



**LETÍCIA SANTOS GARCIA**

**ESTUDOS DE POTENCIAIS INIBIDORES DA  
ENZIMA TIROSINA QUINASE DE BRUTON**

**LAVRAS – MG**

**2014**

**LETÍCIA SANTOS GARCIA**

**ESTUDOS DE POTENCIAIS INIBIDORES DA ENZIMA TIROSINA  
QUINASE DE BRUTON**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Agroquímica, área de concentração em Agroquímica, para a obtenção do título de Mestre.

Orientadora

Dra. Elaine Fontes Ferreira da Cunha

**LAVRAS - MG**

**2014**

**Ficha Catalográfica Elaborada pela Coordenadoria de Produtos e  
Serviços da Biblioteca Universitária da UFLA**

Garcia, Leticia Santos.

Estudos de potenciais inibidores da enzima Tirosina Quinase de  
Bruton / Leticia Santos Garcia. – Lavras : UFLA, 2014.

100 p. : il.

Dissertação (mestrado) – Universidade Federal de Lavras, 2014.

Orientador: Elaine Fontes Ferreira da Cunha.

Bibliografia.

1. Linfócito B. 2. Tirosina Quinase de Bruton. 3. Ancoramento  
molecular. 4. QSAR-4D DR. I. Universidade Federal de Lavras. II.  
Título.

CDD – 574.1925

**LETÍCIA SANTOS GARCIA**

**ESTUDOS DE POTENCIAIS INIBIDORES DA ENZIMA TIROSINA  
QUINASE DE BRUTON**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Agroquímica, área de concentração em Agroquímica, para a obtenção do título de Mestre.

APROVADA em 27 de fevereiro de 2014.

Dra. Daiana Teixeira Mancini      UFLA

Dra. Melissa Soares Caetano      UFOP

Dr. Teodorico de Castro Ramalho      UFLA

Dra. Elaine Fontes Ferreira da Cunha  
Orientadora

**LAVRAS - MG**

**2014**

## AGRADECIMENTOS

A DEUS, senhor de todas as coisas, que me permitiu a conclusão desta etapa de minha vida.

Aos meus pais, Déia e Gaspar, por serem a luz da minha estrada e meus pilares de segurança em tudo que faço. Por estarem sempre comigo, mesmo à distância, por não me deixarem desanimar durante a caminhada, e pela certeza de seu carinho incondicional.

À minha amada e única irmã Michelly, exemplo de amizade, companheirismo, confiança, amor, carinho e compreensão.

Ao meu namorado Juliano, pelo amor, paciência, carinho, por me apoiar sempre e por fazer minha vida mais feliz.

A toda minha família, que sempre torceu para que essa vitória fosse concretizada.

À minha orientadora Elaine, pela oportunidade, por seu constante apoio, confiança, ensinamentos, paciência e pela amizade.

Aos membros da banca, Daiana, Melissa e Teodorico, pela disposição e pelas contribuições que complementarão o trabalho.

Aos companheiros do Laboratório de Química Computacional pelos ensinamentos, incentivo e amizade.

Ao CNPq, UFLA e DQI, pelo suporte e oportunidade de realizar a pesquisa.

## RESUMO

A Tirosina Quinase de Bruton (Btk) é uma enzima importante para o desenvolvimento e diferenciação dos linfócitos B. Além disso, a expressão da Btk é considerada como essencial para a proliferação e sobrevivência dessas células. A inibição da Btk tem se tornado um alvo atrativo para o tratamento de doenças autoimunes, leucemias e linfomas de células B. Objetivando-se propor novos candidatos a fármacos inibidores da Btk, técnicas computacionais foram aplicadas neste estudo. Foi realizado um estudo de ancoramento molecular a fim de avaliar o modo de interação de 96 inibidores nicotinamídicos da Btk (DELUCCA et al., 2010). Os valores dos termos de energia para as melhores conformações ancoradas mostrou um coeficiente de determinação de 0,80. A metodologia de QSAR-4D DR também foi utilizada. As conformações obtidas por simulação da dinâmica molecular foram sobrepostas, de acordo com os três alinhamentos testados, em uma caixa virtual tridimensional composta por células de 2 e 5Å. Os modelos foram gerados pela técnica combinada de algoritmos genéticos (GA, do inglês *Genetic Algorithm*) e mínimos quadrados parciais (PLS, do inglês *Partial Least Square*). O melhor modelo gerou valores de validação cruzada ajustado ( $q^2_{\text{ajustado}}$ ) e coeficiente de determinação ( $R^2$ ) de 0,67 e 0,74, respectivamente. A análise do QSAR-4D DR, além de corroborar com os resultados do *docking*, sugeriu uma boa correlação com os resultados experimentais e, ainda, forneceu informações mais detalhadas do que o ancoramento molecular. A boa potência predita para o composto **M2** proposto indica que este composto é um potencial candidato a inibidor da Btk.

Palavras-chave: Linfócito B. Tirosina Quinase de Bruton. Ancoramento molecular. QSAR-4D DR

## ABSTRACT

Bruton's Tyrosine Kinase (Btk) is an important enzyme in B-lymphocyte development and differentiation. Furthermore, Btk expression is considered essential for the proliferation and survival of these cells. Btk inhibition has become an attractive target for treating autoimmune diseases, and B-cell leukemia and lymphomas. With the objective of proposing new candidates for Btk inhibitor pharmaceuticals, we applied computational techniques in this study. A study using molecular docking was performed in order to evaluate the mode of interaction of 96 nicotinamidic Btk inhibitors (DELUCCA et al., 2010). The energy term values of the best docked conformations showed a coefficient of correlation ( $R^2$ ) equal to 0.80. We also used the receptor-dependent four-dimensional quantitative structure-activity relationship (RD 4D-QSAR) methodology. The conformations obtained by molecular dynamic simulation were overlapped in a virtual three dimensional box comprised of 2 and 5 Å cells, according to the three trial alignments. The models were generated by a combined genetic algorithm (GA) and partial least squares (PLS) regression technique. The best model generated adjusted cross-validate value ( $q^2_{\text{adjusted}}$ ) and correlation coefficient value ( $R^2$ ) of 0.67 and 0.74, respectively. The RD 4D-QSAR analysis, in addition to corroborating with the docking results, suggested a good correlation with the experimental results and provided more detailed information than the molecular docking. The good potency predicted for the proposed **M2** compound indicates this compound as a potential candidate of Btk inhibitor.

Keywords: B-lymphocyte. Bruton's Tyrosine Kinase. Molecular docking. RD 4D-QSAR.

## LISTA DE FIGURA

Figura 1	Fosforilação do resíduo de aminoácido tirosina por uma enzima TK .....19
Figura 2	Modelo esquemático da Btk demonstrando seus diferentes domínios: <i>pleckstrin homology</i> (PH); <i>Tec homology</i> (TH), que é internamente dividido em <i>Btk homology</i> e <i>polyproline region</i> (PRR); <i>Src homology 3</i> (SH3); <i>Src homology 2</i> (SH2); <i>tyrosine kinase</i> , SH1 (TK) .....22
Figura 3	Modelo esquemático da ativação da Btk. Após o acoplamento do receptor de células B (BCR) a Btk é fosforilada pela Lyn ou Syk no resíduo Tyr551, sendo então possível a autofosforilação da Btk no resíduo Tyr223. Assim, a Btk torna-se ativa e fosforila a fosfolipase C $\gamma$ levando à mobilização de cálcio .....23
Figura 4	Mudança conformacional sofrida pela Btk ao interagir com o inibidor CGI746 (amarelo) resultando no sequestro do resíduo Tyr551. Em verde está a Btk na conformação ligada ao inibidor, e em cinza, em sua forma apo. O resíduo Tyr551 em ambas as conformações é mostrado como esferas. Os resíduos de aminoácidos que constituem o bolsão H3 estão demonstrados à direita.....26
Figura 5	Estrutura terciária da Btk obtida do PDB (código 3OCS). As $\alpha$ -hélices estão representadas em vermelho, as folhas- $\beta$ em azul e o sítio ativo em verde .....41
Figura 6	Cavidade da enzima BTK ilustrada em verde.....51



Figura 7	Estrutura do ligante 4-terc-butil-N-[2-metil-3-(4-metil-6-{{4-(morfolin-4-ilcarbonil)fenil] amino} -5-oxi-4,5-dihidropirazin-2-il)fenil]benzamida (CGI1746) que foi obtido do PDB (código 3OCS) .....52
Figura 8	A figura ilustra a superfície eletrostática da Btk e os inibidores <b>14</b> (em verde) e <b>59</b> (em amarelo), na conformação de ancoramento.....55
Figura 9	A figura ilustra o bolsão H3 na Btk e os inibidores: A) <b>14</b> (em verde) e <b>65</b> (em amarelo); B) <b>14</b> (em verde) e <b>75</b> (em amarelo); C) <b>14</b> (em verde) e <b>59</b> (em amarelo), ancorados no sítio ativo da enzima.....56
Figura 10	Ilustração da conformação de ancoramento da molécula <b>59</b> . O átomo de flúor no anel de nicotinamida e o átomo de nitrogênio da piperidina estão marcados em verde .....57
Figura 11	Imagem ilustrativa do corte realizado na enzima Btk (código 3OCS) ao redor do ligante CGI1746, a fim de seleccionar os resíduos de aminoácidos a serem utilizados na metodologia QSAR-4D DR para os 96 compostos .....59
Figura 12	Modelo de QSAR-4D DR representativo do alinhamento <i>1a</i> , mostrando o composto <b>59</b> .....69
Figura 13	Representação da molécula <b>14</b> e do GCOD-1 (esfera branca) obtido do Modelo1 do alinhamento <i>1a</i> .....70
Figura 14	A) Representação da molécula <b>53</b> e do GCOD-2 (esfera branca) obtido do Modelo1 do alinhamento <i>1a</i> ; B) Representação da molécula <b>59</b> e do GCOD-3 (esfera branca) obtido do Modelo1 do alinhamento <i>1a</i> .....73

Figura 15	A) Representação da molécula <b>53</b> e do GCOD-4 (esfera branca) obtido do Modelo1 do alinhamento <i>1a</i> ; B) Representação da molécula <b>49</b> e do GCOD-5 (esfera preta) obtido do Modelo1 do alinhamento <i>1a</i> .....	75
Figura 16	Representação da molécula <b>59</b> e do GCOD-6 (esfera branca) obtido do Modelo1 do alinhamento <i>1a</i> .....	77
Figura 17	A) Representação da molécula <b>55</b> e do GCOD-7 (esfera branca) obtido do Modelo1 do alinhamento <i>1a</i> . B) Representação da molécula <b>61</b> e do GCOD-8 (esfera preta) obtido do Modelo1 do alinhamento <i>1a</i> .....	79
Figura 18	Representação da molécula <b>55</b> e do GCOD-9 (esfera branca) obtido do Modelo1 do alinhamento <i>1a</i> .....	80
Figura 19	Ilustração da molécula <b>59</b> em sua conformação de ancoramento. As ligações de hidrogênio são demonstradas em azul e a distância do átomo de flúor (em cinza) e do resíduo de Lys430 é demonstrada em verde.....	83
Figura 20	Ilustração da molécula <b>59</b> em sua conformação de ancoramento e das representações dos potenciais eletrostático positivo (em azul) e negativo (em vermelho). Os átomos de flúor (cinza) estão indicados em verde.....	84
Figura 21	Ilustração das moléculas propostas: <b>M1</b> e <b>M2</b> .....	85

## LISTA DE GRÁFICOS

Gráfico 1	Valores de $pIC_{50}$ experimentais versus preditos do conjunto de treinamento usando o melhor modelo de QSAR-4D-DR do alinhamento <i>1a</i> .....	63
-----------	--	----

## LISTA DE QUADROS

Quadro 1	Estrutura química e os respectivos valores de $pIC_{50}$ (M) dos 96 análogos nicotinamídicos. Em <b>negrito</b> estão representadas as moléculas selecionadas para compor o grupo treinamento, e em <i>itálico</i> , o grupo teste da metodologia QSAR-4D .....32
----------	---

## LISTA DE TABELAS

Tabela 1	Alinhamentos testados e os átomos de C <sub>α</sub> da proteína selecionados para o alinhamento dos complexos. O composto <b>59</b> está representado em cinza escuro.....	45
Tabela 2	Definição dos elementos de interação farmacofórica, IPEs.....	46
Tabela 3	Valores de pIC <sub>50</sub> experimentais e de energia de impedimento espacial (kcal.mol <sup>-1</sup> ) obtidos do ancoramento com a enzima Btk para as moléculas <b>14, 59, 65, e 75</b> .....	55
Tabela 4	Valores de pIC <sub>50</sub> experimentais e de energia de interação intramolecular (kcal.mol <sup>-1</sup> ) obtidos do ancoramento com a enzima Btk para as moléculas <b>12, 14, 36, 52, 59, 60, e 80</b> .....	57
Tabela 5	Valores de pIC <sub>50</sub> experimentais e de energia de interação eletrostática com raio > 4,5 Å entre a proteína e o ligante (kcal.mol <sup>-1</sup> ) obtidos do ancoramento com a enzima Btk para as moléculas <b>24, 58, 59, 60, 61, e 70</b> .....	58
Tabela 6	Alinhamentos testados no QSAR-4D DR para a Btk .....	60
Tabela 7	Valores estatísticos para os diferentes alinhamentos analisados ....	61
Tabela 8	Matriz de correlação cruzada dos GCODs do melhor modelo do alinhamento <i>1a</i> .....	62
Tabela 9	Valores de pIC <sub>50</sub> experimentais e preditos e valores residuais (pIC <sub>50</sub> predito - pIC <sub>50</sub> experimental) para os compostos do conjunto de treinamento ( <b>negrito</b> ) e o do conjunto de teste ( <i>itálico</i> ).....	64
Tabela 10	Valores de ocupação percentual do GCOD-1 do Modelo 1 do alinhamento <i>1a</i> do QSAR-4D DR para os compostos <b>10, 12, 14, 43 e 96</b> .....	71

Tabela 11	Valores de ocupação percentual do GCOD2 do Modelo 1 do alinhamento <i>1a</i> do QSAR-4D DR para os compostos <b>10, 14, 53, 82 e 85</b> .....	72
Tabela 12	Valores de ocupação percentual do GCOD-3 do Modelo 1 do alinhamento <i>1a</i> do QSAR-4D DR para os compostos <b>10, 14, 38,45 e 59</b> .....	74
Tabela 13	Valores de ocupação percentual dos GCODs 4 e 5 do Modelo 1 do alinhamento <i>1a</i> do QSAR-4D DR para os compostos <b>30, 35, 49, 50 e 53</b> .....	76
Tabela 14	Valores de ocupação percentual do GCOD-6 do Modelo 1 do alinhamento <i>1a</i> do QSAR-4D DR para os compostos <b>10, 14, 17, 59 e 92</b> .....	77
Tabela 15	Valores de ocupação percentual dos GCODs 7 e 8 do Modelo 1 do alinhamento <i>1a</i> do QSAR-4D DR para os compostos <b>17, 40, 42, 55 e 61</b> .....	79
Tabela 16	Valores de ocupação percentual do GCOD-9 do Modelo 1 do alinhamento <i>1a</i> do QSAR-4D DR para os compostos <b>28, 36, 54, 55, 62 e 63</b> .....	81
Tabela 17	Valores de potência inibitória preditos ( $pIC_{50predito}$ ) (M) para os inibidores <b>14, 59, M1e M2</b> .....	86
Tabela 18	Valores de $E_{Est}$ , $E_{Intra}$ , $E_{LigH}$ , $E_{Elet}$ e $E_{EletLong}$ ( $kcal.mol^{-1}$ ) e valores de $IC_{50}$ predito para os inibidores <b>14, 59, M1e M2</b> .....	86

## LISTA DE ABREVIATURAS

Ala	Alanina
Asn	Asparagina
Asp	Aspartato
Cys	Cisteína
Gln	Glutamina
His	Histidina
Leu	Leucina
Lys	Lisina
Met	Metionina
Phe	Fenilalanina
Ser	Serina
Thr	Treonina
Tyr	Tirosina
Val	Valina

## LISTA DE SIGLAS

ATP	Trifosfato de adenosina
BCR	Receptor de célula B
Btk	tirosina quinase de Bruton
CEP	<i>Conformational Ensemble Profile</i>
GA	<i>Genetic algorithm</i>
GCOD	<i>Grid Cell Occupancy Descriptor</i>
GFA	<i>Genetic Function Approximation</i>
IPE	<i>Interaction Pharmacophore Elements</i>
NF-κB	Fator nuclear kappa B
PDB	<i>Protein Data Bank</i> (Banco de Dados de Proteínas)
PH	<i>Pleckstrin homology</i>
PLS	<i>Partial least squares</i>
PPR	<i>Poly Proline Region</i>
QSAR	Relação quantitativa estrutura-atividade
QSAR-4D DR	Quarta dimensão da relação quantitativa estrutura-atividade dependente do receptor
QSAR-4D IR	Quarta dimensão da relação quantitativa estrutura-atividade independente do receptor
SDM	Simulação da dinâmica molecular
SH	<i>Src homology</i>
TH	<i>Tec homology</i>
UFLA	Universidade Federal de Lavras
XLA	Agamaglobulinemia ligada ao X
PLC-γ	Fosfolipase C gamma
TK	Tirosina Quinase



## LISTA DE SÍMBOLOS

Å	Angstrom
IC <sub>50</sub>	concentração capaz de inibir 50% da atividade enzimática
K	Kelvin
µM	Micromolar
M	Molar
kcal mol <sup>-1</sup>	Quilocaloria por mol
pIC <sub>50</sub>	Log de IC <sub>50</sub>

## SUMÁRIO

1	INTRODUÇÃO .....	17
2	REFERENCIAL TEÓRICO .....	19
2.1	Tirosina Quinase.....	19
2.2	Linfócitos B e a Tirosina Quinase de Bruton .....	20
2.3	A bioquímica da Btk e sua ativação .....	21
2.4	Inibidores da Btk.....	24
2.5	Ancoramento molecular ou <i>Docking</i> .....	26
2.6	Relação quantitativa estrutura- atividade.....	28
3	OBJETIVOS .....	31
4	METODOLOGIA.....	32
4.1	Dados Biológicos .....	32
4.2	Cálculo de ancoramento molecular ( <i>docking</i> ).....	39
4.3	QSAR-4D .....	41
4.3.1	Grupos teste e treinamento.....	42
4.3.2	Simulação de dinâmica molecular.....	43
4.3.3	Definição do alinhamento .....	44
4.3.4	Definição do tamanho das células.....	45
4.3.5	Definição dos elementos de interação farmacofórica.....	46
4.3.6	Pré-tratamento dos descritores (GCODs).....	46
4.3.7	Obtenção dos modelos de QSAR-4D .....	47
4.3.8	Validação interna dos modelos .....	48
4.3.9	Validação externa dos modelos .....	49
4.3.10	Identificação dos compostos <i>outliers</i> .....	49
4.3.11	Seleção da conformação “representativa” .....	50
5	RESULTADOS E DISCUSSÃO.....	51
5.1	Ancoramento molecular .....	51
5.2	QSAR-4D .....	59
5.3	Seleção do melhor alinhamento de QSAR-4D.....	59
5.4	Identificação dos <i>outliers</i> .....	63
5.4.1	Análise do alinhamento <i>1a</i> .....	67
5.5	Planejamento de potenciais inibidores da Btk.....	81
6	CONCLUSÕES.....	88
	REFERÊNCIAS.....	89
	APÊNDICE .....	96

## 1 INTRODUÇÃO

A proteína quinase de Bruton (Btk, do inglês Bruton's tyrosine kinase) é uma tirosina quinase citoplasmática importante para o desenvolvimento, diferenciação e sinalização dos linfócitos B. Sua expressão é fundamental para a proliferação e sobrevivência dessas células. Assim, linfócitos B deficientes em Btk não atingem a maturação e são incapazes de produzir anticorpos. Em humanos, mutações no gene da Btk levam à agamaglobulinemia, ligada ao cromossomo X (MOHAMED et al., 2009).

Muitos dos processos prejudiciais à saúde humana que estão associados à ativação desregulada dos linfócitos B e à formação de autoanticorpos podem ser interrompidos ao se inativar a Btk. Sua inibição é utilizada no tratamento de várias doenças alérgicas, autoimunes e inflamatórias, além da terapêutica do linfoma de células B e da leucemia (DELUCCA et al., 2010). Desta forma, o estudo de inibidores capazes de regular a atividade da Btk é importante visando a possibilidade de sua aplicação no reestabelecimento da homeostasia do organismo.

Métodos de química computacional desempenham um papel importante no processo de descoberta e desenvolvimento de novos fármacos. Uma ferramenta amplamente utilizada é o estudo da correlação entre a estrutura e a atividade de fármacos (LILL, 2007). Neste método, considera-se que existe uma relação entre as propriedades de uma molécula e sua estrutura química, e então, buscam-se estabelecer relações matemáticas simples para descrever e prever a atividade de um conjunto de inibidores semelhantes.

Outra metodologia relevante para o desenvolvimento de medicamentos é o ancoramento molecular, que envolve a predição da conformação do ligante e sua orientação dentro do sítio de ligação da proteína.

Assim, estudos de ancoramento molecular e da correlação entre a estrutura e a atividade de inibidores da Btk permitem avaliar as interações de inibidores no sítio ativo, propor modelos que descrevam a atividade de moléculas congêneres e, então, contribuir no desenho e desenvolvimento de novos inibidores específicos para esta enzima, proporcionando o benefício terapêutico no tratamento de diversas doenças relacionadas com a atuação desregulada dos linfócitos B.

Neste trabalho, procura-se entender o modo de interação entre a enzima Btk e 96 compostos nicotinamídicos, por meio da técnica de ancoramento molecular; e realizar estudos de QSAR-4D dependente do receptor, a fim de selecionar características importantes que um potente inibidor deva apresentar. Pretende-se ainda propor candidatos a fármacos inibidores da Btk.

## 2 REFERENCIAL TEÓRICO

### 2.1 Tirosina Quinase

As tirosinas quinases (TKs, do inglês Tyrosine Kinases) são enzimas que catalisam a transferência do grupo  $\gamma$ -fosfato do ATP para resíduos de tirosina em proteínas (Figura 1) (LA MOTTA et al., 2009). A fosforilação de proteínas conduz à ativação de vias de transdução de sinal que desempenham um papel importante em um grande número de processos biológicos (CHEETHAM, 2004; KONDAPALLI; SOLTANI; LACOUTURE, 2005).

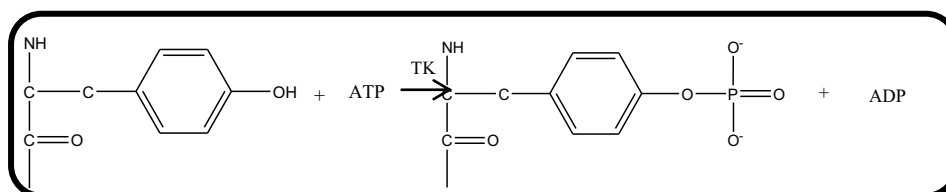


Figura 1 Fosforilação do resíduo de aminoácido tirosina por uma enzima TK

As proteínas TKs são divididas em duas principais classes: TKs transmembrânicas, que possuem um domínio de ligação extracelular e um domínio catalítico intracelular; e as TKs não receptoras, que não são transmembrânicas, sendo encontradas no citosol, no núcleo, e na superfície interna da membrana plasmática (KRAUSE; ETEN, 2005).

Como estas enzimas desempenham funções relevantes no controle da sinalização celular, uma TK com atividade alterada, leva a vias de sinalização alteradas (LA MOTTA et al., 2009), dando origem ao desenvolvimento de várias doenças, como câncer, doenças inflamatórias, neurológicas, cardiovasculares, hormonais e alérgicas. Deste modo, pesquisas para identificar inibidores de TKs que sejam eficazes como agentes terapêuticos contra estas doenças são de grande importância (YILMAZ et al., 2011).

## 2.2 Linfócitos B e a Tirosina Quinase de Bruton

A resposta imune humoral é mediada pelas moléculas presentes no sangue e nas secreções das mucosas, chamadas de anticorpos (ou imunoglobulinas), que são produzidas pelos linfócitos B (ou células B). Os anticorpos reconhecem antígenos microbianos, neutralizam sua infecciosidade e os preparam para serem eliminados por diversos mecanismos efetores (ABBAS; LICHTMAN; PILLAI, 2008).

A enzima Btk é um membro das proteínas tirosina quinases citoplasmáticas (MAO; ZHOU; UCKUN, 2001) e é uma enzima crucial para a diferenciação, sinalização e desenvolvimento dos linfócitos B. A expressão da Btk é necessária para a proliferação e sobrevivência dessas células. Desta forma, linfócitos B deficientes em Btk não se tornam maduros e são incapazes de gerar imunoglobulinas (MOHAMED et al., 2009; THOMAS et al., 1993; TSUKADA et al., 1993; VETRIE et al., 1993).

Inúmeras mutações no gene da Btk, localizado no cromossomo X, são relatadas. Tais mutações levam à deficiências na maturação e na função de linfócitos B maduros, resultando na agamaglobulinemia ligada ao X (XLA, do inglês X-Linked Agammaglobulinemia) (KUGLSTATTER et al., 2011). Clinicamente, os pacientes apresentam quadros infecciosos que se iniciam já no primeiro ano de vida, logo após a queda dos níveis de imunoglobulinas maternas, podendo ocorrer infecções respiratórias, cutâneas e do trato gastrointestinal. Laboratorialmente, há níveis séricos bastante reduzidos de todos os anticorpos, e níveis de linfócitos B circulantes inferiores a 2% (JACOB, 2005).

A inibição da atividade da Btk, por meio de pequenas moléculas indicou que esta enzima é um potencial alvo terapêutico para o tratamento de doenças

inflamatórias, autoimunes, leucemias e linfomas da linhagem de células B (KUGLSTATTER et al., 2011).

### 2.3 A bioquímica da Btk e sua ativação

A enzima Btk contém vários domínios, como ilustrado na Figura 2: um domínio homólogo à pleckstrina aminoterminal (PH, do inglês pleckstrin homology), seguido por um domínio de homologia Tec (TH), um domínio de homologia Src (SH) 3, um SH2, e um domínio catalítico carboxiterminal da TK (SH1) (KUGLSTATTER et al., 2011).

O domínio PH, comumente encontrado em proteínas sinalizadoras eucarióticas, possui aproximadamente 100 aminoácidos, e ocorre em muitas proteínas envolvidas na sinalização intracelular e na constituição do citoesqueleto. O domínio TH contém uma sequência específica de ligação ao zinco, com resíduos de cisteína e histidina conservadas, e geralmente está localizado na porção carboxiterminal do domínio PH. A estrutura cristalográfica dessa região demonstra que os segmentos de uma sequência específica denominada Btk interagem com os domínios PH e esta interação foi denominada *Tec homology* (TH) – (Btk- PH). Os oito resíduos localizados na porção extrema do domínio Btk-PH possuem um núcleo hidrofóbico, que é importante para a estabilidade da proteína. Os 27 resíduos da porção aminoterminal de TH são muito conservados e são seguidos de uma região rica em prolina (*Poly Proline Region*, PPR) (OLIVEIRA, 2008).

O SH3 contém aproximadamente 50 resíduos de aminoácidos e liga-se a regiões ricas em prolina e em aminoácidos hidrofóbicos das proteínas alvo. O domínio SH2 é constituído por cerca de 100 aminoácidos e possui função regulatória da cascata de sinalização intracelular. O domínio quinase está

presente em proteínas que compartilham de um núcleo catalítico conservado com as proteínas serina/treonina e TK (OLIVEIRA, 2008).

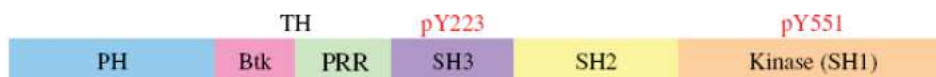


Figura 2 Modelo esquemático da Btk demonstrando seus diferentes domínios: *pleckstrin homology* (PH); *Tec homology* (TH), que é internamente dividido em *Btk homology* e *polyproline region* (PRR); *Src homology* 3 (SH3); *Src homology* 2 (SH2); *tyrosine kinase*, SH1 (TK)

Fonte: Oliveira (2008)

O mecanismo molecular fundamental para a ativação da Btk não é totalmente compreendido. Entretanto, acredita-se que modificações pós traducionais, bem como a localização subcelular desta enzima possa desempenhar uma função crítica. No estado estacionário, a Btk é predominantemente citosólica. Após o acoplamento do receptor de células B (BCR), ela se transloca para a membrana onde é fosforilada pela enzima Syk ou pela Lyn, da família das TKs Src (MOHAMED et al., 2009). Após ser ativada pela transfosforilação na posição Tyr551 do loop de ativação no domínio quinase, a Btk é estimulada a autofosforilar o resíduo Tyr223 no domínio SH3 (KUGLSTATTER et al., 2011; MAO; ZHOU; UCKUN, 2001).

Posteriormente, a Btk ativada fosforila a fosfolipase C $\gamma$  (PLC- $\gamma$ ), o que leva à hidrólise de fosfatidilinositol 4, 5-bifosfato para a produção de inositol 1,4,5-trifosfato e diacilglicerol, causando a mobilização de cálcio e ativando a proteína quinase C, respectivamente (Figura 3) (GUO; WAHL; WITTE, 2006). Desta forma, a Btk regula vários eventos importantes, como o transporte de íons cálcio e a ativação de NF-kB em resposta à ativação do receptor de célula B (BCR) (PAN et al., 2007). Neste contexto, a Btk promove a produção de



anticorpos, diferenciação de células B e a progressão do ciclo celular (FIUME et al., 2009; MAO; ZHOU; UCKUN, 2001).

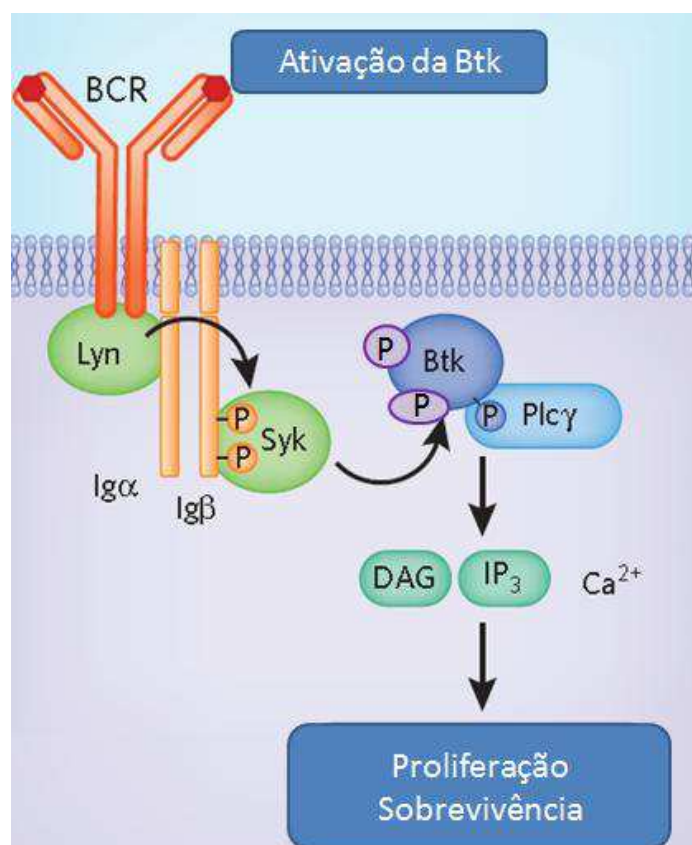


Figura 3 Modelo esquemático da ativação da Btk. Após o acoplamento do receptor de células B (BCR) a Btk é fosforilada pela Lyn ou Syk no resíduo Tyr551, sendo então possível a autofosforilação da Btk no resíduo Tyr223. Assim, a Btk torna-se ativa e fosforila a fosfolipase C $\gamma$  levando à mobilização de cálcio

Fonte adaptado de Hendriks (2011)

Baseado em proteínas terapêuticas, como Rituxan, desenvolvidas para destruir as células B, no tratamento de várias doenças autoimunes e inflamatórias, a inibição da atividade da Btk será útil para inibir os linfócitos B e a atividade patogênica mediada por estas células, podendo estes inibidores ser utilizados no tratamento de doenças alérgicas, inflamatórias, autoimunes, leucemias e linfoma de células B (DELUCCA et al., 2010).

## 2.4 Inibidores da Btk

O Ibrutinib ou PCI-32765 (1-[ (3R)-3-[4-amino-3-(4-fenoxifenil) pirazol [3,4-d]pirimidin-1-il]piperidin-1-il]prop-2-en-1-ona) é um potente inibidor da Btk (usado no tratamento do câncer) e se liga covalentemente ao resíduo de aminoácido Cys481, que se encontra no domínio de ligação do ATP na Btk, com  $IC_{50}$  de 0,5 nM, impedindo sua autofosforilação no resíduo Tyr223. Entretanto, o resíduo de cisteína também está presente em outras quinases com homologia à Btk as quais podem ser irreversivelmente inibidas pelo Ibrutinib (BROWN, 2013; HONIGBERG et al., 2010; VARGAS et al., 2013).

Dasatinib ou BMS-354825 [N-(2-cloro-6-metilfenil)-2-(6-(4-(2-hidroxi-2-hidroxi-1-il) piperazin-1-il)-2-metilpirimidin-4-ilamino)tiazol-5-carboxamida] foi aprovado pelo FDA para o tratamento de pacientes com leucemia mielóide crônica, quando ocorre resistência ao Imatinib. Originalmente utilizado como alvo para BCR/Abl, é um potente inibidor da Btk ( $IC_{50} = 5$  nM), todavia também inibe outras quinases, como os membros da família Src. Em adição, estudos relatam que a mutação no resíduo Thr474-Ile na Btk confere resistência ao Dasatinibe (HANTSCHHEL et al., 2007; MARCOTTE et al., 2010).

CGI1746 (4-terc-butil-N-[2-metil-3-(4-metil-6-{[4-(morfolin-4-ilcarbonil) fenil] amino} -5-oxi-4,5-dihidropirazin-2-il) fenil] benzamida), desenvolvido por Di Paolo et al. (2011) se liga à forma inativa da Btk,

impedindo as etapas de auto- e transfosforilação, necessárias para a ativação da enzima. Tem sido demonstrado que este inibidor regula a artrite reumatoide. É um potente inibidor competitivo da Btk com  $IC_{50}$  de 1,9 nM e tem se mostrado ser um inibidor específico para a Btk, com seletividade de aproximadamente 1000 vezes sobre outras famílias de quinase como Src e Tec. A ligação do CGI1746 induz a uma mudança conformacional na Btk, com a posição do grupo hidroxila da Tyr551, diferindo aproximadamente 18 Å na forma ligada ao inibidor CGI1746 versus sua estrutura apo (não ligada), resultando na formação de um grande bolsão, denominado “H3 pocket”, o qual é ocupado pela porção t-butilfenil do inibidor. O bolsão é formado pelos resíduos Tyr551, Leu542, Val546, Phe413, Asn526, Asp521, Asp539, His519 e Ser543. A Figura 4, ilustra o inibidor CGI1746 ocupando o “bolsão de especificidade” da Btk (DI PAOLO et al., 2011; LOU et al., 2012).

O Desatinib liga-se à Btk com igual afinidade tanto na forma inativa quanto na forma ativada. Já o CGI1746 liga-se à forma inativa da Btk com afinidade 32 vezes maior do que em sua forma ativada. Provavelmente, essa seletividade do CGI1746 seja devido às interações favoráveis com o bolsão H3, no estado inativo da enzima. Essas interações não são possíveis no estado ativo da Btk já que o resíduo de Tyr551 fosforilado previne a formação do bolsão H3 e, assim, o resíduo de Tyr551 não é sequestrado (DI PAOLO et al., 2011).

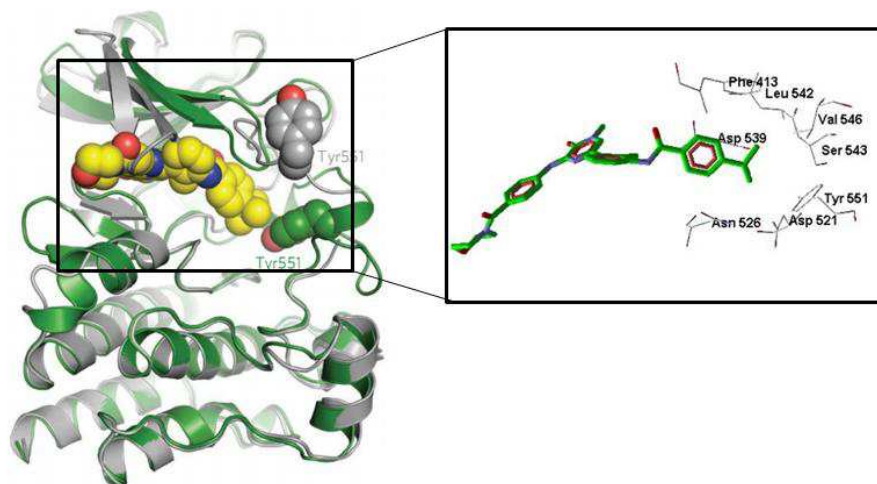


Figura 4 Mudança conformacional sofrida pela Btk ao interagir com o inibidor CGI746 (amarelo) resultando no sequestro do resíduo Tyr551. Em verde está a Btk na conformação ligada ao inibidor, e em cinza, em sua forma apo. O resíduo Tyr551 em ambas as conformações é mostrado como esferas. Os resíduos de aminoácidos que constituem o bolsão H3 estão demonstrados à direita.

Fonte adaptado de Di Paolo et al. (2011).

## 2.5 Ancoramento molecular ou *Docking*

Saber o posicionamento tridimensional da interação de um ligante com a proteína alvo é importante para se entender o mecanismo de interação de um fármaco. A predição da geometria e da energia de ligação é de grande interesse para o planejamento de novos fármacos (HILLISH et al., 2004; IWATA et al., 2000). A simulação computacional por ancoramento molecular (*docking*) é uma das técnicas de investigação das interações moleculares entre uma proteína e um ligante, quando a estrutura tridimensional da proteína já foi elucidada (SILVEIRA, 2003). Esta abordagem *in silico* permite uma identificação mais

rápida e mais barata de promissores candidatos a medicamentos (THOMSEN; CHRISTENSEN, 2006).

O propósito da descoberta de novas drogas é obter moléculas que se ligam mais fortemente à proteína alvo do que o substrato natural. Ao se fazer isso, a reação bioquímica que a molécula alvo catalisa pode ser alterada ou prevenida (THOMSEN; CHRISTENSEN, 2006).

O ancoramento molecular utiliza a estrutura do receptor como molde para o desenvolvimento de novos ligantes, estimando-se a afinidade de ligação entre ligante e receptor. Durante o ancoramento, uma função baseada nos valores de energia é usada para identificar a conformação do ligante energeticamente mais favorável, quando este está ancorado à proteína. Os valores mais baixos de energia representam melhor as interações proteína-ligante quando comparadas aos valores de energia mais elevados. Esse tipo de simulação encontra a energia mais estável do complexo proteína-ligante (THOMSEN; CHRISTENSEN, 2006).

Os valores da função de desempenho do ancoramento (*Docking Scoring Function*),  $E_{score}$ , usados neste estudo são definidos pela Eq. 1:

$$E_{score} = E_{inter} + E_{intra} \quad (1)$$

onde,  $E_{inter}$  é o termo de interação proteína-ligante

$$E_{inter} = \sum_{i \in \text{ligante}} \sum_{j \in \text{proteína}} \left[ E_{PLP}(r_{ij}) + 332.0 \frac{q_i q_j}{4r_{ij}^2} \right] \quad (2)$$

O termo de  $E_{PLP}$  é um “piecewise linear potencial” que usa dois conjuntos diferentes de parâmetros: um para a aproximação do termo de impedimento espacial (van der Waals) entre átomos, e outro é o potencial de

realizar a ligação de hidrogênio. O segundo termo descreve as interações eletrostáticas entre átomos carregados. É um potencial de Coulomb com uma constante dielétrica dependente da distância ( $D(r) = 4r$ ). O valor numérico de 332,0 ajusta as unidades de energia eletrostática em quilocalorias por mol (THOMSEN; CHRISTENSEN, 2006).

O termo  $E_{intra}$ , que corresponde à energia interna do ligante, é determinado de acordo com a Eq. 3:

$$E_{intra} = \sum_{i \in \text{ligante}} \sum_{j \in \text{ligante}} E_{PLP}(r_{ij}) + \sum_{\text{ligações flexíveis}} A[1 - \cos(m \cdot \theta - \theta_0)] + E_{penalizada} \quad (3)$$

O primeiro termo corresponde ao somatório das energias de interação intramolecular entre todos os pares de átomos do ligante, excluindo-se os pares de átomos conectados por duas ligações ou menos. O segundo termo refere-se à energia torcional, onde  $\theta$  é o ângulo de torção da ligação. A média da contribuição de torção da ligação da energia é usada se diversas torsões são determinadas. O último termo,  $E_{penalizada}$ , atribui uma penalidade de 1000 se a distância entre dois átomos pesados (átomos diferentes de hidrogênio) for menor que 2.0 Å, punindo conformações inexistentes do ligante (CAETANO, 2007). Em resumo, estas funções são usadas para ancorar automaticamente uma molécula flexível em uma molécula molde (proteína).

## 2.6 Relação quantitativa estrutura- atividade

A premissa fundamental da relação quantitativa estrutura atividade (QSAR, do inglês Quantitative Structure-Activity Relationship) é que as propriedades de uma molécula são determinadas por sua estrutura molecular (SUTHERLAND; O'BRIEN; WEAVER, 2004). Portanto, a atividade biológica

de um fármaco está relacionada com as características moleculares específicas, que são denominadas de descritores (ANDRADE et al., 2010).

Estudos de QSAR têm amadurecido em uma ferramenta amplamente utilizada, que contribui substancialmente para o processo de descoberta de drogas (LILL, 2007). Essa metodologia economiza recursos e agiliza o processo de desenvolvimento de novas moléculas (KHAN et al., 2009).

Os modelos de QSAR devem ser capazes de prever a atividade de uma série congênere. Além disso, eles são utilizados para fazer previsões sobre novos derivados a fármacos (SUTHERLAND; O'BRIEN; WEAVER, 2004). Uma vez que a estrutura é conhecida, qualquer descritor molecular pode ser calculado, sendo o composto de interesse já sintetizado ou não (YILMAZ et al., 2011).

Os primeiros estudos QSAR basearam-se em uma única propriedade físico-química, como a solubilidade para explicar o efeito biológico da molécula (QSAR 1D). Hansch e Fujita, consideraram que a atividade biológica observada é resultado da contribuição de diversas propriedades físico-químicas (QSAR 2D), e seu modelo tenta identificar a contribuição de cada fator no estabelecimento da resposta biológica (TAVARES, 2004). Informações significativas sobre um conjunto de dados podem ser extraídas, utilizando o QSAR 2D, entretanto, esses descritores não possuem informações relacionadas com as características tridimensionais dos compostos (OPREA, 2002).

No QSAR 3D, considera-se a estrutura tridimensional do ligante, e este modelo fornece uma informação estrutural completa, tendo em conta a topologia, composição e forma espacial da molécula em uma única conformação (KUZ'MIN; ARTEMENKO; MURATOV, 2008). Em 1988, o método de análise comparativa de campo molecular (CoMFA) foi publicado por Cramer et al. (1988). Este método, que é baseado em campo molecular, constituiu o primeiro verdadeiro método de QSAR 3D (KUBINYI, 1997).

O QSAR 4D foi proposto por Hopfinger et al. (1997) e examina o espaço conformacional dos ligantes. Este modelo é semelhante ao modelo 3D, porém, a quarta dimensão na metodologia de QSAR-4D corresponde à amostragem conformacional em função do tempo de simulação pela dinâmica molecular, permitindo assim, identificar a conformação que maximiza o valor da atividade em modelos de QSAR-3D (HOPFINGER et al., 1997; YANMAZ et al., 2011). Assim, após a obtenção de um conjunto de conformações geradas pela simulação da dinâmica molecular, estas devem ser utilizadas para gerar os descritores na metodologia de QSAR-4D.

Na ausência da estrutura do receptor biológico, diferentes protocolos são apresentados no QSAR 5D, como uma solução possível para o problema, permitindo a avaliação simultânea de diferentes cenários de ajuste induzido (VEDANI et al., 2005). Finalmente, o QSAR 6D, a estimativa simultânea de diferentes cenários de solvatação pode ser considerada (VEDANI et al., 2007).



### 3 OBJETIVOS

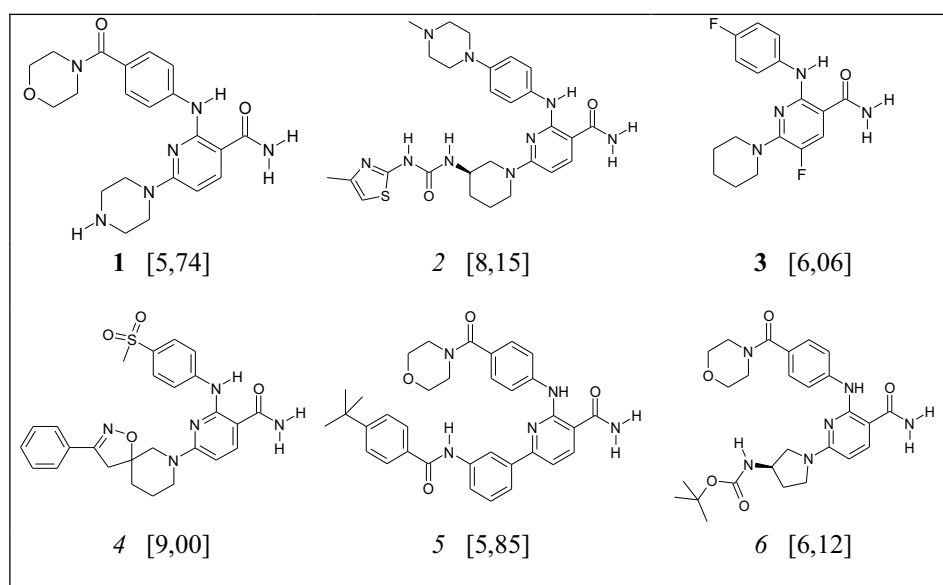
A compreensão das interações específicas entre ligantes e seus receptores auxilia no conhecimento das razões moleculares da atividade dos ligantes. No intuito de estudar novos inibidores da Btk, objetivou-se, neste trabalho:

- a) estudar o modo de interação entre a enzima Btk e 96 análogos nicotinamídicos sintetizados e testados por Delucca et al. (2010) através da técnica de ancoramento molecular;
- b) realizar estudos quantitativos de estrutura-atividade em 4D, dependente do receptor de 71 compostos nicotinamídicos, e propor modelos que permitam prever a atividade de 25 moléculas não testadas;
- c) avaliar as interações dos inibidores no sítio ativo da Btk e comparar os resultados do ancoramento molecular com os de relação estrutura-atividade em 4D, a fim de selecionar as características mais relevantes que um potente inibidor deva apresentar;
- d) propor novos candidatos a fármacos inibidores da Btk.

## 4 METODOLOGIA

### 4.1 Dados Biológicos

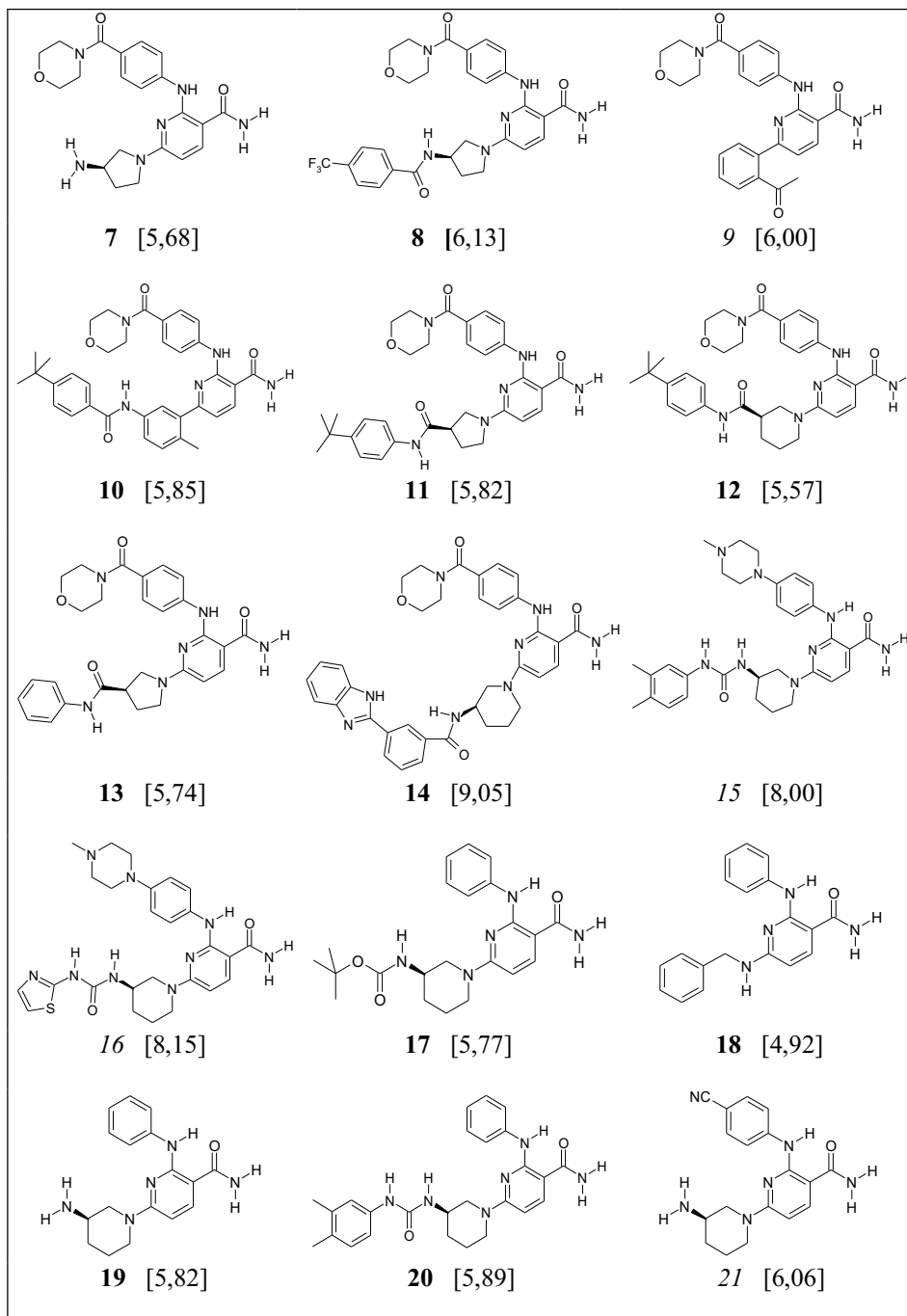
No Quadro 1, estão apresentadas as estruturas dos 96 inibidores da Btk e seus respectivos valores de  $pIC_{50}$  (logaritmo da concentração capaz de inibir 50% da atividade enzimática). Esses compostos foram selecionados, a partir de uma série de análogos nicotinamídicos inibidores da Btk, desenvolvidos por Delucca et al. (2010).



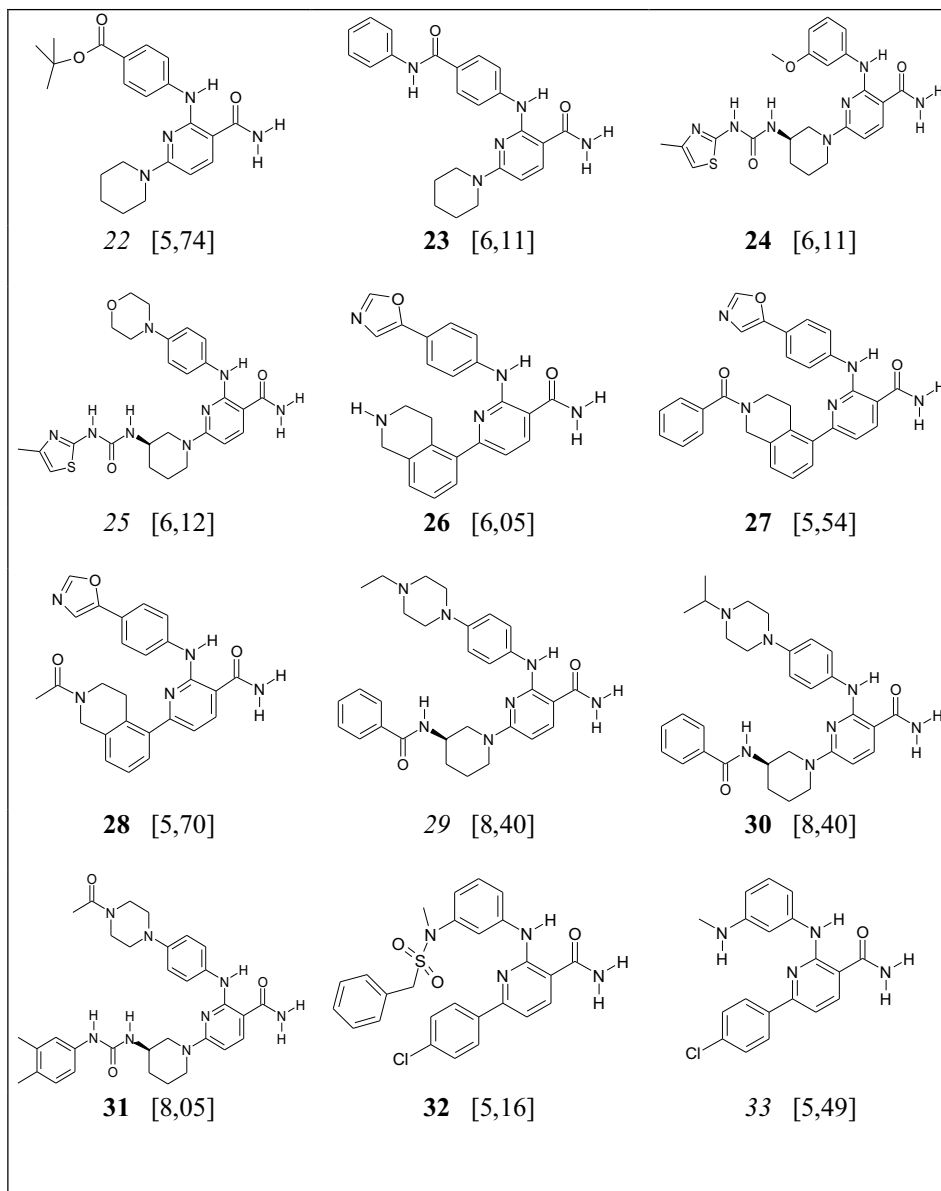
Quadro 1 Estrutura química e os respectivos valores de  $pIC_{50}$  (M) dos 96 análogos nicotinamídicos. Em negrito estão representadas as moléculas selecionadas para compor o grupo treinamento, e em itálico, o grupo teste da metodologia QSAR-4D

(...continua...)

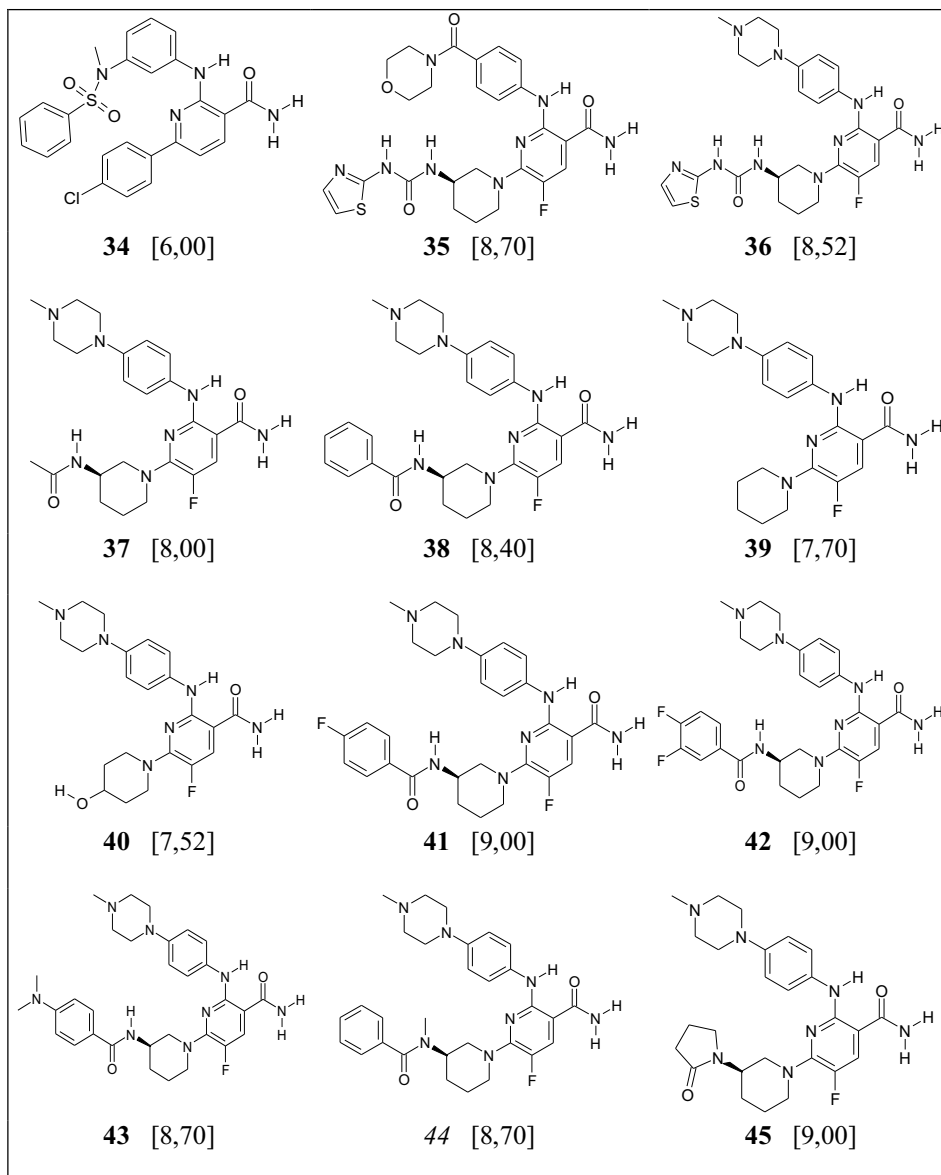
## “Quadro 1, continuação”



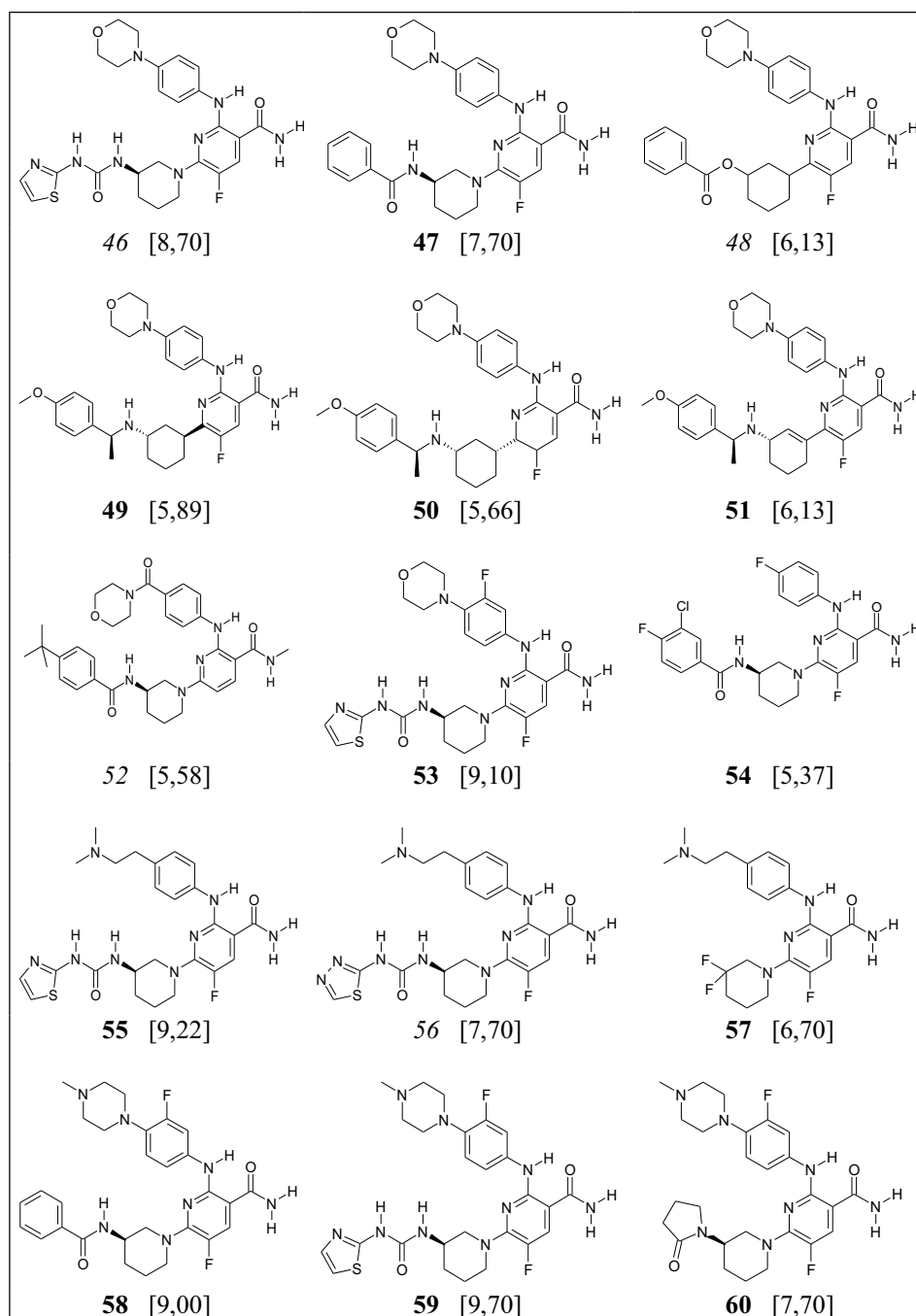
## “Quadro 1, continuação”



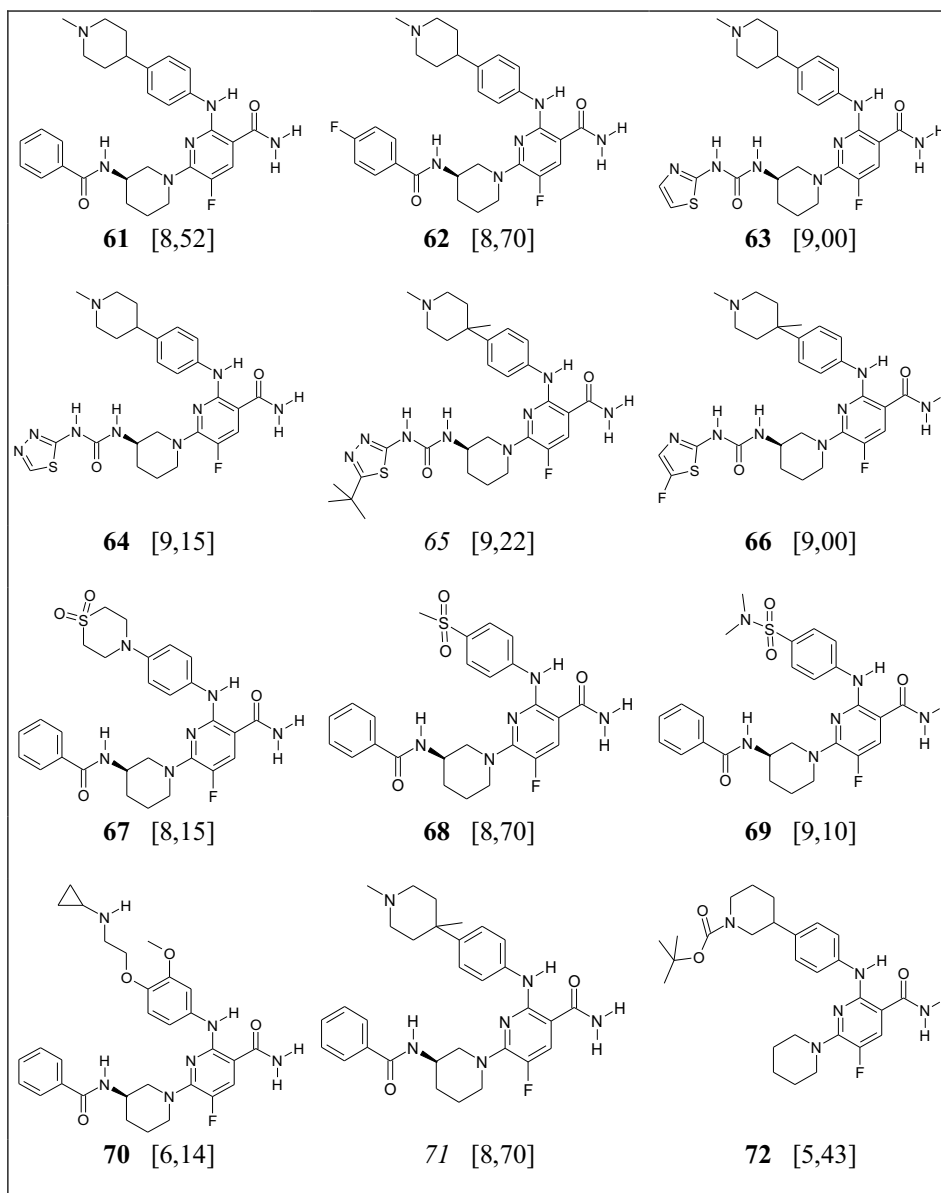
## “Quadro 1, continuação”



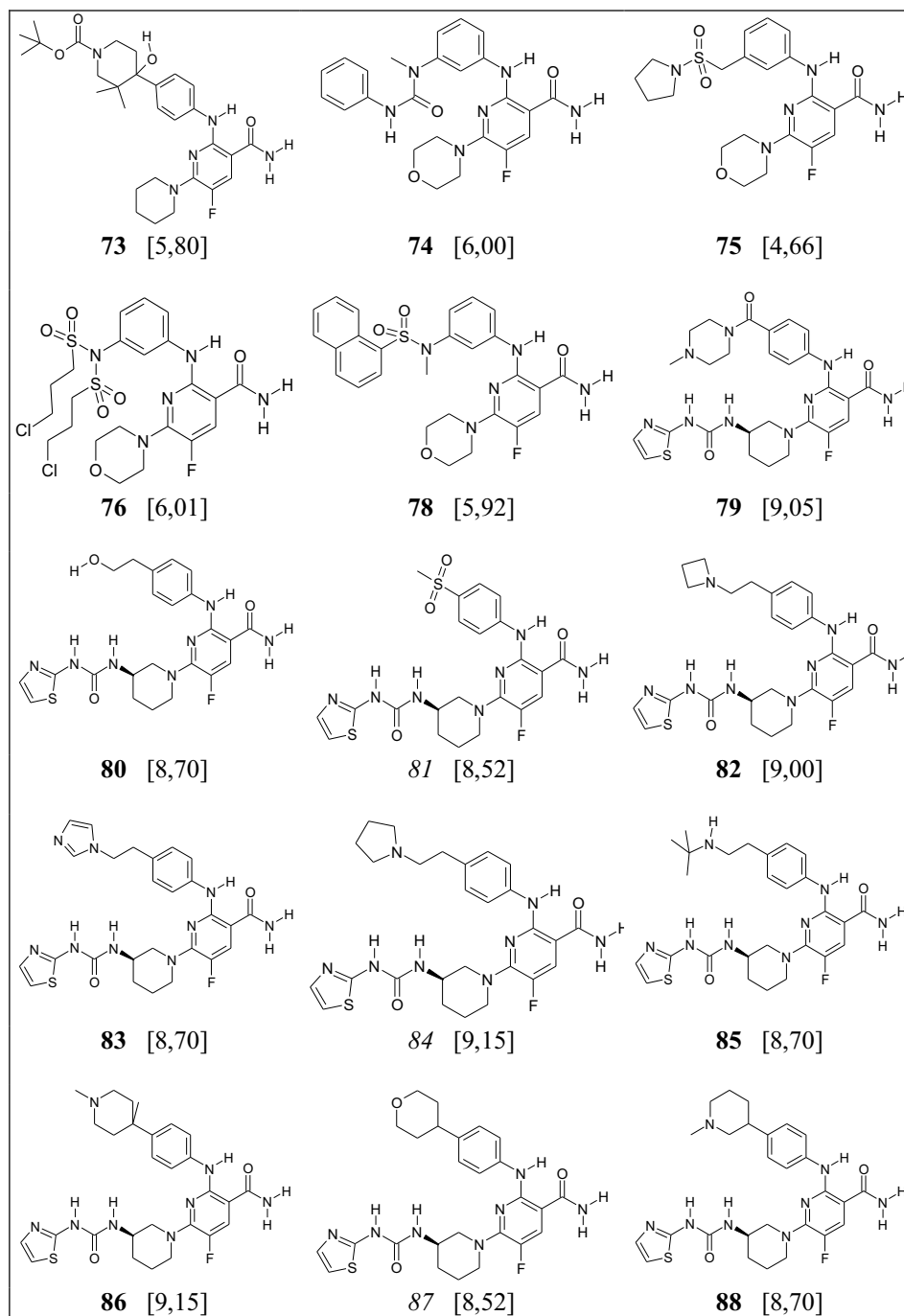
“Quadro 1, continuação”



“Quadro 1, continuação”

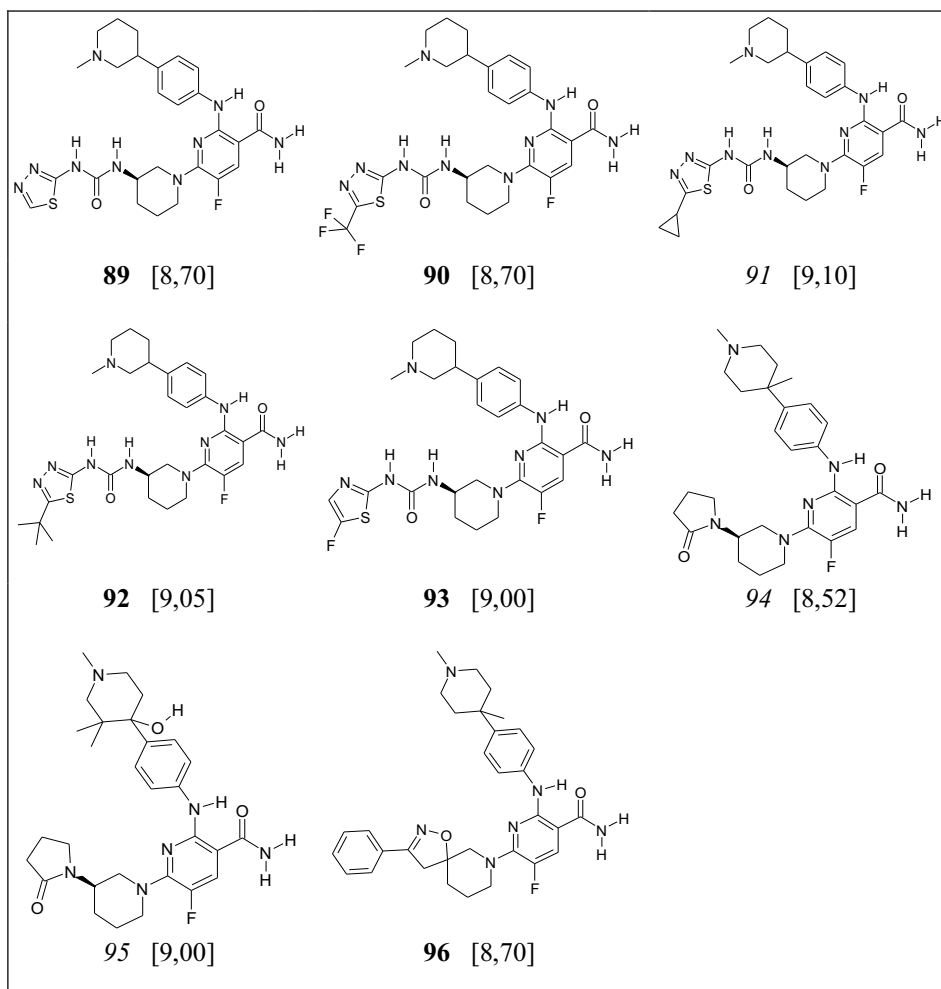


“Quadro 1, continuação”



“Quadro 1, conclusão”





#### 4.2 Cálculo de ancoramento molecular (*docking*)

Com a finalidade de se obter a conformação inicial dos compostos da Tabela 1, necessária para os estudos de ancoramento molecular, suas estruturas tridimensionais foram construídas no programa *HyperChem Release 7.03* (HYPERCHEM, 2002), a geometria foi otimizada, usando-se MM+ como campo de força (ALLINGER, 1977) e foram realizados os cálculos das cargas parciais atômicas através do método semiempírico AM1 (DEWAR et al., 1985).

As coordenadas da estrutura cristalográfica da enzima Btk utilizada foram obtidas a partir do banco de dados *Protein Data Bank* (PDB) (BERMAN et al., 2000), cujo código é o 3OCS, e a resolução do cristal desta proteína é de 1,8 Å.

O programa utilizado para o cálculo das energias do ancoramento molecular foi o *Molegro Virtual Docker* (MVD) (THOMSEN; CHRISTENSEN, 2006), sendo que este permite determinar a conformação mais provável de interação do ligante com uma macromolécula. As estruturas tridimensionais dos compostos previamente minimizadas foram transferidas para o programa MVD e cada inibidor foi ancorado ao sítio ativo da enzima Btk (Figura 5), utilizando como referência o composto 4-terc-butil-N-[2-metil-3-(4-metil-6-{[4-(morfolin-4-ilcarbonil)fenil] amino} -5-oxi-4,5-dihidropirazin-2-il)fenil]benzamida (CGI1746) (Figura 7), por possuir uma estrutura química muito similar à dos compostos estudados. Não foi adicionada flexibilidade aos resíduos de aminoácidos do sítio ativo, pois a cavidade da Btk é extensa e os compostos possuem uma liberdade conformacional consideravelmente grande. Estes dois fatores iriam atrapalhar na escolha da conformação.

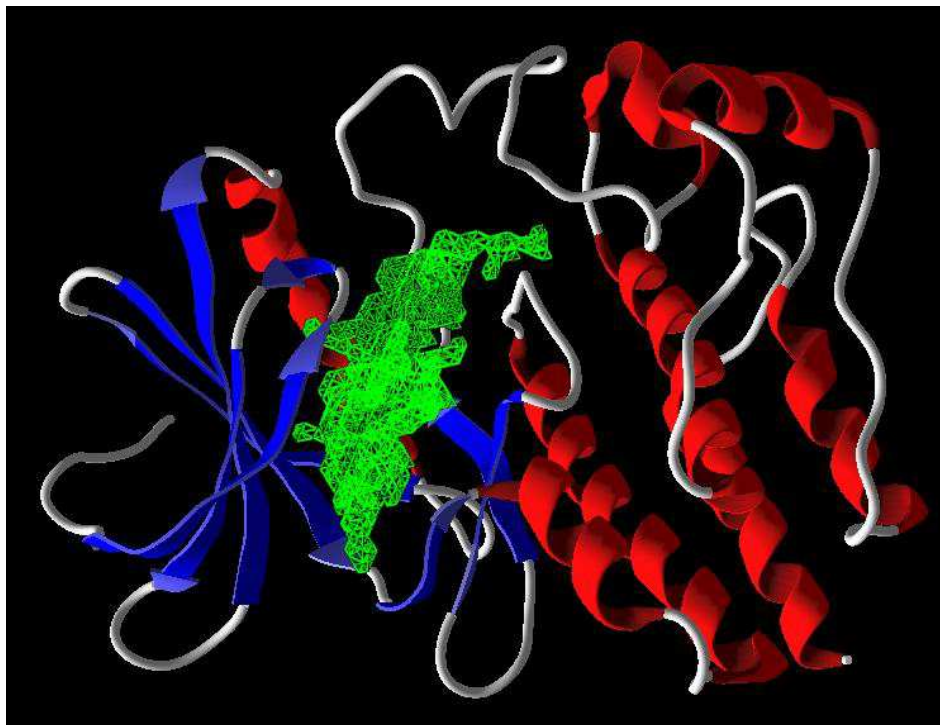


Figura 5 Estrutura terciária da Btk obtida do PDB (código 3OCS). As  $\alpha$ -hélices estão representadas em vermelho, as folhas- $\beta$  em azul e o sítio ativo em verde

### 4.3 QSAR-4D

A abordagem QSAR-4D pode ser aplicada para a análise independente do receptor (IR) ou dependente do receptor (DR). Na primeira, ou a geometria do receptor não está disponível ou esta não é utilizada na análise de QSAR-4D devido à incerteza na geometria do receptor e / ou o modo de ligação. Em contraste, na análise de QSAR-4D DR, a geometria do receptor (ou alvo molecular) disponível é utilizada. Além disso, os modelos de QSAR-4D DR podem "capturar" a existência de interações específicas pelo ajuste induzido

entre o sítio ativo da enzima e cada inibidor específico (SANTOS; HOPFINGER, 2006). A abordagem QSAR-4D DR foi aplicada neste estudo.

#### **4.3.1 Grupos teste e treinamento**

Os modelos de QSAR-4D foram desenvolvidos utilizando um conjunto de 71 compostos (em **negrito** na Tabela 1), denominados de conjunto de treinamento, e validados externamente utilizando um conjunto de 25 compostos, o conjunto de teste (em *itálico* na Tabela 1).

Os valores de  $IC_{50}$  dos 96 compostos expressos em  $\mu M$  foram convertidos para unidade molar e então transformados para  $-\log IC_{50}$  ( $pIC_{50}$ ). Isto altera uma faixa ampla de valores em  $IC_{50}$  para uma estreita faixa em  $pIC_{50}$ , viando a interpretação das análises de QSAR. A divisão entre compostos do conjunto de treinamento e de teste foi feita separando-se os compostos nicotinâmicos, estruturalmente relacionados, em sete subgrupos de acordo com faixas de valores de atividade biológica. Aproximadamente, 25% dos compostos de cada subgrupo foram randomicamente selecionados para compor o conjunto de teste (em *itálico*) (Quadro 1), já os remanescentes correspondem ao conjunto de treinamento (em **negrito**) (Quadro 1). Os subgrupos I, II, III, IV, V, VI e VII correspondem aos compostos com valores de  $pIC_{50}$  entre 4,50-5,50; 5,51-6,00; 6,01-6,50; 7,51-8,00; 8,01-8,50; 8,51-9,00 e 9,01-9,70, respectivamente.

A faixa dos valores de  $pIC_{50}$  (conjuntos de teste e de treinamento) abrange pelo menos cinco ordens de magnitude em unidade logarítmica (4,65 a 9,7) e os valores de atividade biológica apresentam uma distribuição regular sobre toda a extensão.

### 4.3.2 Simulação de dinâmica molecular

Após avaliar os modos de interação de cada ligante no sítio ativo da Btk, pela metodologia de ancoramento molecular, a melhor conformação de cada estrutura foi selecionada. A enzima Btk consiste em 4160 átomos, assim, a demanda computacional exigida pela simulação por dinâmica molecular para todos os 96 complexos proteína-ligante seria muito grande. Então a técnica baseada no corte da enzima em um raio de cerca de 5 Å ao redor do ligante CGI1746 foi empregada. À estrutura obtida foram adicionados os átomos de hidrogênio correspondentes, a fim de manter a integridade geométrica do receptor. Desta forma, esses 96 complexos resultantes foram submetidos ao processo de simulação da dinâmica molecular (SDM), para gerar um perfil de amostragem conformacional (CEP, do inglês Conformational Ensemble Profile), para cada inibidor. A quarta dimensão na metodologia de QSAR-4D corresponde à amostragem conformacional, em função do tempo de simulação de dinâmica molecular (SDM). Essa amostragem conformacional corresponde a um conjunto estatisticamente representativo de conformações distintas de baixa energia (4D-QSAR..., 1997; HOPFINGER et al., 1997).

O processo de SDM foi executado usando o pacote MOLSIM (DOHERTY, 1997), incluso no programa QSAR-4D com campo de força MM2 (ALLINGER; BURKERT, 1982; WEINER et al., 1986), que inclui parametrizações descritas por Hopfinger e Pearlstein (1984). O tempo de simulação da dinâmica molecular é o produto do número de períodos de simulação pelo tempo de duração de cada período. O tempo total da simulação foi de 100 ps, integrada em intervalos de 0,001ps. A temperatura para cada SDM foi de 300K.

Os cálculos de dinâmica foram executados, aplicando-se a constante dielétrica da água no intuito de tentar modelar o efeito de solvatação na ausência do solvente explícito. Adicionalmente, todos os resíduos de aminoácidos

selecionados foram fixados para prevenir uma grande mudança conformacional dos complexos, o que poderia resultar em conformações estericamente proibidas.

### 4.3.3 Definição do alinhamento

O alinhamento é a maneira pela qual os compostos do conjunto de treinamento são comparados. A qualidade do modelo gerado depende do alinhamento testado, sendo assim, esta é considerada uma das etapas mais importantes na metodologia de QSAR-4D. Como os compostos nicotinâmídicos são congêneres, podemos assumir que eles irão apresentar o mesmo mecanismo de inibição no sítio ativo da Btk (ROMEIRO et al., 2005).

Em geral, os alinhamentos são escolhidos de modo a abranger uma estrutura em comum dos compostos dos grupos treinamento e teste. Alinhamentos usando átomos da direita, esquerda e do meio, ou que usam átomos que abrangem toda a região similar dos ligantes devem ser usados para garantir um bom alinhamento (OLIVEIRA; RAMALHO; CUNHA, 2009; ROMEIRO et al., 2005).

A ordem dos átomos escolhidos para cada alinhamento é importante, já que durante o procedimento o primeiro átomo selecionado é posicionado na origem dos três eixos cartesianos e a molécula é girada de modo que o segundo átomo se posicione no eixo x, e o terceiro se posicione no eixo y (HOPFINGER et al., 1997).

Para os três alinhamentos realizados foram escolhidos átomos do carbono-alfa ( $C_{\alpha}$ ) da proteína (Tabela 1) já que estes são comuns a todos os complexos e abrangem a estrutura dos compostos estudados.

Tabela 1 Alinhamentos testados e os átomos de  $C_{\alpha}$  da proteína selecionados para o alinhamento dos complexos. O composto **59** está representado em cinza escuro

Alinhamento	$C_{\alpha}$
1	Met477/Tyr476/Tyr551
2	Tyr551/Met477/Tyr476
3	Met477/Tyr476/Ser543

#### 4.3.4 Definição do tamanho das células

Um perfil de amostragem conformacional (CEP, do inglês Conformational Ensemble Profile), para cada composto foi obtido através da simulação de dinâmica molecular e, em seguida, este conjunto de distintas conformações (500) foi inserido dentro de uma caixa tridimensional virtual. A caixa é composta por células cúbicas que podem ter dimensões variadas. Quanto menor o tamanho da célula, mais refinados serão os modelos, entretanto, o tempo de análise aumenta devido ao maior número de variáveis geradas. Para cada um dos três alinhamentos foram utilizadas células de 2,0Å e 5,0Å de aresta.

#### 4.3.5 Definição dos elementos de interação farmacofórica

De acordo com a metodologia de QSAR-4D (HOPFINGER et al., 1997), os átomos de cada molécula são classificados em sete tipos de elementos de interação farmacofórica (IPE, do inglês *Interaction Pharmacophore Elements*), e correspondem aos tipos de interações específicas que podem ocorrer no sítio ativo e estão relacionados com os grupos farmacofóricos. A Tabela 2 relaciona os seis tipos de IPEs utilizados neste trabalho.

Tabela 2 Definição dos elementos de interação farmacofórica, IPEs

IPEs	Definição	Código
1	interação não polar	np
2	interação polar positiva	p+
3	interação polar negativa	p-
4	acceptor de ligação hidrogênio	alh
5	doador de ligação hidrogênio	dlh
6	sistema aromático	ar

A frequência de ocupação de cada célula da caixa, por cada tipo de IPE, é registrada ao longo das conformações geradas na SDM e forma o conjunto de descritores de ocupação da célula (GCOD, Grid Cell Occupancy Descriptor), que são as variáveis independentes usadas na construção dos modelos de QSAR-4D.

#### 4.3.6 Pré-tratamento dos descritores (GCODs)

A análise de QSAR-4D gera um enorme número de variáveis (os descritores), devido ao grande número de IPEs e células da caixa utilizados. Além disso, existe a possibilidade de haver descritores altamente



correlacionados o que dificulta a interpretação e validação dos modelos (PAN; TSENG; HOPFINGER, 2003; PITA et al., 2012).

A fim de reduzir o número de variáveis independentes (GCODs) e melhorar a qualidade dos GCODs, eliminando os indesejados, efetuaram-se dois tipos de pré-tratamento do banco de dados, antes da etapa de geração dos modelos (equações). No primeiro foi usado o método de regressão dos mínimos quadrados parciais (PLS, do inglês *Partial Least Squares*) (CRAMER et al., 1988; CUNHA et al., 2006), a fim de identificar e selecionar os GCODs de maior contribuição (peso) para a variação da atividade biológica. No segundo tipo de pré-tratamento, empregamos dois critérios para a exclusão de variáveis: no primeiro, os GCODs que apresentavam valor de ocupância igual a zero (variância igual a zero) para todos os compostos foram eliminados; e no segundo, os GCODs, que apresentaram valores de ocupação diferente de zero em dez compostos ou menos, foram excluídos, descartando assim, GCODs pouco representativos em função do número de compostos.

#### **4.3.7 Obtenção dos modelos de QSAR-4D**

Os GCODs selecionados foram processados em conjunto com os valores de atividade biológica (variável dependente), utilizando uma metodologia que combina algoritmos genéticos (GA, do inglês *Genetic Algorithm*) (JIN, 1994; ROGERS, 1994) e mínimos quadrados parciais (PLS, do inglês *Partial Least Squares*), denominada de aproximação da função genética (GFA, do inglês *Genetic Function Approximation*) (ROGERS; HOPFINGER, 1994), implementada no programa 4D-QSAR (HOPFINGER et al., 1997). No GFA, os modelos iniciais são criados pela seleção randômica dos GCODs do conjunto de treinamento, gerando uma população de modelos. Modelos acurados são produzidos por operações de cruzamento (crossover) e mutação (mutation), recombinando os termos dos melhores modelos. A grande vantagem do GFA é a

construção e otimização de múltiplos modelos, ao invés da otimização de apenas um único modelo (CUNHA et al., 2006; ROGERS; HOPFINGER, 1994; ROMEIRO et al., 2005).

Os cálculos de GA-PLS foram iniciados com 10000 modelos (equações) gerados randomicamente, cada um contendo inicialmente quinze variáveis (GCODs), com uma taxa de probabilidade de mutação de 100%. Um número máximo de sete componentes foi usado e o crossover variado entre 10000 e 500000. O smoothing factor, que controla o número de descritores (GCODs) nos modelos de QSAR, foi variado de 1,0 a 3,0, no intuito de se obter equações com no máximo quatorze termos. Equações com mais de quatorze termos possivelmente provocará um superajuste (overfitting) nos dados, pois o número máximo de termos pode ser definido pelo número de compostos utilizado no conjunto de treinamento, dividido por 5 (KUBINYI, 1993).

Os melhores modelos resultantes de cada alinhamento testado no estudo de QSAR-4D foram ordenados pelo valor do fator de desajuste (LOF, do inglês lack-of-fit) de Friedman (KARKI; KULKARNI, 2001), que é a medida dos mínimos quadrados penalizados; quando dois métodos têm o mesmo erro de mínimos quadrados, aquele que possuir o menor número de termos será escolhido (ROGERS; HOPFINGER, 1994).

#### **4.3.8 Validação interna dos modelos**

Os melhores modelos obtidos por GA-PLS foram submetidos à validação interna (validação cruzada). A técnica de validação interna utilizada no programa 4D-QSAR é denominada de “leave-one-out cross-validation” (LOO-cv,  $q^2$ ), e consiste na exclusão de um composto do conjunto de treinamento por vez, e os coeficientes da equação são recalculados com base nos compostos remanescentes (n-1); a nova equação obtida é usada para prever a atividade do composto que foi excluído. Quando todos os compostos forem

eliminados uma vez, é gerado o índice estatístico coeficiente de correlação quadrático da validação cruzada,  $q^2$  (KUBINYI, 1993; SELASSIE, 2003).

Os valores de  $q^2$  devem ser transformados em  $q^2_{ajustado}$ , demonstrado pela Equação 4, que permite a comparação entre modelos com número diferente de variáveis (LIVINGSTONE, 1995).

$$q^2_{ajustado} = \frac{1-(1-q^2)}{\frac{n-1}{n-p}} \quad (4)$$

Nesta equação  $q^2$  corresponde ao quadrado do coeficiente de correlação após a validação cruzada,  $n$  corresponde ao número de compostos do conjunto de treinamento e  $p$  corresponde à quantidade de variáveis do modelo.

#### 4.3.9 Validação externa dos modelos

O melhor modelo de QSAR-4D obtido foi submetido à validação externa, usando o conjunto de teste contendo vinte e cinco compostos que não foram incluídos no desenvolvimento dos modelos de QSAR-4D. Este processo verificará a capacidade do modelo de prever a atividade biológica de moléculas congêneres.

#### 4.3.10 Identificação dos compostos *outliers*

Os compostos que não foram bem preditos pela equação (*outliers*), no método de QSAR-4D, são identificados, a partir da observação dos valores residuais calculados para os compostos do conjunto de treinamento e do conjunto de teste. Neste trabalho, os *outliers* foram definidos como sendo os compostos cujos valores modulares da diferença entre o  $pIC_{50}$  predito e o  $pIC_{50}$  experimental sejam maiores do que o dobro do desvio padrão desses resíduos (HOPFINGER et al., 1997).

#### **4.3.11 Seleção da conformação “representativa”**

Objetivou-se, nesta etapa, prever uma conformação de cada composto, utilizando as conformações de menor energia do perfil da amostragem conformacional (CEP). A metodologia de QSAR-4D assume que esta seria a conformação que maximiza a potência do composto no modelo selecionado. Somente conformações termodinamicamente acessíveis serão consideradas. Um critério é utilizar todas as conformações de cada composto, dentro de um intervalo de energia ( $\Delta E$ ). A conformação foi selecionada dentro de uma faixa de  $5 \text{ kcal.mol}^{-1}$  da conformação de menor energia, que maximiza a atividade biológica (HOPFINGER et al., 1997).

## 5 RESULTADOS E DISCUSSÃO

### 5.1 Ancoramento molecular

Um algoritmo de predição da cavidade, baseado em uma grade tridimensional foi usado para gerar os potenciais sítios de ligação da enzima Btk, por meio do programa *MVD*. O volume da cavidade obtida foi de 291,84 Å, demonstrado na Figura 6.

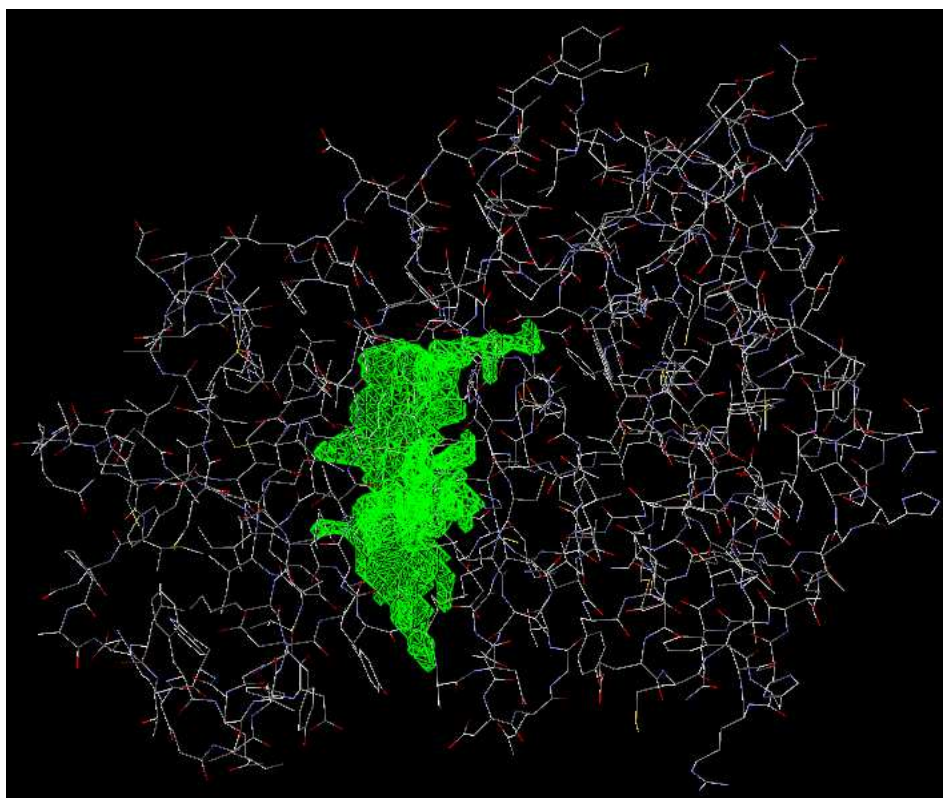


Figura 6 Cavidade da enzima BTK ilustrada em verde

Os compostos nicotinâmicos foram ancorados no sítio ativo da enzima Btk. A melhor conformação de cada ligante foi selecionada de acordo com sua posição no sítio ativo, sendo que a conformação escolhida, para cada um dos inibidores, foi aquela que possuía o maior grau de similaridade espacial com o ligante CGI1746 que se encontrava cristalizado na proteína obtida do PDB como código 3OCS (Figura 7).

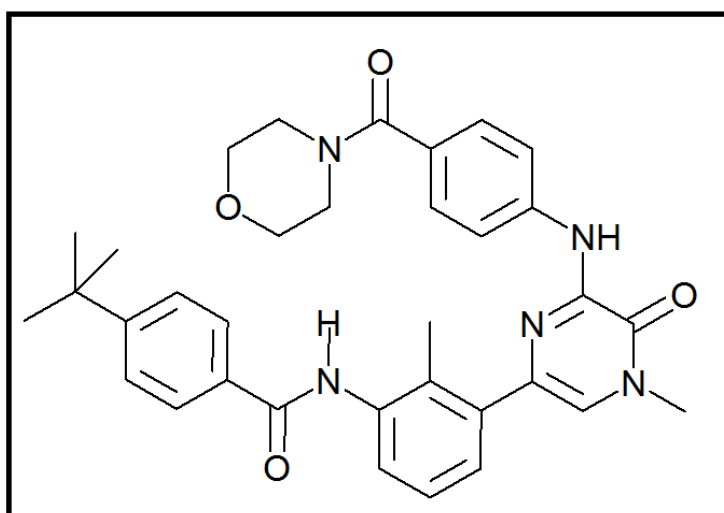


Figura 7 Estrutura do ligante 4-terc-butil-N-[2-metil-3-(4-metil-6-{[4-(morfolin-4-ilcarbonil)fenil] amino} -5-oxi-4,5-dihidropirazin-2-il)fenil]benzamida (CGI1746) que foi obtido do PDB (código 3OCS)

Os resultados teóricos obtidos do ancoramento molecular foram avaliados e comparados com base nos resultados experimentais descritos por Delucca et al. (2010).

As conformações escolhidas foram utilizadas para a análise das interações com os resíduos de aminoácidos do sítio ativo. Para avaliar melhor os fatores mais relevantes para a potência dos compostos, os seguintes parâmetros foram calculados: (a) energia de ligação de hidrogênio entre o ligante e a

proteína ( $E_{\text{Lig.H}}$ ); (b) energia intramolecular do ligante ( $E_{\text{Intra}}$ ); (c) energia eletrostática com raio  $< 4,5 \text{ \AA}$  entre a proteína e o ligante ( $E_{\text{Elet}}$ ); (d) energia eletrostática com raio  $> 4,5 \text{ \AA}$  entre a proteína e o ligante ( $E_{\text{EletLong}}$ ); e (e) energia de impedimento espacial entre a proteína e o ligante ( $E_{\text{Est}}$ ). Foi realizada uma regressão linear múltipla para correlacionar os valores dos termos de energia independentes com os valores de  $\text{pIC}_{50}$  através da seguinte equação:

$$\text{pIC}_{50} = C + C_1 E_{\text{Est}} + C_2 E_{\text{Intra}} + C_3 E_{\text{LigH}} + C_4 E_{\text{Elet}} + C_5 E_{\text{EletLong}} \quad (5)$$

em que  $C$ ,  $C_1$ ,  $C_2$ ,  $C_3$ ,  $C_4$ , e  $C_5$  são os coeficientes obtidos da regressão multivariada:

$$\text{pIC}_{50} = 0,82 - 0,05 E_{\text{Est}} - 0,04 E_{\text{Intra}} + 0,00 E_{\text{LigH}} + 0,04 E_{\text{Elet}} + 0,34 E_{\text{EletLong}} \quad (6)$$

O valor de  $R^2$  (coeficiente de determinação) obtido foi de aproximadamente 0,80. Um  $R^2$  maior do que 0,7 indica que houve uma alta correlação entre os dados (HOPFINGER et al., 1997). Isso sugere que o modelo de ancoramento molecular obtido é satisfatório para avaliar o modo de interação dos 96 ligantes com a enzima Btk. A Equação 6 foi estabelecida com os 96 compostos utilizados. Nesta equação, os valores negativos dos coeficientes representam uma maior estabilidade dos complexos, e favorecem para o aumento na potência dos compostos. Já os coeficientes positivos correspondem às interações desfavoráveis entre o ligante e os resíduos de aminoácidos do sítio ativo da enzima, sendo assim, essas interações fazem com que a potência dos compostos diminua.

Verifica-se que o termo  $E_{\text{Est}}$  tem grande peso na potência dos compostos, visto que apresenta o coeficiente de regressão com valor mais negativo, -0,05

(Eq.6). Assim, para que os compostos aumentem sua potência inibitória é desejável que o fator estérico seja o mais estável possível, ou seja, possua um valor baixo de  $E_{Est}$ . O inibidor **14** apresenta um valor de energia de impedimento espacial menor dentre os 96 compostos estudados. Embora o inibidor **59** seja o mais potente desta série, ele não apresenta um valor de  $E_{Est}$  tão estável quanto o da molécula **14** (Tabela 3). Isto ocorre, provavelmente, devido ao fato de o inibidor **14** preencher uma porção maior do “bolsão H3” na Btk. A Figura 8 ilustra as conformações dos inibidores **14** e **59** ancoradas na Btk. A Figura 9 representa as conformações de ancoramento para as moléculas **65**, **75** e **59**, respectivamente, em comparação com a molécula **14**, e suas ocupações no bolsão H3 da Btk. Podemos perceber que o inibidor **75** não ocupa o bolsão de especificidade na Btk, e assim, possui o valor de  $E_{Est}$  menos estável, enquanto que o inibidor **14** possui o valor de  $E_{Est}$  mais estável, e é o composto que melhor preenche esse bolsão. O inibidor **65** também ocupa uma porção do bolsão H3 um pouco maior do que o inibidor **59**, e então, apresenta um valor mais estável de  $E_{Est}$ . Visto que a contribuição do efeito estérico é um dos fatores que mais influencia para o aumento na potência dos compostos, a ocupação do “bolsão H3” é um fator relevante para a efetividade dos inibidores. Isso reforça a idéia de Di Paolo et al. (2011) de que os inibidores que interagem e estabilizam esse bolsão de especificidade, bloqueiam a fosforilação do loop de ativação da Btk.



Tabela 3 Valores de  $pIC_{50}$  experimentais e de energia de impedimento espacial ( $kcal.mol^{-1}$ ) obtidos do ancoramento com a enzima Btk para as moléculas **14**, **59**, **65**, e **75**

Compostos	$pIC_{50}(M)$	$E_{Est}$
<b>14</b>	9,05	-190,75
<b>59</b>	9,70	-165,92
<b>65</b>	9,10	-182,46
<b>75</b>	4,66	-81,79

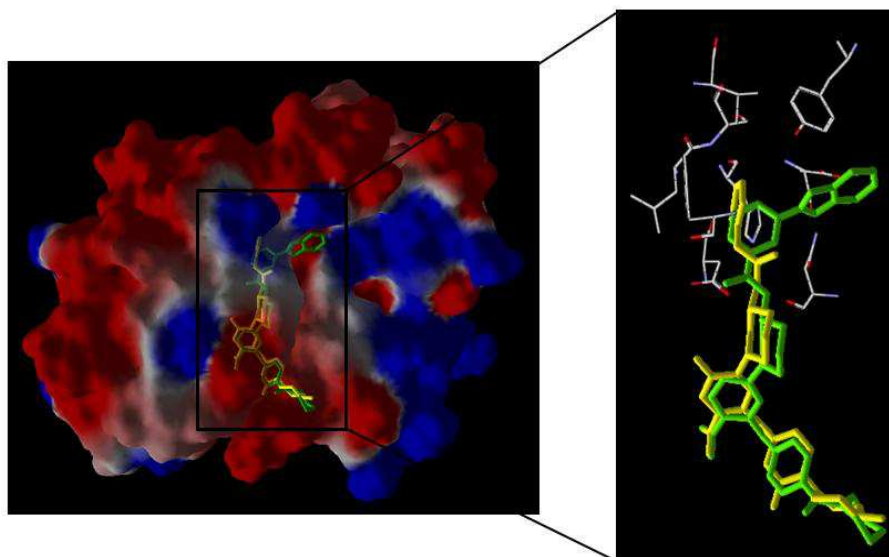


Figura 8 A figura ilustra a superfície eletrostática da Btk e os inibidores **14** (em verde) e **59** (em amarelo), na conformação de ancoramento

Nota: Em vermelho, são representadas as regiões eletrostáticas negativas, e em azul, as regiões eletrostáticas positivas. Os resíduos de aminoácidos mostrados pertencem ao “bolsão H3”.

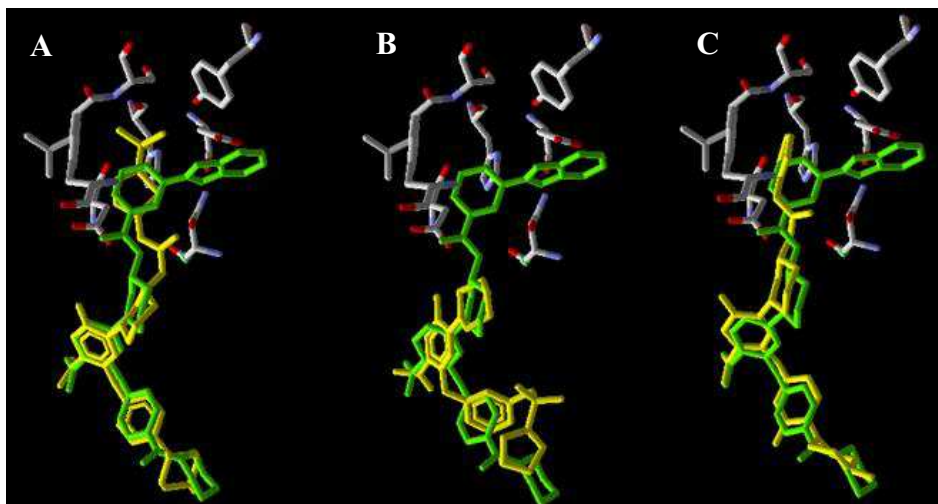


Figura 9 A figura ilustra o bolsão H3 na Btk e os inibidores: A) **14** (em verde) e **65** (em amarelo); B) **14** (em verde) e **75** (em amarelo); C) **14** (em verde) e **59** (em amarelo), ancorados no sítio ativo da enzima

Outro fator importante para o aumento na potência dos inibidores é a  $E_{\text{Intra}}$ , já que possui um coeficiente de regressão de  $-0,04$  (Eq. 6). Assim, os valores mais baixos de  $E_{\text{Intra}}$  correspondem a valores de energia mais estáveis. Dentre os 96 compostos analisados, as moléculas **36**, **59**, **60** e **80** são as que possuem os valores de  $E_{\text{Intra}}$  mais baixos (Tabela 4). Verifica-se que todas elas possuem um átomo de flúor no anel de nicotinamida, e este elemento eletronegativo está localizado próximo ao átomo de nitrogênio da piperidina, como ilustrado na Figura 10. Já as moléculas **12** e **52** são as que possuem os valores de  $E_{\text{Intra}}$  menos estáveis; elas não possuem o átomo de flúor nessa posição. A molécula **14** também não possui um átomo eletronegativo nesse local e, assim, apresenta um valor de energia de interação intramolecular pouco estável. Possivelmente, a presença do átomo de flúor eletronegativo próximo ao átomo de nitrogênio esteja contribuindo para uma interação eletrostática

intramolecular nas moléculas que apresentam os valores mais estáveis de  $E_{\text{Intra}}$ . A molécula **59** é a mais potente e possui um valor de  $E_{\text{Intra}}$  bastante estável.

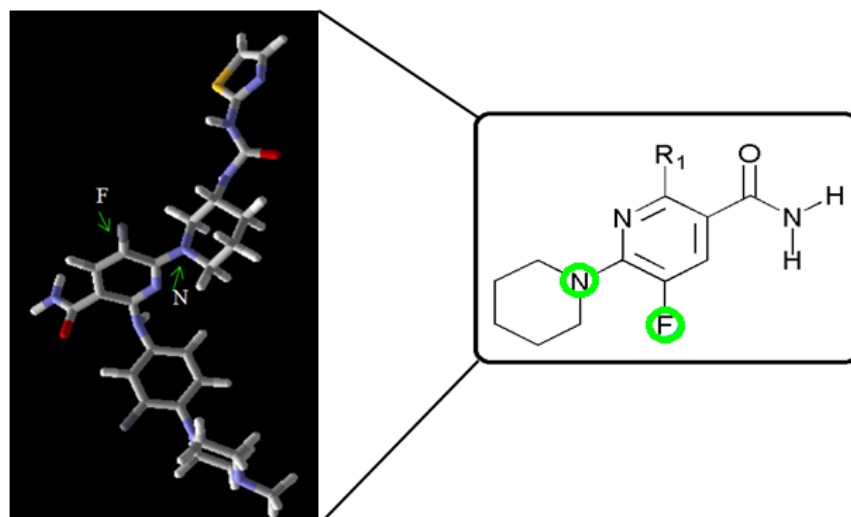


Figura 10 Ilustração da conformação de ancoramento da molécula **59**. O átomo de flúor no anel de nicotinamida e o átomo de nitrogênio da piperidina estão marcados em verde

Tabela 4 Valores de  $pIC_{50}$  experimentais e de energia de interação intramolecular ( $\text{kcal.mol}^{-1}$ ) obtidos do ancoramento com a enzima Btk para as moléculas **12**, **14**, **36**, **52**, **59**, **60**, e **80**

Compostos	$pIC_{50}(M)$	$E_{\text{Intra}}$
<b>12</b>	5,57	0,00
<b>14</b>	9,05	-22,35
<b>36</b>	8,52	-47,77
<b>52</b>	5,58	-8,32
<b>59</b>	9,70	-46,59
<b>60</b>	7,70	-46,70
<b>80</b>	8,70	-47,66

O fator  $E_{\text{EletLong}}$  tem um grande peso na potência dos compostos, pois apresenta um coeficiente de determinação de 0,34 (Eq.6). Coeficientes positivos correspondem a interações desfavoráveis entre o ligante e os resíduos de aminoácidos do sítio ativo da Btk, resultando em um decréscimo na potência dos compostos. Os inibidores **24** e **70** (Tabela 5), por exemplo, possuem baixos valores de potência inibitória e apresentaram também, baixos valores de  $E_{\text{EletLong}}$ , demonstrando uma boa concordância com os resultados observados experimentalmente. Já as moléculas **58**, **60** e **61** (Tabela 5) possuem valores maiores de  $E_{\text{EletLong}}$ , deste modo, o fator interação eletrostática a longa distância (com raio  $> 4,5\text{\AA}$  entre a proteína e o ligante) para esses inibidores, não contribuiu muito para a diminuição de suas potências. O inibidor **59** também possui um valor de  $E_{\text{EletLong}}$  maior, e este apresenta a maior atividade experimental.

Tabela 5 Valores de  $\text{pIC}_{50}$  experimentais e de energia de interação eletrostática com raio  $> 4,5\text{\AA}$  entre a proteína e o ligante ( $\text{kcal.mol}^{-1}$ ) obtidos do ancoramento com a enzima Btk para as moléculas **24**, **58**, **59**, **60**, **61**, e **70**

Compostos	$\text{pIC}_{50}(\text{M})$	$E_{\text{EletLong}}$
<b>24</b>	6,11	-4,47
<b>58</b>	9,00	-0,10
<b>59</b>	9,70	-1,47
<b>60</b>	7,70	0,00
<b>61</b>	8,52	-0,16
<b>70</b>	6,14	-4,32

O fator  $E_{\text{Elet}}$  possui um coeficiente positivo de regressão com menor magnitude do que o fator  $E_{\text{EletLong}}$ . Já para o fator  $E_{\text{LigH}}$  não foi observada

influência para o aumento na potência dos inibidores, pois apresentou um coeficiente de 0,00.

## 5.2 QSAR-4D

A metodologia QSAR-4D dependente do receptor foi aplicada para os 96 complexos ligante-receptor, considerando as conformações dos ligantes obtidas do ancoramento molecular e a estrutura cristalográfica da enzima Btk em um raio de corte com cerca de 5 Å ao redor do ligante CGI1746 (Figura 11).

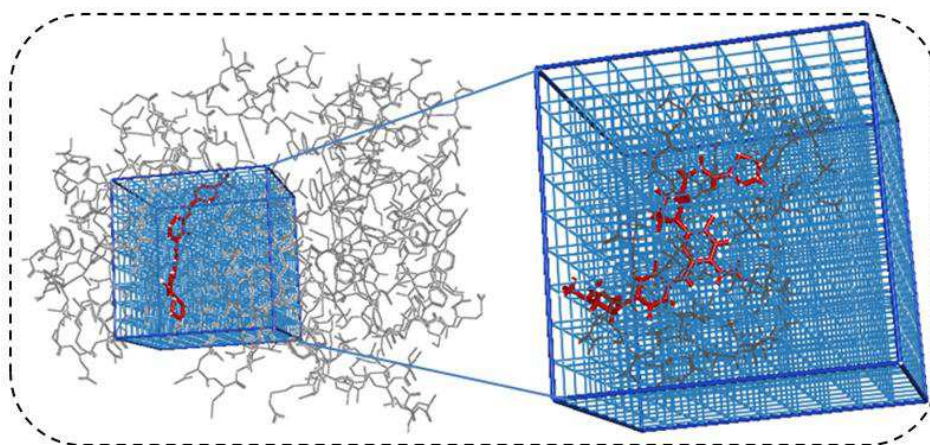


Figura 11 Imagem ilustrativa do corte realizado na enzima Btk (código 3OCS) ao redor do ligante CGI1746, a fim de selecionar os resíduos de aminoácidos a serem utilizados na metodologia QSAR-4D DR para os 96 compostos

## 5.3 Seleção do melhor alinhamento de QSAR-4D

Durante o estudo de QSAR-4D, o espaço conformacional de cada um dos complexos foi avaliado por simulações de dinâmica molecular, mantendo-se

fixos os átomos de toda a estrutura proteica selecionada. Esta etapa teve por finalidade gerar um perfil de amostragem conformacional, produzindo 500 conformações para cada composto. Seis tipos de átomos foram utilizados para definir os elementos de interação farmacofórica (IPEs): np (1), p+ (2), p- (3), alh (4), dlh (5) e ar (6). As conformações de cada um dos complexos foram alinhadas de três formas diferentes, usando dois tamanhos de células cúbicas (Tabela 6):

Tabela 6 Alinhamentos testados no QSAR-4D DR para a Btk

Alinhamento	$C_{\alpha}$	Tamanho das células cúbicas (Å)
<i>1a</i>	Met477/Tyr476/Tyr551	2,0
<i>1b</i>	Met477/Tyr476/Tyr551	5,0
<i>2a</i>	Tyr551/Met477/Tyr476	2,0
<i>2b</i>	Tyr551/Met477/Tyr476	5,0
<i>3a</i>	Met477/Tyr476/Ser543	2,0
<i>3b</i>	Met477/Tyr476/Ser543	5,0

Os modelos de QSAR-4D foram obtidos utilizando uma metodologia que combina algoritmos genéticos (GA) e mínimos quadrados parciais (PLS), implementados no programa 4D-QSAR.

Para evitar um superajuste (*overfitting*) nos dados admite-se que o modelo encontrado possua equações com no máximo quatorze termos. Então os modelos selecionados dos diversos alinhamentos contêm no máximo dez termos.

Na escolha do melhor alinhamento, levaram-se em consideração os parâmetros estatísticos  $R^2$  (coeficiente de determinação) e, principalmente o  $q^2_{\text{ajustado}}$  (coeficiente de correlação linear quadrático ajustado da validação cruzada), que compara a capacidade preditiva do modelo. Para serem

considerados de boa capacidade preditiva, os modelos devem apresentar valores de  $q^2$  maior que 0,5 (KUBINYI, 1993). O valor de  $q^2_{\text{ajustado}}$  é considerado o índice mais adequado para comparar modelos de QSAR, quando estes apresentam números de termos diferentes.

Na Tabela 7, são apresentados os valores de  $R^2$ ,  $q^2$ ,  $q^2_{\text{ajustado}}$ , LSE (do inglês *Least Equares Error*) que é o erro dos mínimos quadrados, e LOF (do inglês *lack-of-fit*) que é o valor do fator de desajuste de Friedman, para os três alinhamentos testados com 2,0 Å e 5,0 Å de aresta. Todos os alinhamentos analisados possuem valores de  $R^2$  maiores que 0,7 e  $q^2_{\text{ajustado}}$  maiores que 0,5.

O alinhamento *1a* foi selecionado para uma análise mais detalhada, já que foi considerado o melhor alinhamento, por possuir o maior valor de  $q^2_{\text{ajustado}}$  (0,67) e de  $R^2$  (0,74) (Tabela 7).

Tabela 7 Valores estatísticos para os diferentes alinhamentos analisados

Alinhamento	Grid cell size (Å)	$R^2$	$q^2$	$q^2_{\text{ajustado}}$	LSE	LOF
<i>1a</i>	2,0	<b>0,74</b>	0,70	<b>0,67</b>	0,56	1,08
<i>1b</i>	5,0	0,71	0,66	0,62	0,62	1,02
<i>2a</i>	2,0	0,70	0,64	0,60	0,65	1,45
<i>2b</i>	5,0	0,72	0,61	0,58	0,61	0,98
<i>3a</i>	2,0	0,71	0,62	0,59	0,62	1,31
<i>3b</i>	5,0	0,71	0,64	0,60	0,63	1,22

A fim de verificar se o Modelo 1 apresenta o mesmo tipo de informação estrutura-atividade, foi calculada a matriz de correlação cruzada entre os valores de ocupação das células (GCODs) (Tabela 8).

Tabela 8 Matriz de correlação cruzada dos GCODs do melhor modelo do alinhamento *1a*

<b>GCOD</b>	<b>0,9,0</b>	<b>2,8,0</b>	<b>4,2,-3</b>	<b>0,4,1</b>	<b>5,1,-3</b>	<b>1,7,2</b>	<b>0,7,1</b>	<b>0,5,1</b>	<b>3,3,-2</b>
	<b>(ar)</b>	<b>(np)</b>	<b>(hba)</b>	<b>(p-)</b>	<b>(np)</b>	<b>(hba)</b>	<b>(hba)</b>	<b>(p-)</b>	<b>(np)</b>
<b>0,9,0</b>	1,00								
<b>(ar)</b>									
<b>2,8,0</b>	0,16	1,00							
<b>(np)</b>									
<b>4,2,-3</b>	-0,03	-0,07	1,00						
<b>(hba)</b>									
<b>0,4,1</b>	-0,05	-0,04	-0,04	1,00					
<b>(p-)</b>									
<b>5,1,-3</b>	-0,01	-0,02	0,20	-0,15	1,00				
<b>(np)</b>									
<b>1,7,2</b>	-0,08	-0,08	0,16	0,28	-0,09	1,00			
<b>(hba)</b>									
<b>0,7,1</b>	-0,10	-0,06	0,10	-0,02	-0,13	-0,14	1,00		
<b>(hba)</b>									
<b>0,5,1</b>	-0,11	0,22	0,04	0,16	0,06	-0,05	-0,02	1,00	
<b>(p-)</b>									
<b>3,3,-2</b>	0,04	-0,11	0,00	0,07	-0,12	0,14	0,03	-0,21	1,00
<b>(np)</b>									

Pode-se verificar que nenhum par de descritor do modelo selecionado está correlacionado, pois apresentam  $r < 0,5$ . Deste modo, os descritores não estão influenciando uns aos outros, e cada descritor contribui de forma única para a construção de modelos robustos de QSAR-4D DR (KUBINYI, 1993).



#### 5.4 Identificação dos *outliers*

Na Tabela 9, estão relacionados os valores de atividade biológica experimentais ( $pIC_{50\text{experimental}}$ ), os valores de atividade biológica calculados ( $pIC_{50\text{predito}}$ ), utilizando a equação do modelo escolhido no alinhamento *1a*, e os respectivos valores residuais.

O Gráfico 1 apresenta os valores de  $pIC_{50}$  experimentais versus  $pIC_{50}$  preditos para os compostos do conjunto de treinamento.

Neste trabalho, foram considerados como *outliers* os compostos com valor residual duas vezes superior ao valor de desvio padrão dos resíduos ( $pIC_{50\text{Exp}} - pIC_{50\text{Pred}}$ ) (CUNHA et al., 2006; HOPFINGER et al., 1997; KUBINYI, 1993). O desvio padrão (DP) para o conjunto de treinamento foi de 0,79. Portanto, o modelo tem apenas um composto do conjunto de treinamento considerado um *outlier*, o composto **31**.

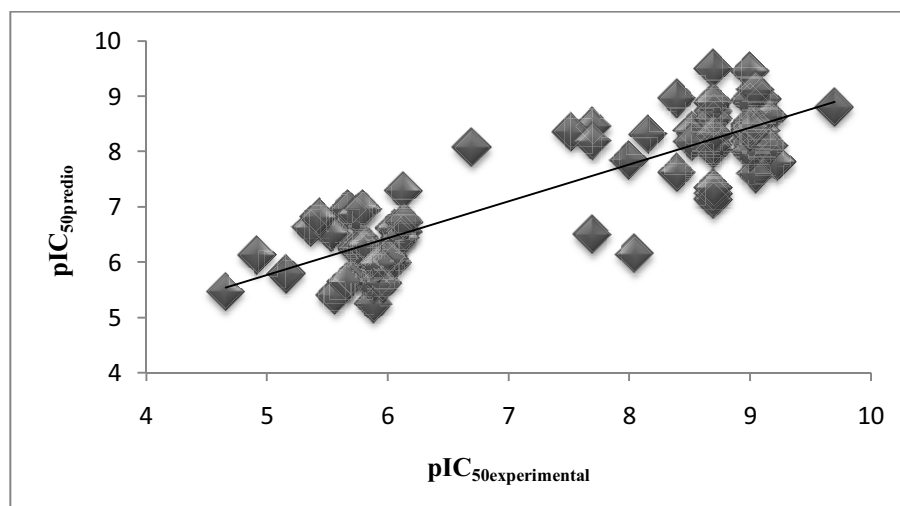


Gráfico 1 Valores de  $pIC_{50}$  experimentais versus preditos do conjunto de treinamento usando o melhor modelo de QSAR-4D-DR do alinhamento *1a*

O conjunto de teste, constituído por 25 compostos (itálico na Tabela 9), por definição, não participa da construção dos modelos e, assim, o cálculo da atividade predita para estes compostos utilizando o Modelo 1, corresponde a um processo de validação externa.

Os compostos *outliers* do grupo de teste foram calculados da mesma forma: os compostos com valor residual duas vezes superior ao valor do DP. O valor do DP encontrado para o grupo de teste foi de 0,64 e apenas a molécula **2** foi considerada um *outlier* no conjunto de teste.

Tabela 9 Valores de pIC<sub>50</sub> experimentais e preditos e valores residuais (pIC<sub>50</sub>predito - pIC<sub>50</sub>experimental) para os compostos do conjunto de treinamento (negrito) e o do conjunto de teste (itálico)

<b>Composto<sup>a</sup></b>	<b>pIC<sub>50</sub> Experimental</b>	<b>pIC<sub>50</sub> Predito</b>	<b>Valores residuais</b>
<b>1</b>	5,74	6,25	0,50
<i>2</i>	8,15	6,38	-1,77
<b>3</b>	6,06	6,61	0,55
<i>4</i>	9,00	9,02	0,02
<i>5</i>	5,85	5,89	0,04
<i>6</i>	6,12	6,17	0,05
<b>7</b>	5,68	6,95	1,27
<b>8</b>	6,13	6,56	0,42
<i>9</i>	6,00	5,33	-0,67
<b>10</b>	5,85	6,28	0,42
<b>11</b>	5,82	6,24	0,41
<b>12</b>	5,57	5,39	-0,17
<b>13</b>	5,74	6,86	1,12
<b>14</b>	9,05	7,60	-1,44
<i>15</i>	8,00	8,86	0,86
<i>16</i>	8,15	8,04	-0,11
<b>17</b>	5,77	6,24	0,47

“Tabela 9, continuação”

<b>Composto<sup>a</sup></b>	<b>pIC<sub>50</sub> Experimental</b>	<b>pIC<sub>50</sub> Predito</b>	<b>Valores residuais</b>
<b>18</b>	4,92	6,13	1,21
<b>19</b>	5,82	6,31	0,49
<b>20</b>	5,89	5,25	-0,64
<b>21</b>	6,06	6,10	0,05
<b>22</b>	5,74	6,61	0,87
<b>23</b>	6,11	6,45	0,34
<b>24</b>	6,11	6,45	0,34
<b>25</b>	6,12	6,63	0,51
<b>26</b>	6,05	6,00	-0,04
<b>27</b>	5,54	6,53	0,99
<b>28</b>	5,70	5,72	0,02
<b>29</b>	8,40	7,63	-0,76
<b>30</b>	8,40	7,63	-0,77
<b>31</b>	8,05	6,15	-1,90
<b>32</b>	5,16	5,79	0,63
<b>33</b>	5,49	5,49	-0,01
<b>34</b>	6,00	6,10	0,10
<b>35</b>	8,70	8,28	-0,42
<b>36</b>	8,52	8,38	-0,14
<b>37</b>	8,00	7,84	-0,15
<b>38</b>	8,40	8,96	0,56
<b>39</b>	7,70	8,46	0,76
<b>40</b>	7,52	8,36	0,84
<b>41</b>	9,00	8,17	-0,83
<b>42</b>	9,00	8,38	-0,62
<b>43</b>	8,70	7,13	-1,57
<b>44</b>	8,70	9,19	0,49
<b>45</b>	9,00	9,45	0,45
<b>46</b>	8,70	8,50	-0,20
<b>47</b>	7,70	6,50	-1,20
<b>48</b>	6,13	6,62	0,49
<b>49</b>	5,89	5,59	-0,30
<b>50</b>	5,66	5,65	-0,01
<b>51</b>	6,13	7,29	1,16
<b>52</b>	5,58	4,89	-0,69
<b>53</b>	9,10	8,94	-0,15
<b>54</b>	5,37	6,63	1,27
<b>55</b>	9,22	7,81	-1,41
<b>56</b>	7,70	8,85	1,15
<b>57</b>	6,70	8,07	1,37
<b>58</b>	9,00	8,49	-0,51

“Tabela 9, conclusão”

<b>Composto<sup>a</sup></b>	<b>pIC<sub>50</sub> Experimental</b>	<b>pIC<sub>50</sub> Predito</b>	<b>Valores residuais</b>
<b>59</b>	9,70	8,81	-0,89
<b>60</b>	7,70	8,19	0,50
<b>61</b>	8,52	8,18	-0,34
<b>62</b>	8,70	9,50	0,80
<b>63</b>	9,00	7,99	-1,00
<b>64</b>	9,15	8,63	-0,53
<b>65</b>	9,22	9,74	0,51
<b>66</b>	9,00	8,93	-0,07
<b>67</b>	8,15	8,32	0,16
<b>68</b>	8,70	8,73	0,04
<b>69</b>	9,10	8,37	-0,72
<b>70</b>	6,14	6,73	0,59
<b>71</b>	8,70	9,56	0,87
<b>72</b>	5,43	6,81	1,37
<b>73</b>	5,80	6,96	1,16
<b>74</b>	6,00	6,12	0,12
<b>75</b>	4,66	5,47	0,82
<b>76</b>	6,01	6,07	0,06
<b>77</b>	5,96	5,62	-0,34
<b>78</b>	5,92	5,98	0,06
<b>79</b>	9,05	8,50	-0,54
<b>80</b>	8,70	7,25	-1,45
<b>81</b>	8,52	8,26	-0,27
<b>82</b>	9,00	8,44	-0,56
<b>83</b>	8,70	8,04	-0,65
<b>84</b>	9,15	9,34	0,18
<b>85</b>	8,70	8,60	-0,10
<b>86</b>	9,15	8,10	-1,05
<b>87</b>	8,52	8,79	0,27
<b>88</b>	8,70	8,90	0,20
<b>89</b>	8,70	8,11	-0,59
<b>90</b>	8,70	8,31	-0,38
<b>91</b>	9,10	9,36	0,26
<b>92</b>	9,05	9,11	0,06
<b>93</b>	9,00	8,36	-0,64
<b>94</b>	8,52	8,73	0,21
<b>95</b>	9,00	8,50	-0,49
<b>96</b>	8,670	7,34	-1,36

a. Compostos *outliers* estão sublinhados.

O composto **31** do conjunto de treinamento, e o composto **2** do conjunto de teste foram calculados como sendo menos potentes do que realmente são. Acredita-se que o motivo de o composto **31** ser um *outlier* pode ser devido à sua estrutura química na porção final da Btk (etanona-piperazinil) ser diferente de todas as outras moléculas do grupo de treinamento, e assim, o modelo não foi capaz de prever sua atividade por não existir células da caixa para descrever esse substituinte nessa posição. Além disso, o inibidor **31** ainda apresenta um grupamento dimetilbenzeno na região do bolsão H3 da Btk, que só está presente (além da molécula **31**) na molécula **20** do conjunto de treinamento. Visto que o inibidor **20** possui um valor menor de pIC<sub>50</sub> do que a molécula **31**, o modelo pode ter se baseado nesta estrutura para prever a atividade da molécula **31**, fazendo com que esta fosse predita como menos potente do que realmente é .

O composto **2** (grupo teste) provavelmente foi considerado um *outlier*, devido à sua ocupação no bolsão H3 por um grupamento metil-tiazol que encontra-se presente apenas na molécula **24** do conjunto de treinamento. Como esta molécula possui um valor de pIC<sub>50</sub> baixo, a comparação dos valores de atividade entre a molécula **2** e a **24** levou a um valor subestimado de atividade predita para a molécula **2**.

#### 5.4.1 Análise do alinhamento *Ia*

O melhor modelo do alinhamento *Ia* (Modelo 1) gerou 9 descritores (GCODs), sendo que sete apresentaram coeficiente positivo, (0,9,0,ar), (4,2,-3,alh), (0,4,1,p-), (5,1,-3,np), (1,7,2,alh), (0,7,1,alh), (0,5,1,p-), e correspondem a interações favoráveis entre os substituintes dos compostos e os resíduos de aminoácidos do sítio ativo da Btk. Assim, substituintes com as características selecionadas (IPes) nessas posições aumentam a potência dos compostos. Os GCODs (2,8,0,np) e (3,3,-2,np) apresentaram coeficientes negativos, correspondendo às interações desfavoráveis entre os substituintes dos compostos

e os resíduos de aminoácidos da Btk. Então, substituintes com características apolares nas regiões desses GCODs, diminuem a potência dos compostos.

A Equação 7 apresenta o Modelo 1 do alinhamento *1a*. Os números entre parênteses indicam as coordenadas cartesianas x, y e z da célula da caixa e as letras entre parênteses indicam o tipo de IPE dos átomos que ocupam a célula.

$$\begin{aligned}
 pIC_{50} = & 6,18 + 5,21 (0,9,0, ar) - 2,92 (2,8,0, np) + 1,98 (4,2, -3, alh) \\
 & + 2,50(0,4,1, p -) + 0,99 (5,1, -3, np) + 2,06 (1,7,2, alh) \\
 & + 1,67 (0,7,1, alh) + 2,48 (0,5,1, p -) - 0,62 (3,3, -2, np) \\
 n = & 71, r^2 = 0,74, q^2 = 0,70, q_{ajustado}^2 = 0,67, LSE = 0,56, LOF = 1,07 (7)
 \end{aligned}$$

Neste trabalho, os GCODs são representados por esferas. As esferas brancas representam os GCODs com coeficientes positivos, e correspondem às regiões na molécula onde a presença de determinados substituintes favorece para aumentar sua potência, e as esferas pretas representam os GCODs com coeficientes negativos, correspondendo às regiões na molécula onde a presença de certo substituinte irá diminuir a potência dos compostos. É importante ressaltar que os GCODs referentes a cada molécula representam a frequência de ocupação das células, considerando todas as conformações de cada um dos compostos obtidos na simulação por dinâmica molecular. Assim, a conformação representada graficamente para cada composto (conformação representativa) não possui, necessariamente, todos os GCODs descritos pelo modelo para a referida molécula.

A Figura 12 ilustra o composto mais ativo (molécula **59**) e os nove descritores obtidos no Modelo1 do alinhamento *1a*.

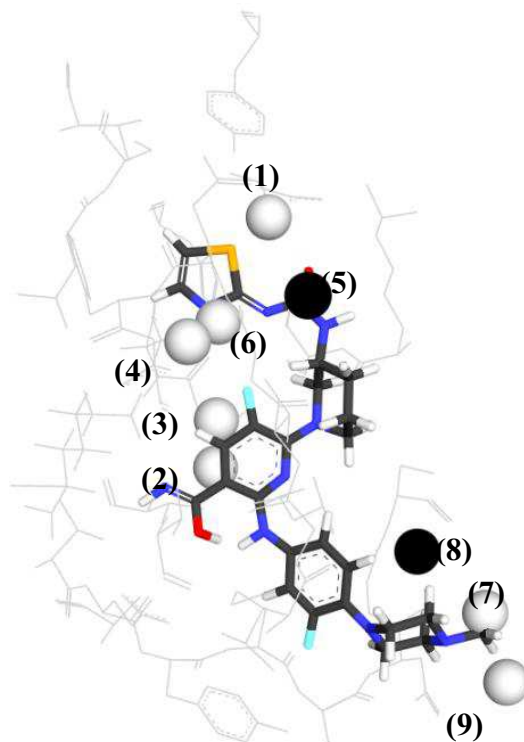


Figura 12 Modelo de QSAR-4D DR representativo do alinhamento *1a*, mostrando o composto **59**

Nota: Os resíduos de aminoácidos da Btk selecionados estão em cinza. A ocupação dos GCODs representados por esferas brancas contribui para o aumento da potência dos compostos e pelas esferas pretas, para a diminuição da potência. Os GCODs descritos são: (1) (0,9,0,ar), (2) (0,4,1,p-), (3) (0,5,1,p-), (4) (1,7,2,alh), (5) (2,8,0,np), (6) (0,7,1,alh), (7) (4,2,-3,alh), (8) (3,3,-2,np), e (9) (5,1,-3,np).

O GCOD-1 (0,9,0,ar) (Figura 13) é o descritor que mais contribui para o aumento na potência dos compostos, já que possui um coeficiente positivo de 5,21 (Eq. 7). Esta célula da caixa mostra alta frequência de ocupação para os compostos **14** e **96** e ocupa o “bolsão de especificidade H3” na Btk. Sabendo-se da importância do resíduo de Tyr551 (presente no bolsão H3) na ativação desta enzima, e observando a ocupação do GCOD-1, que é representado por um IPE do tipo aromático, podemos inferir que este descritor representa uma interação

do tipo  $\pi$ -*stacking* entre a porção aromática de um inibidor, como o composto **14**, e o resíduo de Tyr551. Di Paolo et al. (2011) já haviam mencionado a importância da ocupação do “bolsão H3” na Btk para o aumento de especificidade dos compostos. Podemos ainda observar que, grande parte das moléculas que ocupam esse bolsão (como as moléculas **10** e **43**), não apresenta este descritor (Tabela 10), e os maiores coeficientes são encontrados nas moléculas **14** e **96**, que não só ocupa o bolsão H3, mais também possui uma porção aromática capaz de interagir com o importante resíduo Tyr551. Assim, sugere-se que a interação de um inibidor com o resíduo Tyr551 seja o fator mais relevante para o aumento na potência dos compostos. Deste modo, se um inibidor interagir diretamente com o resíduo de Tyr551, ele será mais eficaz em impedir a transfosforilação da Btk e, assim, irá apresentar uma maior potência, quando comparado com os inibidores que simplesmente ocupam o “bolsão H3”.

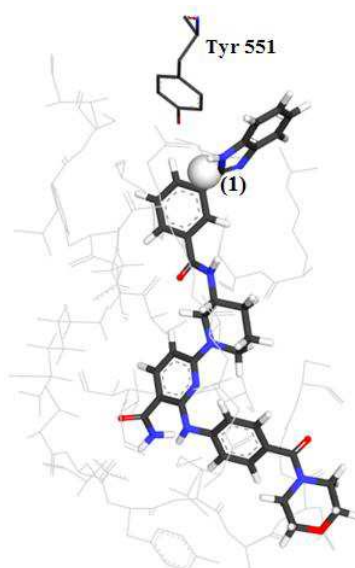


Figura 13 Representação da molécula **14** e do GCOD-1 (esfera branca) obtido do Modelo1 do alinhamento *1a*



Tabela 10 Valores de ocupação percentual do GCOD-1 do Modelo 1 do alinhamento *1a* do QSAR-4D DR para os compostos **10**, **12**, **14**, **43** e **96**

Composto	Porcentagem de ocupação percentual da célula da caixa do GCOD-1 (+) (0,9,0,ar)
<b>10</b>	0,00
<b>14</b>	0,18
<b>43</b>	0,00
<b>59</b>	0,00
<b>96</b>	0,17

Podemos observar ainda, que a molécula mais ativa desta série, **59**, e a molécula mais similar ao inibidor CGI1746, composto **10**, não possuem o descritor que mais contribui para o aumento na potência dos compostos.

O GCOD-2 (0,4,1,p-) (Figura 14A) possui o segundo maior coeficiente positivo, 2,50 (Eq. 7) e, deste modo, também influencia bastante no aumento da potência dos inibidores. Este descritor apresenta um IPE do tipo polar negativo, ou seja, a presença de substituintes eletronegativos nessa posição contribui para aumentar a potência dos compostos. Esta célula localiza-se próximo ao resíduo de aminoácido catalítico Lys430 e ao átomo de flúor nos compostos **53**, **82** e **85**, onde se encontra alta frequência de ocupação para o GCOD-2. Já as moléculas **10** e **14**, por exemplo, não apresentam o átomo de flúor nesta célula da caixa e, portanto, não possuem este descritor que contribui para o aumento na potência dos compostos (Tabela 11).

Tabela 11 Valores de ocupação percentual do GCOD2 do Modelo 1 do alinhamento *1a* do QSAR-4D DR para os compostos **10**, **14**, **53**, **82** e **85**

Composto	Porcentagem de ocupação percentual da célula da caixa do GCOD-2 (+) (0,4,1,p-)
<b>10</b>	0,00
<b>14</b>	0,00
<b>53</b>	0,55
<b>82</b>	0,28
<b>85</b>	0,38

O GCOD-3 (0,5,1,p-) (Figura 14B) também possui um coeficiente positivo, de 2,48, indicando o aumento na potência dos compostos que têm elevada frequência de ocupação para esse descritor. O GCOD-3 apresenta um IPE do tipo polar negativo, ou seja, a presença de substituintes eletronegativos nessa posição contribui para aumentar a potência dos compostos. Assim como o GCOD-2, o GCOD-3 localiza-se próximo ao resíduo de aminoácido catalítico Lys430 e ao átomo de flúor em diversas moléculas, tais como, **38**, **45** e **59**. De forma similar ao GCOD-2, as moléculas **10** e **14**, não apresentam o átomo de flúor e não ocupam esta célula da caixa e, portanto, não possuem o GCOD-3 que favorece um aumento na potência dos inibidores (Tabela 12).

Podemos verificar que os GCOD-2 e GCOD-3 estão localizados muito próximos um do outro e representam o mesmo tipo de interação: entre o resíduo catalítico Lys430 e um átomo eletronegativo de flúor. É possível que um modelo de QSAR-4D DR selecione mais de um descritor que demonstre o mesmo tipo de interação, pois durante o processo de simulação de dinâmica molecular são geradas várias conformações para cada molécula em estudo, e a alta frequência

de ocupação de células vizinhas pode vir a ser registrada e selecionada como um importante fator para a atividade dos compostos.

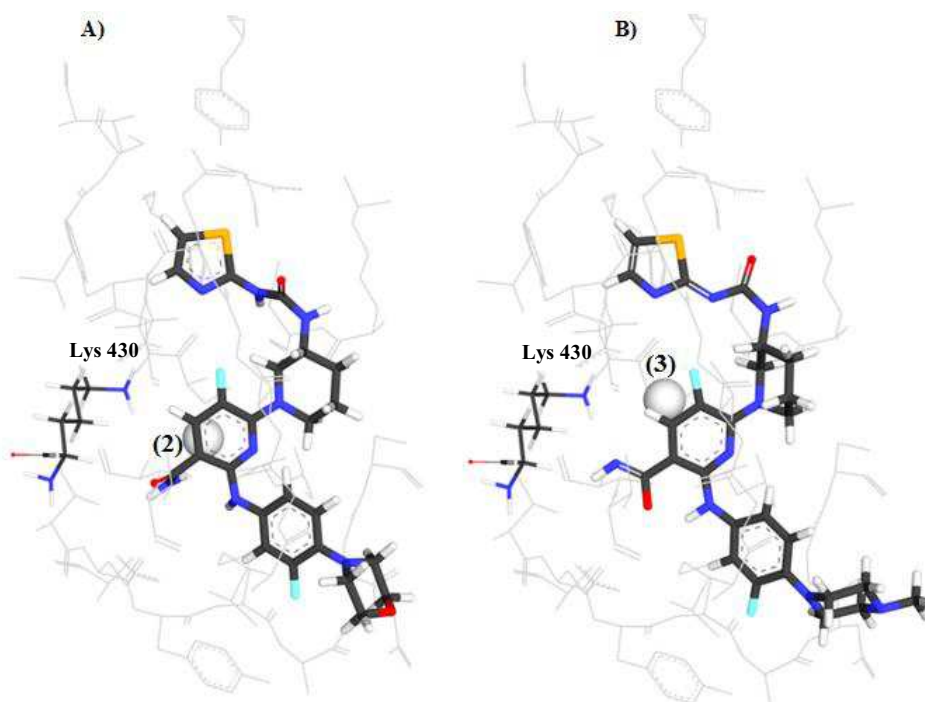


Figura 14 A) Representação da molécula **53** e do GCOD-2 (esfera branca) obtido do Modelo1 do alinhamento *1a*; B) Representação da molécula **59** e do GCOD-3 (esfera branca) obtido do Modelo1 do alinhamento *1a*

Tabela 12 Valores de ocupação percentual do GCOD-3 do Modelo 1 do alinhamento *Ia* do QSAR-4D DR para os compostos **10**, **14**, **38**, **45** e **59**

Composto	Porcentagem de ocupação percentual da célula da caixa do GCOD-3(+) (0,5,1,p-)
<b>10</b>	0,00
<b>14</b>	0,00
<b>38</b>	0,50
<b>45</b>	0,64
<b>59</b>	0,45

O GCOD-4 (1,7,2,alh), Figura 15A, possui um coeficiente positivo de 2,06 (Eq.7) com IPE do tipo acceptor de ligação de hidrogênio, indicando o aumento da potência dos compostos que têm elevada frequência de ocupação para esse descritor. Já o GCOD-5 (2,8,0,np), Figura 15B, possui um coeficiente de -2,92. Visto que o GCOD-5 possui um coeficiente com sinal negativo, a ocupação de átomos apolares nesta célula da caixa reduz a potência dos compostos. Os GCOD-4 e GCOD-5 situam-se próximos do resíduo de aminoácido Gln412. O descritor 4 nos aponta a importância de um átomo que possua um par de elétrons não ligado próximo do resíduo Gln412 capaz, assim, de realizar uma ligação de hidrogênio. As moléculas **30**, **35** e **53** apresentam ocupação nesta célula da caixa e possuem um átomo de oxigênio, acceptor de ligação de hidrogênio, próximo ao resíduo de Gln412. Por outro lado, se ao invés de um átomo acceptor de ligação de hidrogênio houver um átomo apolar nesta localização, descrito pelo GCOD-5, como nas moléculas **49** e **50**, a potência dos compostos irá diminuir como resultado de uma interação por ligação de hidrogênio que o inibidor deixará de realizar (Tabela 13).

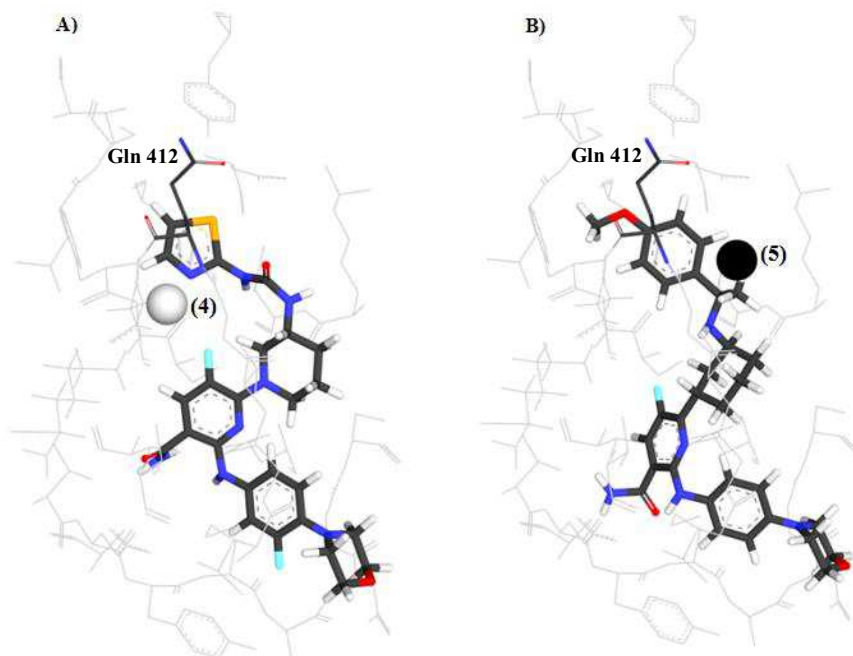


Figura 15 A) Representação da molécula **53** e do GCOD-4 (esfera branca) obtido do Modelo1 do alinhamento *1a* ; B) Representação da molécula **49** e do GCOD-5 (esfera preta) obtido do Modelo1 do alinhamento *1a*

Tabela 13 Valores de ocupação percentual dos GCODs 4 e 5 do Modelo 1 do alinhamento *1a* do QSAR-4D DR para os compostos **30**, **35**, **49**, **50** e **53**

Composto	Porcentagem de ocupação percentual da célula da caixa do GCOD-4(+) (1,7,2,alh)	Porcentagem de ocupação percentual da célula da caixa do GCOD-5(-) (2,8,0,np)
<b>30</b>	0,40	0,00
<b>35</b>	0,84	0,00
<b>49</b>	0,00	0,33
<b>50</b>	0,00	1,16
<b>53</b>	0,95	0,00

O GCOD-6 (0,7,1,alh) (Figura 16) possui um coeficiente positivo de 1,67 e um IPE alh. Então, a ocupação de átomos doadores de ligação de hidrogênio nesta célula da caixa leva a um aumento na potência dos compostos. O GCOD-6 situa-se próximo ao resíduo de aminoácido Asp539. Logo, átomos que possuem um par de elétrons livres neste GCOD podem interagir por ligação de hidrogênio com esse aminoácido. Algumas das moléculas que apresentam ocupação para este descritor e podem interagir com o resíduo de Asp539 por ligação de hidrogênio são: **17**, **59** e **92**. Em contraste, as moléculas **10** e **14** não apresentam ocupação de um átomo aceptor de ligação de hidrogênio nessa célula da caixa e não possuem o GCOD-6 (Tabela 14).

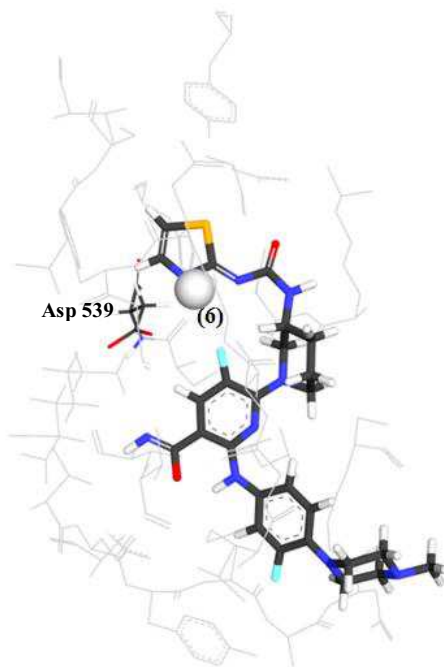


Figura 16 Representação da molécula **59** e do GCOD-6 (esfera branca) obtido do Modelo1 do alinhamento *1a*

Tabela 14 Valores de ocupação percentual do GCOD-6 do Modelo 1 do alinhamento *1a* do QSAR-4D DR para os compostos **10**, **14**, **17**, **59** e **92**

Composto	Porcentagem de ocupação percentual da célula da caixa do GCOD-6 (+) (0,7,1,alh)
<b>10</b>	0,00
<b>14</b>	0,00
<b>17</b>	0,72
<b>59</b>	0,72
<b>92</b>	1,69

O GCOD-7 (4,2,-3,alh), Figura 17A, possui um coeficiente positivo de 1,98 e um IPE do tipo acceptor de ligação de hidrogênio. Desta forma, a ocupação de átomos que possuam um par de elétrons livres nesta célula da caixa leva a um aumento na potência dos compostos. O GCOD-7 encontra-se próximo aos resíduos de aminoácido Asn479 e Tyr476, logo, a presença de átomos aceptores de ligação de hidrogênio nessa célula da caixa favorece o aumento na potência dos compostos já que os mesmos podem realizar uma interação por ligação de hidrogênio com esses resíduos de aminoácido. Algumas das moléculas que apresentam boa ocupação neste descritor são: **40**, **42** e **55** (Tabela 15). Já o GCOD-8 (3,3,-2,np) (Figura 17B) possui um coeficiente de -0,62, e contribui para a diminuição na potência dos inibidores. O GCOD-8 também se situa próximo dos resíduos de aminoácido Asn479 e Tyr476. Este GCOD nos indica que a ocupação desta célula da caixa por átomos apolares resulta em um decréscimo na potência dos compostos. Então, os inibidores que possuem alta frequência de ocupação para o GCOD-8 não podem interagir com os resíduos de Asn479 e Tyr476 e, conseqüentemente, a potência desses inibidores deve ser menor. Alguns compostos que apresentam alta frequência de ocupação para o GCOD-8 são: **17** e **61** (Tabela 15).



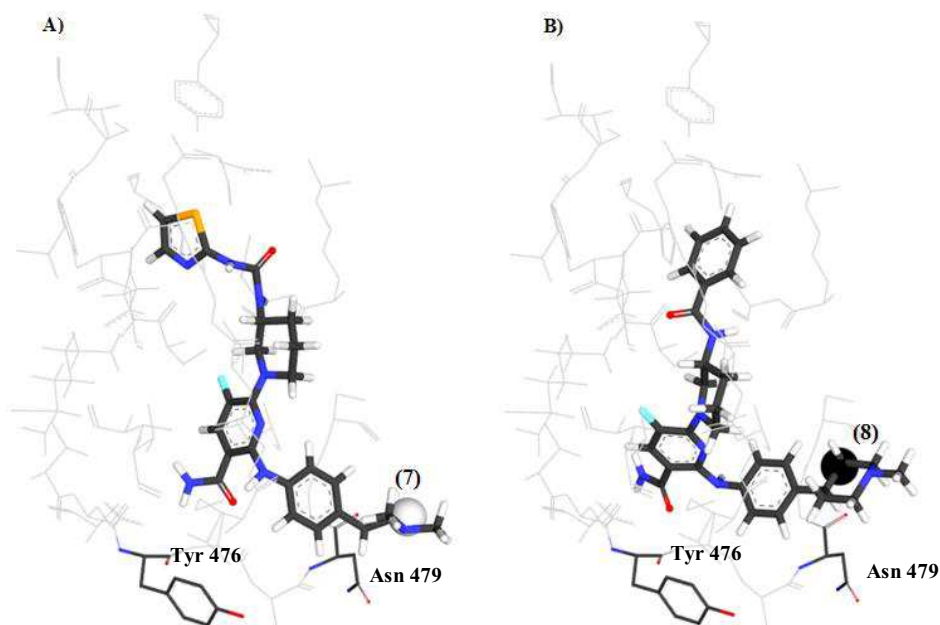


Figura 17 A) Representação da molécula **55** e do GCOD-7 (esfera branca) obtido do Modelo1 do alinhamento *1a*. B) Representação da molécula **61** e do GCOD-8 (esfera preta) obtido do Modelo1 do alinhamento *1a*

Tabela 15 Valores de ocupação percentual dos GCODs 7 e 8 do Modelo 1 do alinhamento *1a* do QSAR-4D DR para os compostos **17**, **40**, **42**, **55** e **61**

Composto	Porcentagem de ocupação percentual da célula da caixa do GCOD-7 (+) (4,2,-3,alh)	Porcentagem de ocupação percentual da célula da caixa do GCOD-8 (-) (3,3,-2,np)
<b>17</b>	0,00	1,82
<b>40</b>	0,46	0,14
<b>42</b>	0,45	0,16
<b>55</b>	0,32	0,07
<b>61</b>	0,05	1,11

Finalmente, o GCOD-9 (5,1,-3,np) (Figura 18) também possui um coeficiente positivo, de 0,99, indicando aumento na potência dos compostos que têm elevada frequência de ocupação para esse descritor. O GCOD-9 apresenta um IPE do tipo não polar e, deste modo, a presença de substituintes apolares nessa célula da caixa, que ocupa a porção final dos inibidores, contribui para aumentar a potência dos compostos. As moléculas **36**, **55**, **62** e **63** (Tabela 16) possuem alta frequência de ocupação para esse descritor, que está localizado próximo dos resíduos de aminoácidos apolares Leu408 e Ala478, favorecendo uma interação entre estes resíduos e a porção apolar das moléculas que possuem alta frequência de ocupação para o GCOD-9. As moléculas **28** e **54** não possuem esta porção apolar nesta célula da caixa e, assim, não possui o GCOD-9, que favorece um discreto aumento na potência dos inibidores.

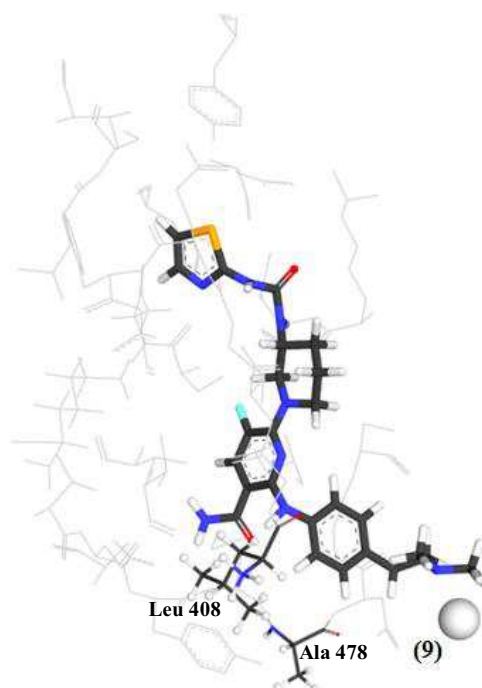


Figura 18 Representação da molécula **55** e do GCOD-9 (esfera branca) obtido do Modelo1 do alinhamento *Ia*

Tabela 16 Valores de ocupação percentual do GCOD-9 do Modelo 1 do alinhamento *1a* do QSAR-4D DR para os compostos **28, 36, 54, 55, 62 e 63**

<b>Composto</b>	<b>Porcentagem de ocupação percentual da célula da caixa do GCOD-6 (+) (0,7,1,alh)</b>
<b>28</b>	0,00
<b>36</b>	0,60
<b>54</b>	0,00
<b>55</b>	0,92
<b>62</b>	1,07
<b>63</b>	0,73

### **5.5 Planejamento de potenciais inibidores da Btk**

Ao analisar os resultados do ancoramento molecular, observamos que muitas informações obtidas estão em concordância com os resultados da metodologia de QSAR-4D DR. As duas metodologias ressaltam que a característica mais importante que um inibidor potente deve possuir é ocupar o bolsão de especificidade H3 na Btk. Além disso, o resultado do QSAR-4D DR ainda complementa indicando a importância da interação de um inibidor potente com o resíduo de aminoácido Tyr551.

A segunda característica mais relevante para o aumento na potência dos fármacos é a presença do elemento flúor. Pela metodologia de QSAR, esta informação é demonstrada como uma interação eletrostática. Também observamos a presença do resíduo de aminoácido catalítico Lys430 (que possui características básicas) muito próximo ao átomo de flúor. Já pela metodologia de ancoramento molecular, esta informação é demonstrada como uma interação

intramolecular favorável nas moléculas que possuem este átomo eletronegativo, também na mesma posição, próximo ao resíduo de aminoácido Lys430.

Ao observar o composto mais potente desta série (a molécula **59**) em sua conformação de ancoramento, Figura 19, notamos que o átomo de flúor encontra-se muito próximo do resíduo catalítico Lys430, com uma distância de apenas 1,94 Å, e ainda, está em uma orientação favorável para a ocorrência de uma interação por ligação de hidrogênio com esse resíduo de aminoácido. É importante notar ainda que, quando visualizamos a superfície eletrostática da molécula **59**, a metodologia de *docking* não reconhece que o átomo eletronegativo de flúor possua um potencial eletrostático negativo (Figura 20). Deste modo, é possível que as metodologias utilizadas neste trabalho não tenham identificado este elemento como um possível acceptor de ligação de hidrogênio, que poderia estar realizando uma interação com o resíduo Lys430, importante na catálise enzimática da Btk.

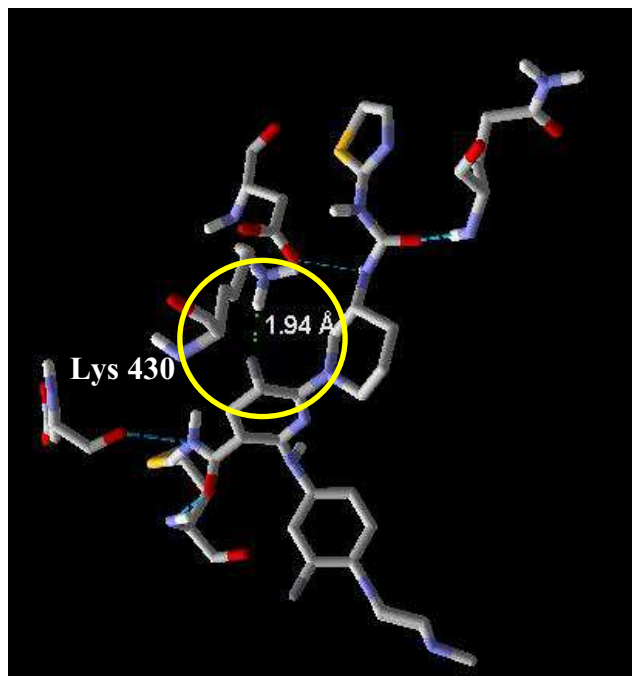


Figura 19 Ilustração da molécula **59** em sua conformação de ancoramento. As ligações de hidrogênio são demonstradas em azul e a distância do átomo de flúor (em cinza) e do resíduo de Lys430 é demonstrada em verde

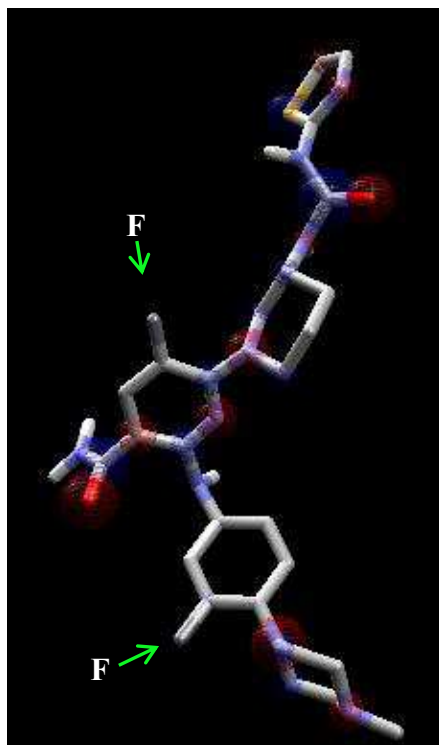


Figura 20 Ilustração da molécula **59** em sua conformação de ancoramento e das representações dos potenciais eletrostático positivo (em azul) e negativo (em vermelho). Os átomos de flúor (cinza) estão indicados em verde

Finalmente, a maioria dos descritores gerados no QSAR está direta ou indiretamente relacionada com as interações por ligação de hidrogênio. Os GCODs 4, 6 e 7 possuem um IPE do tipo alh e, por isso, representam diretamente essa interação. Já os GCODs 5 e 8 estão associados indiretamente, pois representam características prejudiciais, quando os ligantes deixam de possuir alta frequência de ocupação por átomos que poderiam realizar essa interação. E por fim, é possível que os GCODs 2 e 3 também estejam

relacionados com esse tipo de interação, se confirmada uma interação por ligação de hidrogênio do resíduo Lys430 com o átomo de flúor.

Já os estudos de ancoramento molecular não detectaram de forma precisa a importância do fator ligação de hidrogênio, no aumento da potência dos inibidores.

A partir dos resultados obtidos nos estudos de ancoramento molecular e de QSAR, foram propostas novas estruturas para síntese e teste de atividade (Figura 21). Para estas duas moléculas ainda não sintetizadas, foi feito um ensaio virtual de atividade, aplicando a Equação 7 do Modelo 1 de QSAR-4D DR e foram obtidos os valores de energia de ancoramento pelo *docking*.

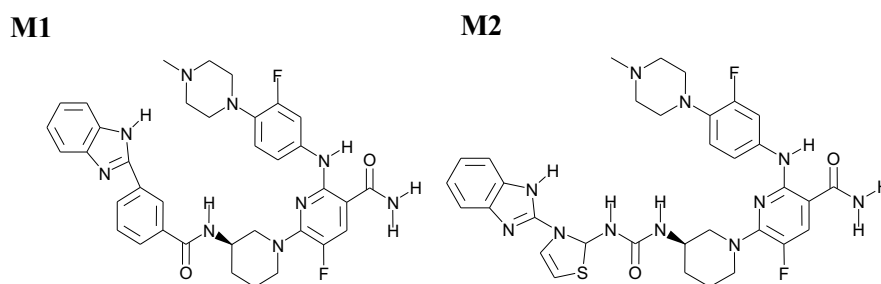


Figura 21 Ilustração das moléculas propostas: **M1** e **M2**

A molécula **M1** foi proposta com base na estrutura **14**, sendo adicionados dois átomos de flúor e substituída a porção final da molécula (piperazinil). Seus valores de potência inibitória preditos foram comparados com os valores de potência inibitória predita do inibidor **14**. A molécula **M2** foi baseada na estrutura **59**, e foi incorporada a porção aromática que ocupa o bolsão H3 na Btk. Seus resultados foram comparados aos valores de potência inibitória do inibidor **59**. Os valores de potência inibitória preditos ( $pIC_{50predito}$ ) para as moléculas **14**, **59**, **M1** e **M2**, encontram-se na Tabela 17:

Tabela 17 Valores de potência inibitória preditos ( $pIC_{50predito}$ ) (M) para os inibidores **14**, **59**, **M1** e **M2**

	<b>14</b>	<b>59</b>	<b>M1</b>	<b>M2</b>
$pIC_{50predito}$	7,60	8,81	8,64	9,06

Os resultados dos valores de  $E_{Est}$ ,  $E_{Intra}$ ,  $E_{LigH}$ ,  $E_{Elet}$  e  $E_{EletLong}$  ( $kcal.mol^{-1}$ ) para os inibidores **14**, **59**, **M1** e **M2** e os valores de  $IC_{50}$  preditos a partir da Eq.6, são apresentados na Tabela 18:

Tabela 18 Valores de  $E_{Est}$ ,  $E_{Intra}$ ,  $E_{LigH}$ ,  $E_{Elet}$  e  $E_{EletLong}$  ( $kcal.mol^{-1}$ ) e valores de  $IC_{50}$  predito para os inibidores **14**, **59**, **M1** e **M2**

	<b>14</b>	<b>59</b>	<b>M1</b>	<b>M2</b>
$E_{Est}$	-190,75	-165,92	-188,87	-189,90
$E_{Intra}$	-22,35	-46,59	-27,51	-27,31
$E_{LigH}$	-6,31	-7,92	-4,06	-5,01
$E_{Elet}$	-4,34	-4,23	-3,39	-4,23
$E_{EletLong}$	-3,12	-1,47	-4,23	-1,07
$pIC_{50}$ predito	10.01	10.31	10.00	10.87

Os resultados de QSAR-4D DR e de ancoramento molecular indicam que o inibidor proposto **M2** deverá apresentar uma potência inibitória maior do que a molécula **59**, que é a mais ativa dentre as moléculas da série avaliada. Apesar de a molécula **M1** ter apresentado valores altos de potência inibitória, a molécula **M2** foi a que apresentou a maior potência inibitória, quando comparada com as moléculas **14**, **59** e **M1**, provavelmente devido à molécula **M2** estar relacionada ao composto mais ativo (molécula **59**) acrescido da informação obtida por meio do descritor positivo que mais contribui para o



aumento na potência dos compostos, segundo a metodologia de QSAR-4D DR.  
Assim, esta molécula pode ser considerada um promissor inibidor da Btk.

## 6 CONCLUSÕES

A análise dos resultados dos cálculos de energia por ancoramento molecular mostrou que existe uma boa correlação entre os resultados teóricos e experimentais para os 96 inibidores estudados com  $R^2$  de 0,80. Os fatores que contribuem para o aumento na potência dos inibidores são a energia de impedimento espacial entre a proteína e o ligante ( $E_{Est}$ ) e a energia intramolecular do ligante ( $E_{Intra}$ ).

No estudo de QSAR-4D DR, os modelos gerados fornecem informações importantes que podem auxiliar no planejamento de fármacos. O melhor modelo, Modelo 1 do alinhamento *1a*, apresentou um  $q^2_{ajustado}$  de 0,67, confirmando sua capacidade preditiva. Os resultados de QSAR-4D confirmam e ampliam os resultados do ancoramento molecular.

As duas metodologias preveem a importância da ocupação do bolsão H3 para o aumento na potência dos compostos e a importância do átomo de flúor próximo do resíduo catalítico Lys430. Partindo dessas observações, foram propostas duas moléculas, sendo que a molécula **M2** apresentou uma melhor potência predita do que o composto **59**, que é o composto mais ativo da série estudada. Os resultados obtidos são passíveis de serem aplicados na busca por novos fármacos inibidores da Btk.

## REFERÊNCIAS

- 4D-QSAR user's manual v.1.0. Lake Forest: The Chem21, 1997. 142 p.
- ABBAS, A. K.; LICHTMAN, A. H.; PILLAI, S. **Imunologia celular e molecular**. 6. ed. Rio de Janeiro: Elsevier, 2008. 564 p.
- ALLINGER, N. L. Conformational-analysis 130 MM2-Hydrocarbon force-field utilizing v1 and v2 torsional terms. **Journal of the American Chemical Society**, Easton, v. 99, p. 8127-8134, 1977.
- ALLINGER, N. L.; BURKERT, U. Pitfall in the use of the torsion angle driving method for the calculation of conformation interconversions. **Journal of Computational Chemistry**, New York, v. 3, n. 1, p. 40-46, 1982.
- ANDRADE, C. H. et al. 4D-QSAR: perspectives in drug design. **Molecules**, Basel, v. 15, n. 5, p. 3281-3294, May 2010.
- BERMAN, H. M. et al. The protein data bank. **Nucleic Acids Research**, Oxford, v. 28, n. 1, p. 235-242, 2000.
- BROWN, J. R. Ibrutinib (PCI-32765), the first BTK (Bruton's Tyrosine Kinase) inhibitor in clinical trials. **Current Hematologic Malignancy Reports**, New York, v. 8, n. 1, p. 1-6, Mar. 2013.
- CAETANO, M. S. **Estudos computacionais entre análogos do glifosato e as enzimas EPSP sintase nativa e mutante Gly96Ala**. 2007. 53 p. Monografia (Graduação em Química) - Universidade Federal de Lavras, Lavras, 2007.
- CHEETHAM, G. M. T. Novel protein kinases and molecular mechanisms of autoinhibition. **Current Opinion in Structural Biology**, London, v. 14, n. 6, p. 700-705, Dec. 2004.
- CRAMER, R. D. et al. Comparative molecular field analysis (CoMFA) 1: effect of shape on binding of steroids to carrier proteins. **Journal of the American Chemical Society**, Easton, v. 110, p. 5959-5967, 1988.
- CUNHA, E. F. F. da et al. Computer-assisted analysis of the interactions of macrocyclic inhibitors with wild type and mutant D168A hepatitis C virus NS3 serine protease. **Letters in Drug Design & Discovery**, Cambridge, v. 3, n. 1, p. 17-28, Feb. 2006.

DELUCCA, G. V. et al. **New substituted nicotinamide compounds are Bruton's tyrosine kinase inhibitors, useful for the treatment or prophylaxis of e.g. ulcerative colitis, asthma, Graves' disease, rheumatoid arthritis, psoriasis, osteoarthritis and osteoporosis.** US n. 144647, 16 dez. 2010. Bristol: Myers Squibb, 2010.

DEWAR, M. J. S. et al. The development and use of quantum-mechanical molecular-models 76: AM1-A new general-purpose quantum-mechanical molecular-model. **Journal of the American Chemical Society**, Easton, v. 107, p. 3902-3909, 1985.

DI PAOLO, J. A. et al. Specific Btk inhibition suppresses B cell- and myeloid cell-mediated arthritis. **Nature Chemical Biology**, New York, v. 7, n. 1, p. 41-50, Jan. 2011.

DOHERTY, D. **MOLSIM: molecular mechanics and dynamics simulation software: userMs guide.** Version 3.2. Lake Forest: The Chem21, 1997. Software.

FIUME, G. et al. Computational analysis and in vivo validation of a microRNA encoded by the IBTK gene, a regulator of B-lymphocytes differentiation and survival. **Computational Biology and Chemistry**, Oxford, v. 33, n. 6, p. 434-439, Dec. 2009.

GUO, S. L.; WAHL, M. I.; WITTE, O. N. Mutational analysis of the SH2-kinase linker region of Bruton's tyrosine kinase defines alternative modes of regulation for cytoplasmic tyrosine kinase families. **International Immunology**, Oxford, v. 18, n. 1, p. 79-87, Jan. 2006.

HANTSCH, O. et al. The Btk tyrosine kinase is a major target of the Bcr-Abl inhibitor dasatinib. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 104, n. 33, p. 13283-13288, 2007.

HENDRIKS, R. W. New Btk inhibitor holds promise. **Nature Chemical Biology**, New York, v. 7, p. 1-2, Jan. 2011.

HILLISCH, A. et al. Dissecting physiological roles of estrogen receptor alpha and beta with potent selective ligands from structure-based design. **Molecular Endocrinology**, Bethesda, v. 18, n. 7, p. 1599-1609, July 2004.

HONIGBERG, L. A. et al. The Bruton tyrosine kinase inhibitor PCI-32765 blocks B-cell activation and is efficacious in models of autoimmune disease and B-cell malignancy. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 107, n. 29, p. 13075-13080, July 2010.

HOPFINGER, A. J. et al. Construction of 3D-QSAR models using the 4D-QSAR analysis formalism. **Journal of the American Chemical Society**, Easton, v. 119, n. 43, p. 10509-10524, Oct. 1997.

HOPFINGER, A. J.; PEARLSTEIN, R. A. Molecular mechanics force-field parameterization procedures. **Journal of Computational Chemistry**, New York, v. 5, n. 5, p. 486-499, 1984.

HYPERCHEM. **Hyperchem program release 7.03 for Windows**. Hypercube, 2002. Software.

IWATA, Y. et al. Conformational analysis and docking study of potent factor XIIIa inhibitors having a cyclopropenone ring. **Journal of Molecular Graphics & Modelling**, New York, v. 18, n. 6, p. 591-599, Dec. 2000.

JACOB, C. M. A. Agamaglobulinemia. **Revista Brasileira de Alergia e Imunopatologia**, São Paulo, v. 28, n. 6, p. 267-272, dez. 2005.

KARKI, R. G.; KULKARNI, V. M. Three-dimensional quantitative structure-activity relationship (3D-QSAR) of 3-aryloxazolidin-2-one antibacterials. **Bioorganic & Medicinal Chemistry**, New York, v. 9, n. 12, p. 3153-3160, Dec. 2001.

KHAN, A. K. R. et al. Comparative QSTR study of saturated alcohols based on topological, constitutional, geometrical, and getaway descriptors. **Medicinal Chemistry Research**, Cambridge, v. 18, n. 9, p. 770-781, Dec. 2009.

KONDAPALLI, L.; SOLTANI, K.; LACOUTURE, M. E. The promise of molecular targeted therapies: protein kinase inhibitors in the treatment of cutaneous malignancies. **Journal of the American Academy of Dermatology**, Saint Louis, v. 53, n. 2, p. 291-302, Aug. 2005.

KRAUSE, D. S.; ETTEN, R. A. van. Tyrosine kinases as targets for cancer therapy. **New England Journal of Medicine**, Waltham, v. 353, n. 2, p. 172-187, July 2005.

KUBINYI, H. QSAR and 3D QSAR in drug design .1. methodology. **Drug Discovery Today**, Kidlington, v. 2, n. 11, p. 457-467, Nov. 1997.

KUBINYI, H. **QSAR**: hanch analysis and related approaches: methods and principles in medicinal chemistry. Weinheim: VHC, 1993. 240 p.

KUGLSTATTER, A. et al. Insights into the conformational flexibility of Bruton's tyrosine kinase from multiple ligand complex structures. **Protein Science**, Cold Springer Harbor, v. 20, n. 2, p. 428-436, Feb. 2011.

KUZ'MIN, V. E.; ARTEMENKO, A. G.; MURATOV, E. N. Hierarchical QSAR technology based on the Simplex representation of molecular structure. **Journal of Computer-Aided Molecular Design**, Dordrecht, v. 22, n. 6/7, p. 403-421, June 2008.

LA MOTTA, C. et al. Computational studies of epidermal growth factor receptor: docking reliability, three-dimensional quantitative structure-activity relationship analysis, and virtual screening studies. **Journal of Medicinal Chemistry**, Washington, v. 52, n. 4, p. 964-975, Feb. 2009.

LILL, M. A. Multi-dimensional QSAR in drug discovery. **Drug Discovery Today**, Kidlington, v. 12, n. 23/24, p. 1013-1017, Dec. 2007.

LIVINGSTONE, D. J. Neural-Network Studies: 1., comparison of overtraining. **Journal of Chemical Information and Computer Sciences**, Washington, v. 35, n. 5, p. 826-833, 1995.

LOU, Y. et al. Bruton's tyrosine kinase inhibitors: approaches to potent and selective inhibition, preclinical and clinical evaluation for inflammatory diseases and B cell malignancies. **Journal of Medicinal Chemistry**, Washington, v. 55, n. 10, p. 4539-4550, May 2012.

MAO, C.; ZHOU, M.; UCKUN, F. M. Crystal structure of Bruton's tyrosine kinase domain suggests a novel pathway for activation and provides insights into the molecular basis of X-linked agammaglobulinemia. **Journal of Biological Chemistry**, Bethesda, v. 276, n. 44, p. 41435-41443, Nov. 2001.

MARCOTTE, D. J. et al. Structures of human Bruton's tyrosine kinase in active and inactive conformations suggest a mechanism of activation for TEC family kinases. **Protein Science**, Cold Springer Harbor, v. 19, n. 3, p. 429-439, Mar. 2010.

MOHAMED, A. J. et al. Bruton's tyrosine kinase (Btk): function, regulation, and transformation with special emphasis on the PH domain. **Immunological Reviews**, Copenhagen, v. 228, p. 58-73, Mar. 2009.

OLIVEIRA, A. A.; RAMALHO, T. C.; CUNHA, E. F. F. da. QSAR study of androstenedione analogs as aromatase inhibitors. **Letters in Drug Design & Discovery**, Cambridge, v. 6, n. 8, p. 554-562, Dec. 2009.

OLIVEIRA, R. R. **Estudo do gene btk (BRUTON'S TYROSINE KINASE) em pacientes com agamaglobulinemia congênita**. 2008. 114 p. Tese (Doutorado em Biotecnologia) - Universidade de São Paulo, São Paulo, 2008.

OPREA, T. I. On the information content of 2D and 3D descriptors for QSAR. **Journal of the Brazilian Chemical Society**, São Paulo, v. 13, n. 6, p. 811-815, Nov./Dec. 2002.

PAN, D. H.; TSENG, Y. F.; HOPFINGER, A. J. Quantitative structure-based design: formalism and application of receptor-dependent RD-4D-QSAR analysis to a set of glucose analogue inhibitors of glycogen phosphorylase. **Journal of Chemical Information and Computer Sciences**, Washington, v. 43, n. 5, p. 1591-1607, Sept./Oct. 2003.

PAN, Z. et al. Discovery of selective irreversible inhibitors for bruton's tyrosine kinase. **ChemMedChem**, Weinheim, v. 2, n. 1, p. 58-61, Jan. 2007.

PITA, S. S. D. et al. Receptor-dependent 4D-QSAR analysis of peptidemimetic inhibitors of trypanosoma cruzi trypanothione reductase with receptor-based alignment. **Chemical Biology & Drug Design**, Oxford, v. 79, n. 5, p. 740-748, May 2012.

ROGERS, D.; HOPFINGER, A. J. Application of genetic function approximation to quantitative structure-activity-relationships and quantitative structure-property relationships. **Journal of Chemical Information and Computer Sciences**, Washington, v. 34, n. 4, p. 854-866, July/Aug. 1994.

ROMEIRO, N. C. et al. Construction of 4D-QSAR models for use in the design of novel p38-MAPK inhibitors. **Journal of Computer-Aided Molecular Design**, Dordrecht, v. 19, n. 6, p. 385-400, 2005.

SANTOS, O. A.; HOPFINGER, A. J. Structure-based QSAR analysis of a set of 4-hydroxy-5,6-dihydropyrones as inhibitors of HIV-1 protease: an application of the receptor-dependent (RD) 4D-QSAR formalism. **Journal of Chemical Information and Modeling**, Washington, v. 46, n. 1, p. 345-354, Jan./Feb. 2006.

SELASSIE, C. D. History of quantitative structure-activity relationship. In: ABRAHAM, D. J. (Ed.). **Burger's medicinal chemistry and drug discovery**. 6<sup>th</sup> ed. New York: J. Wiley, 2003. v. 1, p. 1-48.

SILVEIRA, R. L. V. de A. Fitotoxicidade de glifosato em *Eucalyptus*. **Addubare**, Piracicaba, v. 2, n. 9, p. 4-7, out./dez. 2003.

SUTHERLAND, J. J.; O'BRIEN, L. A.; WEAVER, D. F. A comparison of methods for modeling quantitative structure-activity relationships. **Journal of Medicinal Chemistry**, Washington, v. 47, n. 22, p. 5541-5554, Oct. 2004.

TAVARES, L. C. QSAR: a abordagem de Hansch. **Química Nova**, São Paulo, v. 27, n. 4, p. 631-639, jul./ago. 2004.

THOMAS, J. D. et al. Colocalization of X-linked agammaglobulinemia and X-linked immunodeficiency genes. **Science**, New York, v. 261, n. 5119, p. 355-358, July 1993.

THOMSEN, R.; CHRISTENSEN, M. H. MolDock: a new technique for high-accuracy molecular docking. **Journal of Medicinal Chemistry**, Washington, v. 49, n. 11, p. 3315-3321, June 2006.

TSUKADA, S. et al. Deficient expression of a B-cell cytoplasmic tyrosine kinase in human X-linked agammaglobulinemia. **Cell**, Cambridge, v. 72, n. 2, p. 279-290, Jan. 1993.

VARGAS, L. et al. Inhibitors of BTK and ITK: state of the new drugs for cancer, autoimmunity and inflammatory diseases. **Scandinavian Journal of Immunology**, Oxford, v. 78, n. 2, p. 130-139, Aug. 2013.

VEDANI, A. Novel ligands for the chemokine receptor-3 (CCR3): a receptor-modeling study based on 5D-QSAR. **Journal of Medicinal Chemistry**, Washington, v. 48, n. 5, p. 1515-1527, Mar. 2005.



VEDANI, A. et al. Predicting the toxic potential of drugs and chemicals in silico: a model for the peroxisome proliferator-activated receptor-gamma (PPAR gamma). **Toxicology Letters**, Amsterdam, v. 173, n. 1, p. 17-23, Aug. 2007.

VETRIE, D. et al. The gene involved in X-linked agammaglobulinemia is a member of the src family of protein-tyrosine kinases. **Nature**, London, v. 361, n. 6409, p. 226-233, Jan. 1993.

WEINER, S. J. et al. An all atom force-field for simulations of proteins and nucleic acids. **Journal of Computational Chemistry**, New York, v. 7, p. 230-252, 1986.

YANMAZ, E. et al. 4D-QSAR analysis and pharmacophore modeling: electron conformational-genetic algorithm approach for penicillins. **Bioorganic & Medicinal Chemistry**, New York, v. 19, n. 7, p. 2199-2210, Apr. 2011.

YILMAZ, H. et al. 4D-QSAR study of p56(lck) protein tyrosine kinase inhibitory activity of flavonoid derivatives using MCET method. **Bulletin of the Korean Chemical Society**, Seoul, v. 32, n. 12, p. 4352-4360, Dec. 2011.

## APÊNDICE

**APÊNDICE A** - Tabela com os coeficientes dos GCODs 1 (D-1), 2 (D-2), 3 (D-3), 4 (D-4), 5 (D-5), 6 (D-6), 7 (D-7), 8 (D-8) e 9 (D-9) para as 96 moléculas estudadas

<b>Mol.</b>	<b>D-1</b>	<b>D-2</b>	<b>D-3</b>	<b>D-4</b>	<b>D-5</b>	<b>D-6</b>	<b>D-7</b>	<b>D-8</b>	<b>D-9</b>
1	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,17	0,18
2	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,07	0,28
3	0,00	0,04	0,14	0,00	0,00	0,00	0,00	0,03	0,00
4	0,55	0,00	0,00	0,00	0,00	0,00	0,01	0,11	0,00
5	0,00	0,00	0,00	0,00	0,22	0,00	0,00	0,44	0,62
6	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,00
7	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,35	1,01
8	0,00	0,00	0,00	0,45	0,00	0,00	0,00	0,89	0,00
9	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,41	0,03
10	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,28	0,24
11	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,09	0,11
12	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,26	0,00
13	0,18	0,00	0,00	0,04	0,00	0,00	0,00	1,93	0,85
14	0,18	0,00	0,01	0,02	0,00	0,00	0,00	0,51	0,78
15	0,00	0,00	0,00	0,98	0,00	0,00	0,11	0,24	0,61
16	0,00	0,00	0,00	0,89	0,00	0,00	0,09	0,23	0,00
17	0,00	0,00	0,00	0,00	0,00	0,72	0,00	1,82	0,00
18	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,08	0,00
19	0,00	0,00	0,00	0,00	0,00	0,14	0,00	0,19	0,00
20	0,01	0,00	0,00	0,00	0,00	0,00	0,00	1,54	0,00
21	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,12	0,00
22	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,45
23	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,02	0,26
24	0,00	0,00	0,00	0,00	0,00	0,21	0,00	0,13	0,00
25	0,00	0,00	0,00	0,00	0,00	0,31	0,00	0,14	0,01
26	0,00	0,00	0,00	0,00	0,03	0,00	0,00	0,17	0,02
27	0,02	0,00	0,00	0,37	0,00	0,00	0,00	0,84	0,01
28	0,00	0,00	0,00	0,25	0,02	0,03	0,00	1,56	0,00

## “Apêndice A, continuação”

<b>Mol.</b>	<b>D-1</b>	<b>D-2</b>	<b>D-3</b>	<b>D-4</b>	<b>D-5</b>	<b>D-6</b>	<b>D-7</b>	<b>D-8</b>	<b>D-9</b>
29	0,00	0,00	0,00	0,28	0,00	0,00	0,38	1,26	0,91
30	0,00	0,00	0,00	0,40	0,00	0,01	0,31	0,40	0,23
31	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,11	0,04
32	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,02	0,25
33	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,11	0,00
34	0,00	0,00	0,00	0,00	0,00	0,00	0,31	1,12	0,01
35	0,00	0,13	0,04	0,84	0,00	0,00	0,00	0,47	0,23
36	0,00	0,00	0,67	0,00	0,01	0,00	0,00	0,07	0,60
37	0,00	0,73	0,00	0,00	0,00	0,00	0,00	0,33	0,03
38	0,00	0,09	0,50	0,27	0,00	0,00	0,39	1,05	0,61
39	0,00	0,10	0,47	0,00	0,00	0,00	0,30	0,09	0,33
40	0,00	0,09	0,32	0,00	0,00	0,00	0,46	0,14	0,34
41	0,00	0,01	0,46	0,36	0,00	0,00	0,11	0,36	0,07
42	0,00	0,00	0,00	0,65	0,00	0,00	0,45	0,16	0,04
43	0,00	0,04	0,00	0,00	0,00	0,00	0,36	0,18	0,26
44	0,20	0,08	0,13	0,26	0,05	0,01	0,41	0,11	0,26
45	0,00	0,25	0,64	0,73	0,00	0,00	0,03	0,86	0,03
46	0,00	0,00	0,47	0,00	0,00	0,69	0,00	0,00	0,00
47	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,33
48	0,05	0,00	0,00	0,10	0,00	0,02	0,00	0,21	0,07
49	0,00	0,02	0,12	0,00	0,33	0,00	0,00	0,06	0,05
50	0,14	0,05	0,83	0,00	1,16	0,00	0,00	0,13	0,03
51	0,00	0,18	0,29	0,00	0,00	0,00	0,00	0,22	0,09
52	0,00	0,00	0,00	0,00	0,00	0,00	0,02	2,12	0,00
53	0,00	0,55	0,00	0,95	0,00	0,00	0,00	0,90	0,00
54	0,00	0,00	0,00	0,24	0,00	0,00	0,00	0,05	0,00
55	0,00	0,01	0,01	0,03	0,00	0,02	0,32	0,07	0,92
56	0,00	0,30	0,42	0,00	0,00	0,15	0,22	0,20	0,29
57	0,00	0,14	0,59	0,00	0,00	0,00	0,00	0,06	0,14
58	0,01	0,11	0,63	0,35	0,00	0,01	0,09	0,82	0,00

## “Apêndice A, continuação”

<b>Mol.</b>	<b>D-1</b>	<b>D-2</b>	<b>D-3</b>	<b>D-4</b>	<b>D-5</b>	<b>D-6</b>	<b>D-7</b>	<b>D-8</b>	<b>D-9</b>
<b>59</b>	0,00	0,00	0,45	0,00	0,00	0,72	0,00	0,20	0,42
<b>60</b>	0,00	0,00	0,72	0,00	0,14	0,01	0,01	0,00	0,61
<b>61</b>	0,00	0,09	0,68	0,28	0,00	0,02	0,05	1,11	0,04
<b>62</b>	0,00	0,08	0,74	0,16	0,00	0,00	0,0	0,32	1,07
<b>63</b>	0,00	0,01	0,44	0,00	0,00	0,00	0,01	0,05	0,73
<b>64</b>	0,00	0,00	0,65	0,00	0,01	0,07	0,00	0,07	0,80
<b>65</b>	0,00	0,00	0,00	0,00	0,00	1,64	0,43	0,06	0,00
<b>66</b>	0,00	0,04	0,24	0,00	0,00	0,86	0,51	0,64	0,00
<b>67</b>	0,00	0,15	0,59	0,32	0,00	0,02	0,03	0,77	0,00
<b>68</b>	0,01	0,11	0,70	0,27	0,00	0,01	0,00	0,15	0,00
<b>69</b>	0,00	0,19	0,63	0,24	0,00	0,01	0,04	0,73	0,00
<b>70</b>	0,24	0,01	0,00	0,00	0,00	0,00	0,00	1,13	0,00
<b>71</b>	0,00	0,11	0,67	0,31	0,00	0,01	0,61	1,14	0,26
<b>72</b>	0,00	0,01	0,13	0,00	0,00	0,00	0,00	0,11	0,34
<b>73</b>	0,00	0,06	0,24	0,00	0,00	0,00	0,00	0,34	0,26
<b>74</b>	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,12	0,02
<b>75</b>	0,00	0,01	0,02	0,00	0,00	0,00	0,00	1,25	0,01
<b>76</b>	0,00	0,09	0,01	0,00	0,00	0,00	0,00	1,19	0,39
<b>77</b>	0,00	0,01	0,01	0,00	0,00	0,00	0,00	1,01	0,02
<b>78</b>	0,00	0,02	0,06	0,00	0,00	0,00	0,00	0,66	0,01
<b>79</b>	0,00	0,23	0,64	0,00	0,00	0,22	0,00	0,38	0,00
<b>80</b>	0,00	0,12	0,34	0,04	0,00	0,01	0,00	0,30	0,01
<b>81</b>	0,00	0,02	0,84	0,00	0,00	0,00	0,00	0,07	0,00
<b>82</b>	0,00	0,28	0,63	0,00	0,00	0,06	0,01	0,18	0,01
<b>83</b>	0,00	0,00	0,35	0,01	0,00	0,73	0,05	0,58	0,02
<b>84</b>	0,00	0,03	0,86	0,00	0,00	0,60	0,01	0,13	0,00
<b>85</b>	0,00	0,38	0,40	0,01	0,00	0,28	0,00	0,21	0,1
<b>86</b>	0,00	0,13	0,61	0,00	0,00	0,07	0,00	0,34	0,18
<b>87</b>	0,00	0,69	0,00	0,71	0,00	0,00	0,00	0,94	0,00
<b>88</b>	0,00	0,15	0,69	0,00	0,00	0,08	0,11	0,08	0,34
<b>89</b>	0,00	0,00	0,48	0,00	0,00	0,54	0,01	0,44	0,08

