.* **11**, 1052–1065 (2014).
 65. Verli, H. *Bioinformática : da Biologia à Flexibilidade Molecular.* (Sociedade Brasileira de Bioquímica e Biologia Molecular - SBBq, 2014).
 66. Haddad, Y., Adam, V. & Heger, Z. Ten quick tips for homology modeling of high-resolution protein 3D structures. *PLoS Comput. Biol.* **16**, 1–19 (2020).
 67. Baker, D. & Sali, A. Protein structure prediction and structural genomics. *Science (80-.).* **294**, 93–96 (2001).

68. Šali, A. & Blundell, T. L. Comparative Protein Modelling by Satisfaction of Spatial Restraints. *J. Mol. Biol.* **234**, 779–815 (1993).
69. Fiser, A., Do, R. K. G. & Šali, A. Modeling of loops in protein structures. *Protein Sci.* **9**, 1753–1773 (2000).
70. Webb, B. & Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc. Bioinforma.* **54**, 100–106 (2016).
71. Altschul, S. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
72. Berman, H. M. *et al.* The protein data bank. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **58**, 899–907 (2002).
73. Shen, M. & Sali, A. Statistical potential for assessment and prediction of protein structures. *Protein Sci.* **15**, 2507–2524 (2006).
74. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 12–21 (2010).
75. Williams, C. J. *et al.* MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315 (2018).
76. Colovos, C. & Yeates, T. O. Verification of protein structures: Patterns of nonbonded atomic interactions. *Protein Sci.* **2**, 1511–1519 (1993).
77. Bowie, J., Luthy, R. & Eisenberg, D. A method to identify protein sequences that fold into a known three-dimensional structure. *Science (80-)*. **253**, 164–170 (1991).
78. Lüthy, R., Bowie, J. U. & Eisenberg, D. Assessment of protein models with three-dimensional profiles. *Nature* **356**, 83–85 (1992).
79. Sippl, M. J. Recognition of errors in three-dimensional structures of proteins. *Proteins Struct. Funct. Bioinforma.* **17**, 355–362 (1993).
80. Wiederstein, M. & Sippl, M. J. ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* **35**, 407–410 (2007).
81. Eidhammer, I., Jonassen, I. & Taylor, W. R. *Protein Bioinformatics: An Algorithmic Approach to Sequence and Structure Analysis*. Wiley (John Wiley & Sons, Inc., 2003). doi:10.1002/9780470092620.
82. Bateman, A. *et al.* UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Res.* **49**, D480–D489 (2021).
83. Sonnhammer, E. L. L., Eddy, S. R. & Durbin, R. Pfam: A comprehensive database of protein domain families based on seed alignments. *Proteins Struct. Funct. Genet.* **28**, 405–420 (1997).
84. Mistry, J. *et al.* Pfam: The protein families database in 2021. *Nucleic Acids Res.* **49**, D412–D419 (2021).
85. Grisard, E. C. *et al.* Trypanosoma cruzi Clone Dm28c Draft Genome Sequence. *Genome Announc.* **2**, 2–3 (2014).
86. Aslett, M. *et al.* TriTrypDB: A functional genomic resource for the Trypanosomatidae.

- Nucleic Acids Res.* **38**, 457–462 (2009).
87. El-Sayed, N. M. *et al.* The Genome Sequence of *Trypanosoma cruzi*, Etiologic Agent of Chagas Disease. *Science* (80-.). **309**, 409–415 (2005).
 88. Benson, D., Boguski, M., Lipman, D. J. & Ostell, J. The National Center for Biotechnology Information. *Genomics* **6**, 389–391 (1990).
 89. Agarwala, R. *et al.* Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **46**, D8–D13 (2018).
 90. Guex, N. & Peitsch, M. C. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis* **18**, 2714–2723 (1997).
 91. Waterhouse, A. *et al.* SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303 (2018).
 92. Bultet, L. A. *et al.* The SIB Swiss Institute of bioinformatics' resources: Focus on curated databases. *Nucleic Acids Res.* **44**, D27–D37 (2016).
 93. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, (2011).
 94. Feig, M. Local Protein Structure Refinement via Molecular Dynamics Simulations with locPREFMD. *J. Chem. Inf. Model.* **56**, 1304–1312 (2016).
 95. Ramachandran, G. N. & Sasisekharan, V. Conformation of Polypeptides and Proteins. *Adv. Protein Chem.* **23**, 283–437 (1968).
 96. Kufareva, I. & Abagyan, R. Methods of protein structure comparison. *Methods Mol. Biol.* **857**, 231–257 (2012).
 97. The PyMOL Molecular Graphics System, Version 2.4.1 Schrödinger, LLC.
 98. Hayward, S. & Berendsen, H. J. C. Systematic analysis of domain motions in proteins from conformational change: New results on citrate synthase and T4 lysozyme. *Proteins Struct. Funct. Genet.* **30**, 144–154 (1998).
 99. Volkamer, A., Griewel, A., Grombacher, T. & Rarey, M. Analyzing the topology of active sites: On the prediction of pockets and subpockets. *J. Chem. Inf. Model.* **50**, 2041–2052 (2010).
 100. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
 101. Tu, X., Kumar, P., Li, Z. & Wang, C. C. An aurora kinase homologue is involved in regulating both mitosis and cytokinesis in *Trypanosoma brucei*. *J. Biol. Chem.* **281**, 9677–9687 (2006).
 102. Li, Z. & Wang, C. C. Changing roles of aurora-B kinase in two life cycle stages of *Trypanosoma brucei*. *Eukaryot. Cell* **5**, 1026–1035 (2006).
 103. Li, Z. *et al.* Identification of a novel chromosomal passenger complex and its unique localization during cytokinesis in *Trypanosoma brucei*. *PLoS One* **3**, (2008).
 104. Siman-Tov, M. M., Ivens, A. C. & Jaffe, C. L. Identification and cloning of Lmairk, a member of the Aurora/Ipl1p protein kinase family, from the human protozoan parasite *Leishmania*. *Biochim. Biophys. Acta - Gene Struct. Expr.* **1519**, 241–245 (2001).

105. Sessa, F. *et al.* Mechanism of Aurora B activation by INCENP and inhibition by hesperadin. *Mol. Cell* **18**, 379–391 (2005).
106. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci.* **98**, 10037–10041 (2001).
107. Dolinsky, T. J., Nielsen, J. E., McCammon, J. A. & Baker, N. A. PDB2PQR: An automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* **32**, 665–667 (2004).
108. Bayliss, R., Sardon, T., Vernos, I. & Conti, E. Structural basis of Aurora-A activation by TPX2 at the mitotic spindle. *Mol. Cell* **12**, 851–862 (2003).
109. Saldivia, M. *et al.* A CLK1-KKT2 Signaling Pathway Regulating Kinetochore Assembly in *Trypanosoma brucei*. *MBio* **12**, e0068721 (2021).
110. Saldivia, M. *et al.* Targeting the trypanosome kinetochore with CLK1 protein kinase inhibitors. *Nat. Microbiol.* **5**, 1207–1216 (2020).
111. Gourguechon, S. & Wang, C. C. CRK9 contributes to regulation of mitosis and cytokinesis in the procyclic form of *Trypanosoma brucei*. *BMC Cell Biol.* **10**, 68 (2009).
112. Badjatia, N., Ambrósio, D. L., Lee, J. H. & Günzl, A. Trypanosome cdc2-Related Kinase 9 Controls Spliced Leader RNA cap4 Methylation and Phosphorylation of RNA Polymerase II Subunit RPB1. *Mol. Cell. Biol.* **33**, 1965–1975 (2013).
113. Badjatia, N., Park, S. H., Ambrósio, D. L., Kirkham, J. K. & Günzl, A. Cyclin-Dependent Kinase CRK9, Required for Spliced Leader trans Splicing of Pre-mRNA in Trypanosomes, Functions in a Complex with a New L-Type Cyclin and a Kinetoplastid-Specific Protein. *PLoS Pathog.* **12**, 1–27 (2016).
114. Schulze-Gahmen, U. *et al.* The AFF4 scaffold binds human P-TEFb adjacent to HIV Tat. *Elife* **2013**, 1–14 (2013).
115. Tahirov, T. H. *et al.* Crystal structure of HIV-1 Tat complexed with human P-TEFb. *Nature* **465**, 747–751 (2010).

APÊNDICE A - Tabela de classificação em grupos e famílias das 171 quinases de *T. cruzi* da cepa Dm28c 2014, e detalhamento do mapeamento das cisteínas em sítios e subsítios.

Código TriTryp	Código Uniprot	Classificação		Mapeamento Cisteínas			
		Grupo	Família	Sítio	Subsítio	Resíduo PDB 1ATP	Resíduo Dm28c 2014
TCDM_06552	V5DCI0	Unique		P-Loop P	P1	L49	C468
				Beta Sheet 4	Beta 4-4	L106	C520
TCDM_01720	V5BNT1	STE	STE11	P-Loop P	P2	S53	C651
				P-Loop P	P3	F54	C652
				Hinge	H2	V123	C726
				Others	O2	F187	C872
TCDM_01664	V5DQG8	Unique		P-Loop P	P3	F54	C786
				Hinge	H2	V123	C854
				Beta Sheet 4	Beta 4-4	L106	C835
				Helix	Helix-11	V98	C827
TCDM_05014	V5BP88	STE	STE11	P-Loop P	P3	F54	C204
				Front	F3	F129	C279
TCDM_00599	V5C106	Other	CAMKK	P-Loop P	P3	F54	C208
				Front	F4	R133	C297
TCDM_08213	V5B871	Other	PLK	P-Loop P	P4	V57	C62
				Hinge	H2	V123	C126
				Others	O3	C199	C205
TCDM_06718	V5DC27	Other	NEK	P-Loop P	P4	V57	C18
TCDM_06298	V5AX03	CMGC	MAPK	DFG	D1	T183	C173
TCDM_09858	V5B8U8	CMGC	MAPK	DFG	D1	T183	C173
TCDM_01142	V5BQE9	CAMK	CAMKL	DFG	D1	T183	C150
				Front	F4	R133	C99
				Others	O3	C199	C168
TCDM_02136	V5BSG8	CMGC	MAPK	DFG	D1	T183	C188
TCDM_06097	V5BLT9	CMGC	MAPK	DFG	D1	T183	C167
TCDM_00215	V5BTI4	STE	STE20	DFG	D1	T183	C162
				Others	O2	A188	C167
TCDM_02524	V5BR38	CMGC	GSK	DFG	D1	T183	C238
TCDM_03065	V5BK24	Other	NAK	DFG	D1	T183	C212
				Hinge	H2	V123	C147
TCDM_11995	V5CZ20	CMGC	MAPK	DFG	D1	T183	C156
				Beta Sheet 4	Beta 4-4	L106	C76
TCDM_08982	V5BFK9	Unique		DFG	D1	T183	C153
				Hinge	H1	Y122	C85
TCDM_03791	V5BN23	CMGC	CLK	DFG	D1	T183	C316
				Front	F1	G126	C243
				Others	O2	A188	C321

Código TriTryp	Código Uniprot	Classificação		Mapeamento Cisteínas			
		Grupo	Família	Sítio	Subsítio	Resíduo PDB 1ATP	Resíduo Dm28c 2014
TCDM_07669	V5DA09	CMGC	MAPK	DFG	D1	T183	C175
TCDM_02318	V5B6E9	CMGC	GSK	DFG	D1	T183	C170
TCDM_00436	V5BBP6	STE	STE11	DFG	D1	T183	C444
				Others	O2	F187	C448
TCDM_05476	V5BIQ0	CMGC	CDK	DFG	D1	T183	C501
				Hinge	H2	V123	C424
TCDM_11672	V5B472	Other	NAK	DFG	D1	T183	C189
TCDM_03670	V5BN97	CAMK	CDPK	DFG	D1	T183	C161
TCDM_06063	V5BCN4	CAMK	CAMKL	DFG	D1	T183	C154
TCDM_04773	V5BFQ4	Unique		DFG	D1	T183	C1181
				Others	O1	V104	C1038
TCDM_06874	V5AVM7	CMGC	MAPK	DFG	D1	T183	C260
TCDM_05453	V5BIN3	CMGC	CLK	DFG	D1	T183	C175
				Others	O1	V104	C86
TCDM_05440	V5AZ01	CMGC	DYRK	Roof R	R1	A70	C447
TCDM_11204	V5D179	STE	STE11	Roof R	R1	A70	C370
				Others	O2	F187	C489
TCDM_04234	V5DIC3	CMGC	DYRK	Roof R	R1	A70	C287
TCDM_04627	V5B0J4	STE	(na)	Hinge	H1	Y122	C1342
TCDM_01147	V5BQF4	STE	STE7	Hinge	H1	Y122	C133
TCDM_02221	V5BRZ6	Unique		Hinge	H1	Y122	C141
				Hinge	H2	V123	C142
				Catalytic Loop	Catalytic-2	L162	C180
TCDM_03274	V5DKM4	Unique		Hinge	H2	V123	C233
TCDM_00296	V5C234	Other	NEK	Hinge	H2	V123	C138
				Others	O2	A188	C207
TCDM_06053	V5DDS8	Unique		Hinge	H2	V123	C1351
TCDM_01433	V5B8Z2	Unique		Hinge	H2	V123	C743
				Catalytic Loop	Catalytic-2	L162	C781
				Hinge	H2	V123	C181
TCDM_00078	V5BTA3	Unique		Catalytic Loop	Catalytic-2	L162	C220
TCDM_01191	V5BQH6	AGC	NDR	Hinge	H2	V123	C142
TCDM_10053	V5D488	CMGC	CDK	Hinge	H2	V123	C86
TCDM_01417	V5DQL7	CMGC	CDK	Hinge	H2	V123	C126
TCDM_00536	V5DTK9	CMGC	(na)	Hinge	H2	V123	C369
TCDM_07838	V5D9M1	Other	NEK	Hinge	H2	V123	C89
				Beta Sheet 4	Beta 4-4	L106	C70
TCDM_07319	V5DAW8	Other	WEE	Hinge	H2	V123	C270
				Others	O2	F187	C338

Código TriTryp	Código Uniprot	Classificação		Mapeamento Cisteínas			
		Grupo	Família	Sítio	Subsítio	Resíduo PDB 1ATP	Resíduo Dm28c 2014
TCDM_00456	V5BBR1	Other	NEK	Hinge	H2	V123	C83
				Others	O3	C199	C161
TCDM_06526	V5DCN7	CMGC	CDK	Hinge	H2	V123	C137
TCDM_09001	V5BAZ8	AGC	(na)	Hinge	H2	V123	C98
				Others	O3	C199	C203
				Helix	Helix-11	V98	C73
TCDM_00579	V5BB10	Other	PEK	Hinge	H2	V123	C433
TCDM_04577	V5DH61	Other	ULK	Hinge	H2	V123	C78
TCDM_04630	V5BKL1	CAMK	CAMKL	Hinge	H2	V123	C90
				Others	O2	A188	C178
TCDM_11145	V5B141	STE	STE20	Hinge	H2	V123	C98
				Helix	Helix-11	V98	C73
TCDM_07156	V5BJQ6	STE	STE11	Hinge	H2	V123	C812
TCDM_00410	V5BXH2	Unique		Hinge	H2	V123	C446
TCDM_13183	V5A3J0	CMGC	CDK	Hinge	H2	V123	C128
				Catalytic Loop	Catalytic-9	L172	C179
TCDM_13551	V5CI11	Other	NEK	Hinge	H2	V123	C108
TCDM_05578	V5DEK0	Other	SCY1	Hinge	H2	V123	C104
				Beta Sheet 4	Beta 4-4	L106	C87
TCDM_11759	V5AZK0	Other	CAMKK	Hinge	H2	V123	C386
				Catalytic Loop	Catalytic-9	L172	C433
TCDM_02055	V5BSA6	Other	PLK	Hinge	H2	V123	C257
				Others	O3	C199	C336
TCDM_06317	V5DCX1	Other	NEK	Hinge	H2	V123	C138
TCDM_04589	V5BG98	Other	Aur	Hinge	H2	V123	C109
				Others	O3	C199	C186
TCDM_09216	V5B639	STE	ST11	Hinge	H2	V123	C489
				Others	O2	F187	C557
TCDM_04306	V5DI55	Other	Aur	Hinge	H2	V123	C202
				Others	O3	C199	C313
TCDM_08140	V5B8E0	Other	NEK	Hinge	H2	V123	C115
TCDM_05599	V5BID9	Unique		Hinge	H2	V123	C335
TCDM_04864	V5BPK4	AGC	(na)	Hinge	H3	A124	C332
				Others	O3	C199	C410
TCDM_02924	V5BUV8	CMGC	DYRK	Hinge	H3	A124	C83
TCDM_01763	V5BTF4	STE	(na)	Hinge	H3	G125	C960
				Others	O2	F187	C1022
TCDM_01732	V5BTE3	STE	STE11	Hinge	H3	G125	C222
				Others	O2	F187	C304

Código TriTryp	Código Uniprot	Classificação		Mapeamento Cisteínas			
		Grupo	Família	Sítio	Subsítio	Resíduo PDB 1ATP	Resíduo Dm28c 2014
TCDM_03768	V5DJD7	CMGC	CDK	Hinge	H3	G125	C464
				Others	O2	A188	C526
TCDM_00502	V5DTI3	CAMK	(na)	Hinge	H3	G125	C123
				Front	F3	F129	C127
TCDM_02571	V5DMJ4	CAMK	(na)	Front	F3	F1129	C90
TCDM_00262	V5BX61	STE	(na)	Front	F4	R133	C150
TCDM_01931	V5DPP5	AGC	NDR	Front	F4	R133	C161
TCDM_06510	V5BL06	CAMK	CAMKL	Others	O1	V104	C85
				Others	O3	C199	C182
TCDM_06936	V5BFI8	STE	STE11	Others	O1	V104	C231
TCDM_00633	V5BB60	STE	(na)	Others	O2	F187	C266
				Beta Sheet 4	Beta 4-4	L106	C172
TCDM_04422	V5BQQ4	STE	(na)	Others	O2	F187	C528
TCDM_02619	V5BLL5	Unique		Others	O2	A188	C225
TCDM_01982	V5B7K6	Unique		Others	O2	A188	C394
				Others	O3	C199	C423
TCDM_01679	V5BPE2	CMGC	RCK	Others	O2	A188	C150
TCDM_00275	V5BSK0	Unique		Others	O2	A188	C358
TCDM_05504	V5DF06	STE	(na)	Others	O2	A188	C696
				Beta Sheet 4	Beta 4-4	L106	C610
TCDM_01526	V5DQT5	AGC	(na)	Others	O3	C199	C292
TCDM_01501	V5BU83	AGC	(na)	Others	O3	C199	C245
TCDM_06376	V5BGP8	CAMK	(na)	Others	O3	C199	C84
TCDM_04375	V5DHW8	Other	NEK	Others	O3	C199	C208
TCDM_07404	V5DAM3	CAMK	(na)	Others	O3	C199	C652
TCDM_08224	V5BH88	CAMK	(na)	Others	O3	C199	C189
				Catalytic Loop	Catalytic-9	L172	C159
TCDM_08522	V5BC61	Other	Aur	Others	O3	C199	C234
TCDM_05312	V5DFA0	Unique		Others	O3	C199	C511
TCDM_05306	V5BEK5	AGC	PKA	Others	O3	C199	C176
TCDM_03208	V5BPJ0	Other	NEK	Others	O3	C199	C205
				Beta Sheet 4	Beta 4-4	L106	C95
				Helix	Helix-11	V98	C87
TCDM_08762	V5D7G4	Other	NEK	Others	O3	C199	C582
				Helix	Helix-11	V98	C472
TCDM_02579	V5BR65	Other	NEK	Others	O3	C199	C230
				Helix	Helix-11	V98	C100
TCDM_01769	V5BNV8	CAMK	CAMKL	Others	O3	C199	C173
TCDM_09739	V5D579	Other	NEK	Others	O3	C199	C184
				Helix	Helix-11	V98	C73

Código TriTryp	Código Uniprot	Classificação		Mapeamento Cisteínas			
		Grupo	Família	Sítio	Subsítio	Resíduo PDB 1ATP	Resíduo Dm28c 2014
TCDM_02360	V5BWC2	CAMK	CAMKL	Others	O3	C199	C166
TCDM_04221	V5BLW4	Other	NEK	Others	O3	C199	C164
TCDM_08577	V5BC28	Other	NEK	Others	O3	C199	C267
				Helix	Helix-11	V98	C157
TCDM_00285	V5C224	Other	NEK	Others	O3	C199	C184
TCDM_07910	V5BI43	CAMK	(na)	Others	O3	C199	C209
TCDM_07266	V5DAY8	Unique		Others	O3	C199	C219
TCDM_08513	V5BC57	CAMK	(na)	Others	O3	C199	C302
TCDM_06093	V5DDI8	STE	STE11	Others	O3	C199	C501
				Beta Sheet 4	Beta 4-4	L106	C386
TCDM_07107	V5BFB5	AGC	(na)	Beta Sheet 4	Beta 4-4	L106	C418
TCDM_06732	V5DC40	STE	STE11	Beta Sheet 4	Beta 4-4	L106	C277
TCDM_04729	V5B0G8	Unique		Beta Sheet 4	Beta 4-4	L106	C202
TCDM_03374	V5BNY0	CMGC	(na)	DFG Peptide	DFG-4	G186	C132
TCDM_00304	V5BSM0	Other	(na)	Helix	Helix-11	V98	C179
TCDM_03205	V5DKY9	Other	NEK	Helix	Helix-11	V98	C182
TCDM_00033	V5DUH0	STE	STE20	Catalytic Loop	Catalytic-2	L162	C394

Subtotal de quinases: 112

Total de Cys:

172

TCDM_09327	V5BA65	CMGC	MAPK	Sem Cys mapeadas			
TCDM_05941	V5DDT1	CMGC	MAPK	Sem Cys mapeadas			
TCDM_06012	V5BCX5	CMGC	CK2	Sem Cys mapeadas			
TCDM_09017	V5D6X4	CMGC	CDK	Sem Cys mapeadas			
TCDM_11907	V5AIW9	CK1	CK1	Sem Cys mapeadas			
TCDM_08467	V5AS81	STE	STE11	Sem Cys mapeadas			
TCDM_07411	V5AUL7	Unique		Sem Cys mapeadas			
TCDM_06460	V5AWK7	CMGC	DYRK	Sem Cys mapeadas			
TCDM_05589	V5AYC9	Other	NEK	Sem Cys mapeadas			
TCDM_05407	V5AYY2	CMGC	DYRK	Sem Cys mapeadas			
TCDM_04734	V5B0H0	Other	PEK	Sem Cys mapeadas			
TCDM_04648	V5B0K7	STE	(na)	Sem Cys mapeadas			
TCDM_04350	V5B1E3	Other	(na)	Sem Cys mapeadas			
TCDM_03927	V5B2C1	CMGC	CLK	Sem Cys mapeadas			
TCDM_08566	V5B7L8	Other	TLK	Sem Cys mapeadas			
TCDM_01842	V5B7W5	STE	STE11	Sem Cys mapeadas			
TCDM_01637	V5B8S3	CMGC	DYRK	Sem Cys mapeadas			
TCDM_09547	V5B9R6	STE	(na)	Sem Cys mapeadas			
TCDM_01174	V5B9Y9	CMGC	MAPK	Sem Cys mapeadas			
TCDM_02063	V5B752	CMGC	SRPK	Sem Cys mapeadas			
TCDM_00566	V5BBX2	Other	CAMKK	Sem Cys mapeadas			

Código TriTryp	Código Uniprot	Classificação		Mapeamento Cisteínas			
		Grupo	Família	Sítio	Subsítio	Resíduo PDB 1ATP	Resíduo Dm28c 2014
TCDM_00324	V5BC71	Unique		Sem Cys mapeadas			
TCDM_08305	V5BCH2	AGC	(na)	Sem Cys mapeadas			
TCDM_07995	V5BDF5	STE	STE7	Sem Cys mapeadas			
TCDM_07421	V5BEN8	Unique		Sem Cys mapeadas			
TCDM_06238	V5BH34	CMGC	(na)	Sem Cys mapeadas			
TCDM_03919	V5BHT9	Unique		Sem Cys mapeadas			
TCDM_08004	V5BHW8	Other	NEK	Sem Cys mapeadas			
TCDM_07765	V5BIE1	Other	ULK	Sem Cys mapeadas			
TCDM_07031	V5BJT8	CMGC	CDK	Sem Cys mapeadas			
TCDM_04883	V5BK43	AGC	(na)	Sem Cys mapeadas			
TCDM_03118	V5BK52	STE	STE11	Sem Cys mapeadas			
TCDM_04752	V5BKJ3	STE	(na)	Sem Cys mapeadas			
TCDM_04353	V5BLF8	Unique		Sem Cys mapeadas			
TCDM_03992	V5BMH5	STE	(na)	Sem Cys mapeadas			
TCDM_02072	V5BMQ3	CK1	CK1	Sem Cys mapeadas			
TCDM_02945	V5BPV2	STE	STE11	Sem Cys mapeadas			
TCDM_01285	V5BPV5	CMGC	(na)	Sem Cys mapeadas			
TCDM_00973	V5BQS8	STE	(na)	Sem Cys mapeadas			
TCDM_00601	V5BRJ3	Other	CAMKK	Sem Cys mapeadas			
TCDM_02123	V5BSF9	STE	STE11	Sem Cys mapeadas			
TCDM_03427	V5BTL4	CMGC	CK2	Sem Cys mapeadas			
TCDM_01586	V5BTT8	STE	STE11	Sem Cys mapeadas			
TCDM_02794	V5BV23	CK1	TTBK	Sem Cys mapeadas			
TCDM_01189	V5BV49	CMGC	CDK	Sem Cys mapeadas			
TCDM_00829	V5BVV5	Unique		Sem Cys mapeadas			
TCDM_00586	V5BW58	CMGC	CDK	Sem Cys mapeadas			
TCDM_00504	V5C1S0	STE	STE11	Sem Cys mapeadas			
TCDM_00010	V5C2I6	Unique		Sem Cys mapeadas			
TCDM_06927	V5DBJ0	Unique		Sem Cys mapeadas			
TCDM_06777	V5DBX9	CMGC	MAPK	Sem Cys mapeadas			
TCDM_06605	V5DCD4	CMGC	RCK	Sem Cys mapeadas			
TCDM_05093	V5DFW9	Other	NEK	Sem Cys mapeadas			
TCDM_04327	V5DHT6	STE	STE11	Sem Cys mapeadas			
TCDM_03390	V5DKD4	Unique		Sem Cys mapeadas			
TCDM_00747	V5DSD2	Unique		Sem Cys mapeadas			
TCDM_00507	V5DTI8	CMGC	RCK	Sem Cys mapeadas			
TCDM_00351	V5DU02	AGC	(na)	Sem Cys mapeadas			
TCDM_00176	V5DUQ7	STE	(na)	Sem Cys mapeadas			

**Subtotal de
quinases: 59**

(na) - not assigned to a Family

APÊNDICE B - Tabela de classificação das 75 quinases de *T. cruzi* da cepa CL Brener El que não apresentaram cisteínas em nenhuma das 25 posições mapeadas.

Código TriTryp	Código Uniprot	Classificação		
		Grupo	Família	Nome
TcCLB.504167.30	Q4CSB9	CMGC	MAPK	KFR1
TcCLB.503513.10	Q4CSX2	CMGC	CK2	CK2A2
TcCLB.506519.60	Q4DFF4	CMGC	CLK	-
TcCLB.504181.40	Q4CRI4	CMGC	CDK	CRK1
TcCLB.510597.9	Q4CSS0	Other	TLK	-
TcCLB.509231.20	Q4D0S5	CMGC	RCK	-
TcCLB.504929.10	Q4CN82	CK1	CK1	CK1.2
TcCLB.509775.20	Q4DCQ6	Unique		
TcCLB.506543.10	Q4D7T4	STE	(na)	-
TcCLB.477079.10	Q4CLH7	CMGC	SRPK	-
TcCLB.506457.50	Q4DGA1	Other	NEK	NEK19
TcCLB.507213.20	Q4CZ67	Unique		
TcCLB.508179.60	Q4D3P9	Unique		
TcCLB.506825.180	Q4DZE5	STE	STE11	-
TcCLB.506777.70	Q4DC69	CMGC	CDK	CRK8
TcCLB.508181.4	Q4DRA8	AGC	RSK	-
TcCLB.508515.40	Q4DM73	AGC	(na)	-
TcCLB.510741.70	Q4DRL8	STE	STE11	-
TcCLB.508995.10	Q4D9R4	STE	(na)	-
TcCLB.506211.180	Q4DYK0	CMGC	MAPK	MPK9
TcCLB.480785.10	Q4CLM3	CMGC	DYRK	-
TcCLB.511393.60	Q4D3B9	STE	STE7	-
TcCLB.506945.110	Q4DX87	CK1	CK1	-
TcCLB.508951.4	Q4D3P7	Unique		
TcCLB.506211.210	Q4DYJ7	STE	STE11	CBPK1
TcCLB.506739.190	Q4DRQ3	CMGC	DYRK	-
TcCLB.507927.4	Q4D592	AGC	PKB	PKB
TcCLB.511715.20	Q4D5A5	Other	NEK	NEK21
TcCLB.506677.20	Q4CPK2	CMGC	CDK	CRK3
TcCLB.507809.14	Q4DST0	Unique		
TcCLB.508541.240	Q4DN94	CK1	CK1	CK1.2
TcCLB.504047.60	Q4D7T6	STE	(na)	-
TcCLB.509475.10	Q4CKS6	CMGC	MAPK	MPK3
TcCLB.511143.6	Q4CR33	Other	NEK	-
TcCLB.508541.230	Q4DN94	CK1	CK1	CK1.2
TcCLB.484949.9	Q4CLI4	AGC	RSK	-
TcCLB.503409.10	Q4CP45	Unique		

Código TriTryp	Código Uniprot	Classificação		
		Grupo	Família	Nome
TcCLB.506505.40	Q4D3I1	Other	CAMKK	-
TcCLB.506775.190	Q4DRG6	Other	CAMKK	-
TcCLB.511395.90	Q4D9P8	Other	NEK	NEK14
TcCLB.508541.225	Q4DN96	CK1	CK1	CK1.2
TcCLB.511075.60	Q4CZ32	STE	(na)	MKK4
TcCLB.508541.220	Q4DN97	CK1	CK1	CK1.1
TcCLB.504125.90	Q4DFR4	Unique		
TcCLB.510295.50	Q4CZ09	CMGC	MAPK	MPK6
TcCLB.427247.10	Q4CN51	CMGC	DYRK	-
TcCLB.510533.50	Q4DX09	Unique		
TcCLB.510431.140	Q4DXY0	CMGC	DYRK	DYRK2
TcCLB.504035.110	Q4DI80	Unique		
TcCLB.503955.110	Q4DH84	Unique		
TcCLB.510247.20	Q4CRN6	CK1	CK1	CK1.3
TcCLB.506343.10	Q4D334	STE	(na)	-
TcCLB.510231.10	Q4CTA4	Other	PEK	EIF2AK2
TcCLB.506633.10	Q4D8Q8	STE	STE7	-
TcCLB.509773.20	Q4CMU3	Unique		
TcCLB.509167.190	Q4DQP8	STE	STE11	-
TcCLB.503955.100	Q4DH85	Unique		
TcCLB.506945.220	Q4DX78	CMGC	SRPK	SRPK
TcCLB.511585.170	Q4E164	CMGC	(na)	-
TcCLB.510089.130	Q4DQZ6	CMGC	CLK	-
TcCLB.511395.60	Q4D9Q1	Other	NEK	NEK14
TcCLB.510761.60	Q4DG40	CMGC	CK2	CK2A1
TcCLB.510291.40	Q4CTM2	AGC	NDR	-
TcCLB.504089.54	Q4DE56	Other	ULK	-
TcCLB.510609.70	Q4DJM4	CMGC	CDK	CRK2
TcCLB.506887.100	Q4D338	Unique		
TcCLB.510759.40	Q4DS91	Unique		
TcCLB.507809.170	Q4DSR4	Unique		
TcCLB.510089.80	Q4DR01	Unique		
TcCLB.503903.79	Q4CZR3	Unique		
TcCLB.506467.40	Q4D0G5	Unique		
TcCLB.506165.16	Q4CPF2	Unique		
TcCLB.506811.180	Q4DUZ9	STE	STE11	-
TcCLB.504057.170	Q4DSD6	CK1	TTBK	-
TcCLB.504421.10	Q4CR24	Unique		