



LETÍCIA CRISTINA DE ASSIS

**MODELOS DE QSAR APLICADOS AO ESTUDO
DE INIBIDORES IMIDAZOPIRIDINÍCOS DA
TIROSINA QUINASE DO BAÇO**

LAVRAS – MG

2015

LETÍCIA CRISTINA DE ASSIS

**MODELOS DE QSAR APLICADOS AO ESTUDO DE INIBIDORES
IMIDAZOPIRIDÍNICOS DA TIROSINA QUINASE DO BAÇO**

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

2015

**Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca
Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).**

Assis, Leticia Cristina De.

Modelos de QSAR aplicados ao estudo de inibidores
imidazopiridínicos da tirosina quinase do baço / Leticia Cristina
De Assis. – Lavras : UFLA, 2015.

111 p. : il.

Dissertação (mestrado acadêmico)–Universidade Federal de
Lavras, 2015.

Orientador(a): Elaine Fontes Ferreira da Cunha.

Bibliografia.

1. Tirosina Quinase do Baço. 2. Compostos de
imidazopiridina. 3. QSAR4D-IR. I. Universidade Federal de
Lavras. II. Título

LETÍCIA CRISTINA DE ASSIS

**MODELOS DE QSAR APLICADOS AO ESTUDO DE INIBIDORES
IMIDAZOPIRIDÍNICOS DA TIROSINA QUINASE DO BAÇO**

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 2015.

Dra. Ana Paula Guimarães	UFV
Dra. Daiana Teixeira Mancini	UFLA
Dr. Teodorico de Castro Ramalho	UFLA

Dra. Elaine Fontes Ferreira da Cunha
Orientadora

LAVRAS – MG

2015

AGRADECIMENTOS

A Deus, por me amparar nos momentos difíceis, por me dar força interior para superar as dificuldades, mostrar o caminho nas horas incertas e me suprir em todas as minhas necessidades.

À minha mãe Agda, ao meu padrasto Joel e aos meus irmãos, João Victor, Bruna e Rafaela, aos quais amo muito. O meu coração vibra de gratidão pelo carinho, apoio, confiança, por me incentivar e acreditar em mim, por todo suporte em todos os momentos da minha vida, pelo dia a dia único com vocês, por serem tão especiais.

E minha sincera gratidão, à minha orientadora Prof^a Elaine, por toda atenção, paciência, cuidado e créditos concedidos a mim. Mas, principalmente, pela convivência enriquecedora, pela amizade, pela orientação, pelos conselhos, e por me fazer mergulhar em um novo mundo de conhecimentos e descobertas. Foi uma pessoa essencial para que esse trabalho pudesse ser realizado.

Aos meus amigos do Grupo de Química Computacional, a “Era do Gelo” que sempre me encorajavam a seguir adiante. Em especial, às meninas do QSAR-4D, Dani, Tamiris, Letícia e Jovas, pela ajuda e disponibilidade sempre. A Dai, pelos sábios conselhos, pessoa que sempre admirei. Ao Stephen e o Alexandre Castro, pela ajuda em tantos momentos. A Fátima, pelos estudos nos finais de semana, pelas brincadeiras como “Alinhamento de Assis”. A Silvana, pela amizade. A vocês, a minha gratidão pelos bons momentos.

Aos professores Teodorico e Matheus, que sempre estiveram dispostos a ajudar, pelo convívio, pelos incentivos e pelo carinho.

Ao Prof. Teodorico, Prof.^a Ana Paula, Dr.^a Daiana, por aceitarem o convite de participação na banca e suas contribuições.

À Capes pela concessão da bolsa de estudos, e à Universidade Federal de Lavras, pela oportunidade de realização deste trabalho.

RESUMO

A análise da relação quantitativa estrutura atividade em quarta dimensão independente do receptor (QSAR4D-IR), foi aplicada numa série de 97 compostos análogos de imidazopiridinas sintetizados e avaliados farmacologicamente por Blomgren e colaboradores. Esses compostos são inibidores da tirosina quinase do baço (Syk, do inglês *Spleen tyrosine kinase*), enzima responsável pela transdução de sinal de imunoreceptores clássicos. A desregulação da Syk está associada a diversas patologias, dentre as quais se destaca o crescimento descontrolado de células tumorais. O inibidor mais promissor da Syk, fostamatinib, demonstrou eficácia em várias indicações terapêuticas, mas a sua evolução clínica ainda está em processo. Nesse contexto, torna-se necessária a busca por novos inibidores potentes, assim, neste trabalho, desenvolvemos e validamos modelos de QSAR4D-IR, a fim de obter características farmacofóricas relevantes para essa classe de compostos. As conformações obtidas por simulação de dinâmica molecular foram sobrepostas, de acordo com os seis alinhamentos testados, em uma caixa virtual tridimensional composta por células de 1Å. 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_{ajustado}) e coeficiente de determinação (R^2) de 0,60 e 0,75, respectivamente. Analisando os descritores, pode-se observar que grande parte da contribuição dessas variáveis apresenta coeficiente negativo, e é do tipo interação não polar, sugerindo que grupos hidrofílicos, em todas as regiões que possuem substituintes, são importantes para o aumento do potencial inibitório.

Palavras-chave: Tirosina Quinase do Baço. Compostos de imidazopiridina. QSAR4D-IR.

ABSTRACT

The analysis of the receptor-independent 4D-QSAR has been applied to a series of 97 analogous compounds of imidazopyridines synthesized and pharmacologically evaluated by Blomgren and collaborators. These compounds are inhibitors of Spleen tyrosine kinase (Syk), enzyme responsible for the signal transduction of classic immunoreceptors. The deregulation of Syk is associated with many pathologies, among which we highlight the uncontrolled growth of tumor cells. The most promising Syk inhibitor, R406 (fostamatinib), has proved efficient in multiple therapeutic indications, however, its clinical evolution is still in process. In this context, it becomes necessary to search for new potent inhibitors therefore, in this work, we developed and validated 4D-QSAR models in order to obtain relevant pharmacophore features for this class of compounds. The conformations obtained by molecular dynamic simulation were overlapped, according to the six trial alignments, in a virtual three-dimensional box comprised of 1 Å cells. The models were generated by the combined techniques of genetic algorithm (GA) and partial least squares (PLS). The best model generated adjusted cross-validation values (q^2_{adjusted}) and coefficient of determination values (R^2) of 0.60 and 0.75, respectively. Analyzing the descriptors, we can observe that most contributions of these variables present a negative coefficient and is of the non-polar interaction type, suggesting that hydrophilic groups in all regions, which have substituents, are important for increasing the inhibitory potential.

Keywords: Spleen Tyrosine Kinase. Imidazopyridine compounds. IR-4DQSAR.

LISTA DE FIGURAS

- Figura 1 Mecanismo de transferência do fosfato terminal do ATP para o grupo hidroxila do resíduo de tirosina.....22
- Figura 2 Modelo esquemático demonstrando a estrutura linear da Syk com as tirosinas que são fosforiladas após a sua ativação. Os principais locais de autofosforilação da Syk são identificados nos domínios ligantes A e B. O domínio ligante B age como uma região de encaixe que fornece resíduos fosfotirosina para os domínios de ligação SH2 de outras moléculas de sinalização26
- Figura 3 Modelo esquemático da ativação da Syk. Após o acoplamento de um receptor com o antígeno (exemplo receptores de célula B), ocorre a fosforilação do ITAM pela Lyn, que fosforila a Syk no resíduo Tyr130, sendo, então, possível a sua autofosforilação.....27
- Figura 4 Estruturas gerais de diferentes classes de inibidores da Syk: diaminopirimidina **(A)**, imidazo [1,2a] piridina **(B)**, 1,6-naftiridinas **(C)**, aminopirimidina **(D)** e oxidóis **(E)**28
- Figura 5 Inibidor da quinase Syk. **(F)** R788/fostamatinib é a forma de pró-fármaco de **(G)** R406 e é o melhor caracterizado dentre os inibidores da Syk em doentes. Ambos desenvolvidos pela empresa Farmacêutica Rigel29
- Figura 6 Drogas derivadas de imidazo [1,2- α] piridina. **(H)** Imidazol [1,2, α]piridina, **(I)** Alpidem (BERSON et al., 2001; SALETU; GRUNBERGER; LINZMAYER, 1986) e **(J)** Zolpidem (HARRISON; KEATING, 2005) ambos são usados para tratar a depressão e insônia, **(K)** Necopidem (BOERNER; MOLLER, 1997) essa droga possui efeitos sedativos31

Figura 7	Estrutura geral dos compostos desenvolvidos por Blomgren e colaboradores	42
Figura 8	Representação da estrutura do ligante OVG, N-{6 - [(2S)-2-metilpirrolidin-1-il] piridin-2-il}-6-fenilimidazo[1,2-b]piridazin-8-amina, cristalizado no sítio ativo da enzima Sky	51
Figura 9	Três letras (a,b,c) ordenando os átomos de carbono utilizados nos estudos de QSAR-4D definindo seis alinhamentos testados. Composto 19 ($pIC_{50\text{experimental}}=5,69$), o inibidor mais potente da série de Blomgren e colaboradores, é usado para definir o código de letra do átomo	55
Figura 10	Molécula 8 da série estudada, representando o descritor em amarelo que está localizado na parte comum de todos os compostos do conjunto de treinamento	66
Figura 11	Molécula 19 mais ativa da série, representando a distância do par de GCODs correlacionados	72
Figura 12	Composto 19 com os descritores gerados pelo modelo M1 . Cada GCOD é descrito como “x, y, z, IPE”, que representam as coordenadas cartesianas da célula de grade selecionada (x, y, z) e o respectivo tipo de átomo (IPE). Os GCODs descritos são: (1) (0, 6, -2, np), (2) (0, 3, -3, np), (3) (0, 10,3, np), (4) (2, 1, -8, np), (5) (-1, 9, 0, np), (6) (0, -3, -5,np), (7) (0, 0,-3,alh), (8) (0, 9,-2, np), (9) (1,-1,-4, np), (10) (-1, 6, 3, np), (11) (0,1,-3, np).....	79
Figura 13	Modelo M1 de QSAR-4D-IR representativo do alinhamento 1, mostrando o GCOD com coeficiente negativo (0,6,-2,np) (esfera laranja) localizado no anel indol do composto 80 ,no qual apresentou maior frequência de ocupação para esse GCOD	80

Figura 14	Estruturas químicas dos compostos de imidazopiridinas e suas respectivas atividades biológicas experimentais.	80
Figura 15	Representação da interação do NH livre, do ligante CG9, cocrystalizado no sítio da Syk obtida do PDB, cujo código: 4PUZ, fazendo interação de hidrogênio com o resíduo de aminoácido Asp512.....	81
Figura 16	Modelo M1 de QSAR4D-IR representativo do alinhamento 1, mostrando os GCODs com coeficiente negativo (esferas laranjas).....	82
Figura 17	Estruturas químicas dos compostos de imidazopiridinas e suas respectivas atividades biológicas experimentais	83
Figura 18	Estruturas químicas dos compostos de imidazopiridinas e suas respectivas atividades biológicas experimentais	83
Figura 19	Modelo M1 de QSAR-4D-IR representativo do alinhamento 1, mostrando o GCOD com coeficiente negativo (2,1,-8, np) (esfera laranja) localizado no substituinte etil do composto 79 que apresentou maior frequência de ocupação para esse GCOD	84
Figura 20	Estruturas químicas dos compostos de imidazopiridinas e suas respectivas atividades biológicas experimentais	84
Figura 21	Representação do composto 48 no sítio da enzima Syk, observa-se que o resíduo de aminoácido lisina se encontra próximo a essa região, reforçando o argumento de que compostos polares, nessa região, contribuem com a atividade desse inibidor.....	85
Figura 22	Modelo M1 de QSAR-4D-IR representativo do alinhamento 1, mostrando o GCOD com coeficiente negativo (0,10,3,np) (esfera laranja) localizado no grupo indol do composto 79 ,no	

	qual apresentou maior frequência de ocupação para esse GCOD	86
Figura 23	Modelo M1 de QSAR-4D-IR representativo do alinhamento 1, mostrando o GCOD com coeficiente negativo (0,0,-3,alh) (esfera laranja), localizado no átomo de nitrogênio do grupo pirimidina do composto 12 , no qual apresentou maior frequência de ocupação para esse GCOD	87
Figura 24	Modelo M1 de QSAR-4D-IR representativo do alinhamento 1, mostrando o GCOD com coeficiente negativo (0, 9, -2, np) (esfera laranja), localizado no grupo indol do composto 76 , no qual apresentou maior frequência de ocupação para esse GCOD	88
Figura 25	Estruturas químicas dos compostos de imidazopiridinas e suas respectivas atividades biológicas experimentais	88
Figura 26	Representação do composto 76 dentro do sítio da enzima Syk obtida do PDB.Observa-se que os resíduos de aminoácidos Serina e Arginina se encontram próximos a essa região, reforçando o argumento de que, compostos polares, nessa região, contribuem com a atividade desse inibidor	89
Figura 27	Modelo M1 de QSAR-4D-IR representativo do alinhamento 1, mostrando o GCOD com coeficiente positivo (1,-1,-4,np) (esfera verde) localizado no átomo de oxigênio do anel oxazina do composto 53 ,composto este que apresentou maior frequência de ocupação para esse GCOD.A direita o composto 19 ,mostrando que durante a DM esse descritor se posicionou no átomo CH ₂ próximo a essa região,esse fato explica a presença desse GCOD.....	90

Figura 28	Modelo M1 de QSAR-4D-IR representativo do alinhamento 1, mostrando o GCOD com coeficiente negativo (-1,6,3,np) (esfera laranja), localizado no grupo indazol do composto 77 , onde apresentou maior frequência de ocupação para esse GCOD	91
Figura 29	Estruturas químicas dos compostos de imidazopiridinas e suas respectivas atividades biológicas experimentais	91
Figura 30	Resumo geral das informações retiradas do Modelo M1, resultante da metodologia QSAR-4D.....	92
Figura 31	Estrutura hipotética (P1), proposta para estudos posteriores, baseados nos resultados obtidos no estudo de QSAR4D-IR de inibidores imidazopiridínicos. À direita, composto 19 mais ativo da série.....	93
Figura 32	Estrutura hipotética (P2), proposta para estudos posteriores, baseados nos resultados obtidos no estudo de QSAR4D-IR de inibidores imidazopiridínicos	93
Figura 33	Estrutura hipotética (P3), proposta para estudos posteriores baseados nos resultados obtidos no estudo de QSAR4D-IR de inibidores imidazopiridínicos	94

LISTA DE GRÁFICOS

Gráfico 1	Distribuição dos valores de $pIC_{50\text{experimental}}$ para os compostos do conjunto de treinamento e de teste	50
Gráfico 2	Representação dos valores de energia total para DM de 100.000 ps referente ao composto 19	61
Gráfico 3	Valores de $pIC_{50\text{Experimental}}$ <i>versus</i> $pIC_{50\text{Predito}}$ do conjunto de treinamento usando o melhor modelo de QSAR-4D representativo do alinhamento 1.....	75

LISTA DE QUADROS

Quadro 1	Estrutura química e os respectivos valores de $pIC_{50}(M)$ dos 97 compostos imidazopiridínicos. Em negrito e sublinhado estão representadas as moléculas selecionadas para compor o grupo teste, as demais moléculas, participam do grupo de treinamento da metodologia QSAR-4D42
----------	---

LISTA DE TABELAS

Tabela 1	Parâmetros Estatísticos utilizados na validação de QSAR-4D.....	59
Tabela 2	Representação dos dez modelos de QSAR-4D obtidos do Alinhamento 1, e célula de 1Å.....	67
Tabela 3	Matriz de correlação cruzada dos resíduos de atividade entre pares de modelos para os dez modelos obtidos na análise de QSAR-4D usando o alinhamento 1 com valor de corte por variância igual a zero e célula de 1Å.....	69
Tabela 4	Matriz de correlação cruzada dos GCODs do melhor modelo do Alinhamento 1.....	71
Tabela 5	Valores de $pIC_{50Experimental}$ e $pIC_{50predito}$ e valores residuais ($pIC_{50predito} - pIC_{50experimental}$) para os compostos do conjunto de treinamento.....	73
Tabela 6	Valores de $pIC_{50Experimental}$ e $pIC_{50predito}$ e valores residuais ($pIC_{50predito} - pIC_{50experimental}$) para os compostos do conjunto de teste.....	76
Tabela 7	Valores de energia de interação intermolecular proteína/inibidores.....	94
Tabela 8	Valores de energia de interação intermolecular proteína/inibidores, energias de ligações de hidrogênio e número de ligações de hidrogênio.....	95

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

ABL	Tirosina quinase Abelson
ATP	Trifosfato de adenosina
BCR	Receptor de célula B
BTK	Tirosina quinase de Bruton
CEP	<i>Conformational Ensemble Profile</i>
CoMFA	<i>Comparative Molecular Field Analysis</i>
DP	Desvio Padrão
EGFR	Receptor do fator de crescimento Epidérmico
FGFR	Receptor do fator de crescimento de Fibroblastos
GA	<i>Geneticalgorithm</i>
GFRs	Receptores do fator de Crescimento
GCOD	<i>Grid Cell Occupancy Descriptor</i>
GFA	<i>Genetic Function Approximation</i>
GFRs	Receptores do fator de crescimento
IGFR	Receptor de insulina
IPE	<i>Interaction Pharmacophore Elements</i>
ITAM	Sequência de ativação do imunorreceptor baseado em tirosina
ITKs	Inibidores das tirosinas quinases
LSE	<i>least-squares error</i>
LOF	<i>lack-of-fit</i>
LOO-cv	<i>leave-one-out cross validation</i>
PDB	<i>Protein Data Bank</i> (Banco de Dados de Proteínas)
PDGFR	Receptor do fator de crescimento derivado das plaquetas
PH	Domínio de homologia com a pleckstrina
PLS	<i>Partial least squares</i>

PTKs	Proteínas tirosina quinase
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>
SH2	Domínio de homologia a src 2
SH3	Domínio de homologia a src 3
Syk	Tirosina Quinase do Baço
TK	Tirosina Quinase
TKRs	Receptor de Tirosina Quinase
TKNRs	Não Receptor de Tirosina Quinase
UFLA	Universidade Federal de Lavras
ZAP-70	Cadeia zeta associada a proteína quinase 70

LISTA DE SÍMBOLOS

Å	Angstrom
IC ₅₀	concentração capaz de inibir 50% da atividade enzimática
µM	Micromolar
M	Molar
pIC ₅₀	Log de IC ₅₀
kDa	kilo Dalton

SUMÁRIO

1	INTRODUÇÃO	20
2	REFERENCIAL TEÓRICO	22
2.1	Tirosina Quinase	22
2.2	Tirosina quinase Syk.....	244
2.3	Mecanismo de ativação da Syk	25
2.4	Inibidores da Syk	28
2.5	Imidazopiridinas	30
2.6	Química medicinal computacional no desenvolvimento de fármacos.....	31
2.7	Relação Quantitativa Estrutura-Atividade (QSAR).....	33
2.8	Evolução da Metodologia de QSAR	34
2.9	Simulação de Dinâmica Molecular	37
3	OBJETIVOS	40
4	METODOLOGIA	41
4.1	Programas.....	41
4.2	Dados Biológicos.....	41
4.3	Construção e Otimização das Estruturas Tridimensionais.....	51
4.4	Metodologia do QSAR-4D: Divisão do conjunto de treinamento e de teste.....	52
4.5	Simulação por Dinâmica Molecular	52
4.6	Definição dos elementos de interação farmacofórica.....	53
4.7	Definição do alinhamento e tamanho das células.....	54
4.8	Construção do modelo QSAR-4D: Redução de dados e abordagem PLS-GA.....	56
4.9	Validação dos modelos de QSAR-4D.....	58
4.10	Identificação dos compostos <i>outliers</i>	60
4.11	Seleção da conformação “representativa”	60
5	RESULTADOS E DISCUSSÃO	61
5.1	Análise de simulação por Dinâmica.....	61
5.2	Seleção dos melhores alinhamentos de QSAR-4D.....	62
5.3	Análise do Alinhamento 1.....	66
5.4	Análise do melhor modelo do Alinhamento 1.....	70
5.5	Análise dos Valores Residuais e Identificação dos <i>Outliers</i>	73
5.6	Análise dos Descritores	78
6	CONCLUSÃO	96
	REFERÊNCIAS	97
	APÊNDICE.....	108

1 INTRODUÇÃO

A tirosina quinase do baço (Syk do inglês, spleentyrosinekinase) é uma enzima citoplasmática, que está envolvida na transdução do sinal de imunorreceptores. Esses imunorreceptores contêm uma sequência de ativação baseada em tirosina que, quando fosforilados, transportam o local de ligação específico para os domínios de ativação da Syk, iniciando, assim, uma série de acontecimentos biológicos, tais como: a liberação de citocinas, diferenciação e a proliferação. Por essas razões, a desregulação dessa enzima está associada a diversas patologias, como, por exemplo: doenças alérgicas e doenças mediadas por anticorpos autoimunes como a artrite reumatóide, asma e rinite alérgica (KAUR; BAHIA; SILAKARI, 2012; RICCABONI; BIANCHI; PETRILHO, 2010). Além disso, a Syk também desempenha um papel importante no crescimento descontrolado de células tumorais (KAUR; SINGH; SILAKARI, 2013).

Assim, essa quinase é considerada um dos alvos biológicos promissores para o planejamento de novas moléculas candidatas a fármacos e estudar novas classes de compostos que apresentam características importantes na inibição dessa enzima, pode ser interessante no auxílio de uma estratégia de tratamento para essas doenças. Além disso, o inibidor mais promissor da Syk, denominado fosfataminib, demonstrou eficácia em várias indicações terapêuticas, mas a sua evolução clínica tem sido dificultada por diversos efeitos prejudiciais, como neutropenia e hipertensão (CURRIE et al., 2014), logo, fortalece a importância da procura de novos inibidores da tirosina quinase do baço.

Dessa forma, estudos que envolvem processos evolutivos de reconhecimento molecular em sistemas biológicos são indispensáveis para a descoberta de um fármaco, pois favorecem a compreensão de algumas propriedades físico-químicas como, por exemplo, potência e a afinidade. Com

isso, as ferramentas biotecnológicas associadas aos métodos de química medicinal computacional são importantes no desenvolvimento de novos compostos com atividade biológica (GUIDO; ANDRICOPULO; OLIVA, 2010; MANCINI et al., 2012).

Atualmente, algumas ferramentas teóricas, como, por exemplo, Modelagem Comparativa, Dinâmica Molecular, *Docking* e QSAR, vêm sendo utilizadas para propor novos alvos moleculares com o intuito de direcionar o planejamento de novos candidatos a fármacos para o tratamento de diversas doenças (ASSIS et al., 2013, 2014; MANCINI et al., 2012; SILVA; RAMALHO; CUNHA, 2014).

Dentre essas ferramentas, um método muito utilizado é o estudo da relação quantitativa entre a estrutura química e a atividade biológica (QSAR) (LILL, 2007). Essa metodologia leva em consideração a relação entre as propriedades de uma molécula e sua estrutura química, buscando estabelecer relações matemáticas simples para descrever e prever a atividade de um conjunto de inibidores congêneres (ANDRADE et al., 2010).

Em 2011, Blomgren e seus colaboradores desenvolveram uma série de compostos derivados de imidazopiridinas que possuem atividade inibitória para a Syk (BLOMGREN et al., 2011). Essa classe de compostos exibe uma variedade de propriedades biológicas como, por exemplo, antimicrobiana, antiprotozoária, portanto, sendo considerados compostos que apresentam grande importância farmacológica (ZENG et al., 2012).

Nesse contexto, conduziu-se este trabalho, com o objetivo de empregar o método de correlação quantitativa entre a estrutura-atividade em quatro dimensões (QSAR-4D) em uma série de análogos imidazopiridinas sintetizados e avaliados farmacologicamente por Blomgren e colaboradores. Com isso, foi possível identificar características que possam contribuir com o aumento do poder inibitório desses compostos.

2 REFERENCIAL TEÓRICO

2.1 Tirosina Quinase

Nos últimos anos, o interesse científico pela classe de proteínas tirosina quinase (PTK) como alvo para o desenvolvimento de fármacos tem aumentado (MARZARO et al., 2011), uma vez que são moléculas cruciais na comunicação intra e inter-celular e possuem um papel central nos processos de transdução de sinal (BAHIA et al., 2007). Essas proteínas são conhecidas por estarem relacionadas às diversas funções biológicas, incluindo processos celulares, tais como: a adesão, a motilidade, proliferação, controle do ciclo celular ou a morte celular. Bem como as funções biológicas ao nível de todo o organismo, por exemplo, no crescimento e desenvolvimento, metabolismo ou defesa imune (PUERTA; MÓCSAI, 2013).

As PTKs pertencem a uma grande família de enzimas homólogas, transmembranas e citoplasmáticas, que catalisam a transferência do grupo γ -fosfato do trifosfato de adenosina (ATP) a grupos hidroxila de resíduos de tirosina em substratos de proteínas, conforme demonstrado na Figura 1 (LA MOTTA et al., 2009).

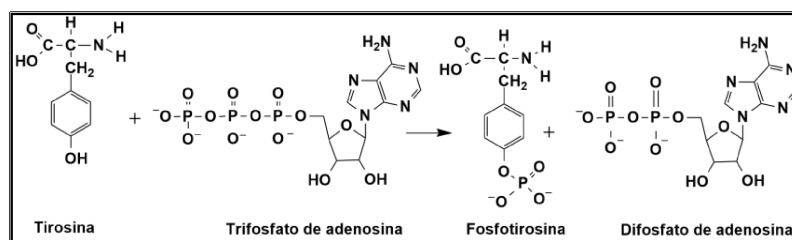


Figura 1 Mecanismo de transferência do fosfato terminal do ATP para o grupo hidroxila do resíduo de tirosina

Tendo em vista o seu papel importante na proliferação, migração e sobrevivência celular, a desregulação das PTKs está associada a diversos tipos

de doenças, inclusive o câncer (PUERTA; MÓCSAI, 2013). Dessa forma, essa classe de proteínas representa um alvo muito atraente para o desenvolvimento de novos agentes quimioterápicos (MARZARO et al., 2011; LA MOTTA et al., 2009).

As PTKs, em razão da organização celular diferenciada, são, geralmente, divididas em receptores de tirosina-quinase (TKRs) e não receptores de tirosina-quinase (TKNRs) (AVILA; ROMEIRO, 2010; KAUR et al., 2013).

As TKRs são proteínas transmembranas, reguladoras-chave de processos celulares. Estruturalmente, possuem um domínio extracelular de ligação ao substrato e um domínio quinase intracelular catalítico (RANIERI et al., 2013). Alguns exemplos das TKRs são o receptor de insulina (IGFR), receptores do fator de crescimento (GFRs), receptor do fator de crescimento epidérmico (EGFR), receptor do fator de crescimento de fibroblastos (FGFR), receptor do fator de crescimento derivado das plaquetas (PDGFR) e receptor do fator de crescimento do endotélio vascular (VEGFR) (AVILA; ROMEIRO, 2010; RANIERI et al., 2013).

Por outro lado, as TKNRs são proteínas citoplasmáticas, conhecidas por desempenharem um papel chave na proliferação e diferenciação celular, também estão envolvidas na migração das células, adesão e em rearranjos do citoesqueleto (BAHIA et al., 2007). Estruturalmente, as TKNRs apresentam um domínio quinase e possuem vários domínios de interação proteína-proteína, como SH2 (domínio de homologia a src 2), SH3 (domínio de homologia a src 3) e PH (domínio de homologia com a pleckstrina). Exemplos de TKNRs são as proteínas BTK (tirosina quinase de Bruton), ABL (tirosina quinase abelson), Syk (tirosina quinase do baço) e ZAP-70 (Cadeia zeta associada à proteína quinase 70) (AVILA; ROMEIRO, 2010).

2.2 Tirosina quinase Syk

A tirosina quinase do baço (Syk) e ZAP-70 são os membros da família da quinase Syk do tipo não receptora (KAUR; BAHIA; SILAKARI, 2012). Inicialmente, a Syk era considerada um fragmento proteolítico de 40kDa, derivada de uma tirosinaquinase p72 presente no baço e pulmão. Entretanto, somente em 1991, é que esta enzima foi clonada, a partir do baço de porco, com 72kDa (KAUR; SINGH; SILAKARI, 2013). Ela é expressa na maioria das células hematopoiéticas, incluindo plaquetas, linfócitos B, mastócitos, exceto nas células T em que a ZAP-70 é relacionada à quinase Syk dominante (CASTILLO et al., 2012; PAMUK; TSOKOS, 2010; PUERTA; MÓCSAI, 2013). Também pode ser expressa em outros tecidos como no tecido epitelial, fibroblastos, células neuronais, e outros tipos de células (LUCAS et al., 2012; RICCABONI; BIANCHI; PETRILHO, 2010).

A Syk é reconhecida como mediadora chave na transdução do sinal de inúmeros imunorreceptores, tais como: receptores de Fc (proteína encontrada na superfície dos macrófagos, mastócitos e neutrófilos, são classificados com base no tipo de anticorpos que reconhecem), e, principalmente, receptores de células B (essenciais para a resposta imune adaptativa e a produção de anticorpos) (RICCABONI; BIANCHI; PETRILHO, 2010). Ela, também, está envolvida na sinalização de receptores de adesão, incluindo as selectinas (glicoproteínas transmembranares que estão envolvidas no rolamento de leucócitos no endotélio) (MÓCSAI; RULAND; TYBULEWICZ, 2010) as integrinas (receptores transmembranares, que desempenham papéis fundamentais das funções efectoras de neutrófilos e de leucócitos) (CHAKRAVARTY et al., 2013; MÓCSAI; RULAND; TYBULEWICZ, 2010) e em células mielóides (subtipo de glóbulos brancos, que são fundamentais em respostas fisiológicas à invasão do

patógeno ou danos no tecido) (KULATHU; GROTHE;RETH, 2009; KYTTARIS et al., 2012; PUERTA; MÓCSAI, 2013).

É importante notar que, a tirosina quinase Syk, está relacionada como desenvolvimento dos osteoclastos e a reabsorção óssea mediada pelos osteoclastos (MÓCSAI et al., 2004),o que demonstra que a Syk também pode estar envolvida em erosões do osso, durante a inflamação das articulações. Além disso, essa enzima também tem sido implicada na transdução de sinal, em células sinoviais de origem não hematopoiética, bem como em função das células endoteliais (PUERTA; MÓCSAI, 2013).

Em resumo, a Syk desempenha papéis cruciais na imunidade adaptativa e mecanismos de inflamação, e também é necessária para o desenvolvimento e a função dos osteoclastos e a transdução de sinal em células sinoviais (PUERTA; MÓCSAI, 2013).

Por essas razões, a Syk tem emergido como um alvo terapêutico em doenças alérgicas, doenças inflamatórias e doenças tumorais (KAUR; SINGH; SILAKARI, 2013), portanto, estudar a inibição dessa enzima pode contribuir para uma estratégia de tratamento para essas doenças.

2.3 Mecanismo de ativação da Syk

A Syk apresenta um domínio catalítico quinase C-terminal e dois domínios (SH2) de homologia Src. Esses domínios estão ligados por duas regiões conhecidos como interdomínio A (localizado como tubo helicoidal entre os dois domínios SH2) e interdomínio B (localizado na região de ligação que une o domínio SH2 para o domínio quinase) (KAUR et al., 2013) como podemos observar na Figura 2.

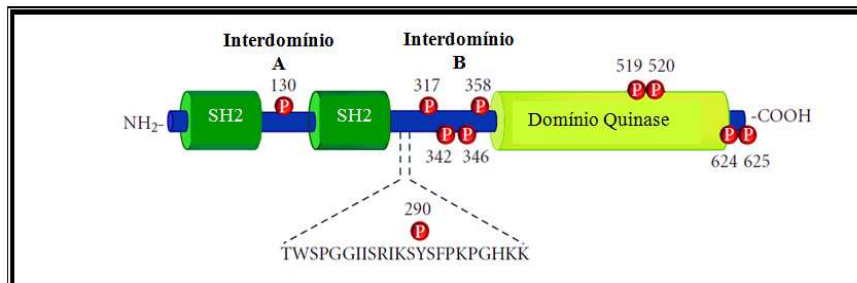


Figura 2 Modelo esquemático demonstrando a estrutura linear da Syk com as tirosinas que são fosforiladas após a sua ativação. Os principais locais de autofosforilação da Syk são identificados nos domínios ligantes A e B. O domínio ligante B age como uma região de encaixe que fornece resíduos fosfotirosina para os domínios de ligação SH2 de outras moléculas de sinalização

Fonte adaptado de Castro (2011).

Os domínios SH2s e ligam a resíduos de tirosinafosforilados dentro da sequência de ativação do imunorreceptor, baseado em tirosina (ITAM em inglês, *immunoreceptor tyrosine-based activating motif*) presente na região citoplasmática dos receptores. O ITAM é uma sequência peptídica curta, composta por dois resíduos de tirosina espaçados que transporta o local de ligação específico para os domínios SH2. O domínio SH2 possui uma elevada flexibilidade conformacional que lhe permite "avaliar" a distância relativa entre fosfotirosinas de ITAM, pertencentes a diferentes receptores e ajustar a orientação relativa dos seus domínios SH2, para ajustar o espaçamento (FANG et al., 2010). Essa particularidade permite à Syk interagir e modular a atividade de vários imunorreceptores diferentes (KAUR; SINGH; SILAKARI, 2013; RICCABONI; BIANCHI; PETRILHO, 2010).

Interações de imunorreceptores específicos presentes sobre a membrana inicia a ativação da enzima Syk (KAUR; SINGH; SILAKARI, 2013; RICCABONI; BIANCHI; PETRILHO, 2010). De acordo com a Figura 3, após a ligação dos receptores de células B com o antígeno, a proteína tirosinaquinase

Lyn, um membro da família Src, é ativada e fosforila o motivo ITAM na cauda citoplasmática do receptor. Essa associação, através do ITAM, fosforila os domínios SH2 da Syk, levando-a a ativação por autofosforilação (KAUR; SINGH; SILAKARI, 2013).

A ativação da enzima Syk fosforila vários outros substratos, muitos dos quais são proteínas do tipo adaptadoras, como SLP-65, que, por sua vez, ativam outras moléculas, iniciando uma cascata de acontecimentos que induzem a várias respostas celulares, tais como: a consequente síntese e liberação de uma quantidade enorme de moduladores inflamatórios responsáveis pela reação alérgica aguda (KAUR; SINGH; SILAKARI, 2013).

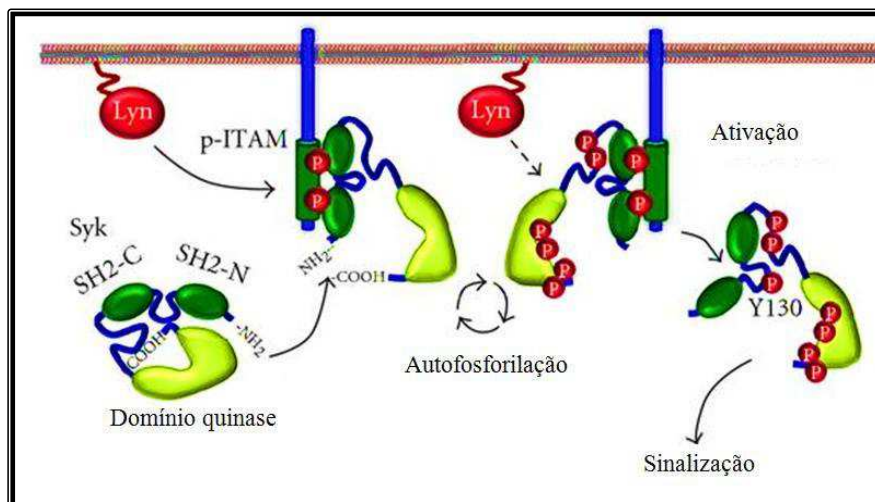


Figura 3 Modelo esquemático da ativação da Syk. Após o acoplamento de um receptor com o antígeno (exemplo receptores de célula B), ocorre a fosforilação do ITAM pela Lyn, que fosforila a Syk no resíduo Tyr130, sendo, então, possível a sua autofosforilação

Fonte adaptado de Castro (2011).

2.4 Inibidores da Syk

Nas últimas décadas, têm-se pesquisado muitas estratégias para o desenvolvimento de inibidores para a enzima Syk. Esse fato ocorre, pois a Syk está posicionada como foco na via de sinalização celular, portanto, as estratégias de seleção de inibidores dirigidas à tirosina quinase do baço podem ser bastante vantajosas.

As diversas classes de potenciais inibidores da Syk têm sido relatadas na literatura, tais como: compostos de diaminopirimidinacarboxamida (LIDDLE et al., 2011), imidazo[1,2-a]piridina (BYTH et al., 2006), 1,6-naftiridinas (CYWIN et al., 2003), aminopiridinas (CASTILLO et al., 2012), e compostos oxidóis (LAI et al., 2003). As estruturas químicas gerais dessas classes são apresentadas na Figura 4.

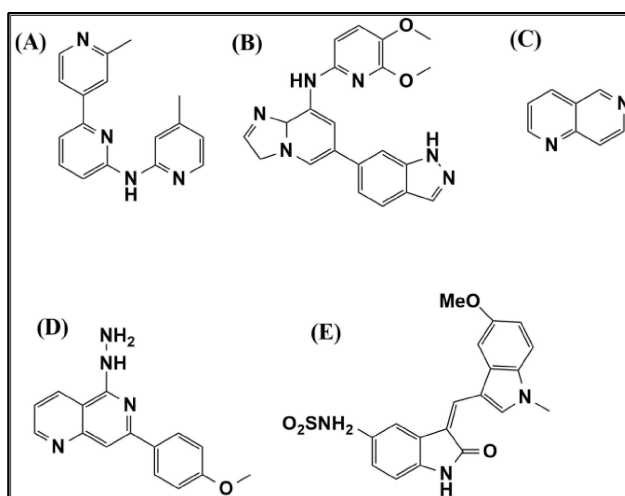


Figura 4 Estruturas gerais de diferentes classes de inibidores da Syk: diaminopirimidina (A), imidazo [1,2a] piridina (B), 1,6-naftiridinas (C), aminopirimidina (D) e oxidóis (E)

Observando as estruturas desses inibidores, a maioria deles apresentam núcleos nitrogenados. Essa característica farmacofórica em comum foi inserida

nessa classe de substâncias, visando a mimetizar o anel adenina do substrato ATP e facilitar a ligação na região de dobradura, conectando os domínios N- e C- terminal do sítio catalítico, já que, a maioria dos inibidores das tirosinas quinases (ITKs), inclusive a quinase Syk, competem pelo sítio de ligação do ATP no sítio catalítico (AVILA; ROMEIRO, 2010; CRUZ; UCKUN, 2013).

Um exemplo de inibidor da Syk, submetido a estudos pré-clínicos que, recentemente, completou os ensaios da fase II para o tratamento da artrite reumatóide (LUCAS et al., 2012) e que compete pelo sítio de ligação do ATP é o Fostamatinib (R788). O R788 é um pró-fármaco que, após a administração oral, é rapidamente convertido ao inibidor R406 (Figura 5) (NIJJAR et al., 2013). O Fostamatinib demonstrou eficácia em várias indicações terapêuticas, mas, a sua evolução clínica tem sido dificultada por vários efeitos colaterais, como diarreia, náuseas, gastrite, neutropenia, entre outros (CURRIE et al., 2014; KAUR; SINGH; SILAKARI, 2013).

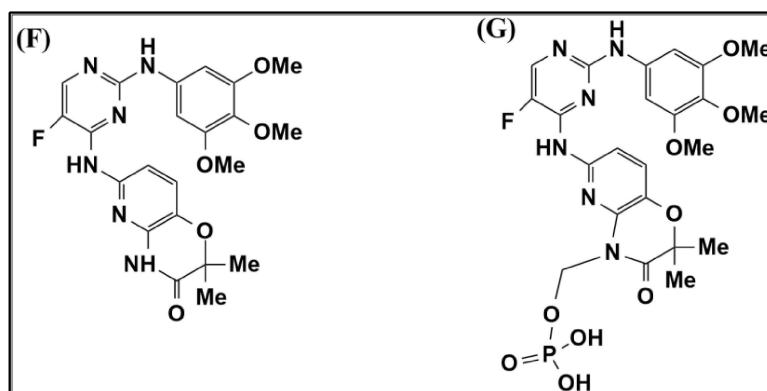


Figura 5 Inibidor da quinase Syk. **(F)** R788/fostamatinib é a forma de pró-fármaco de **(G)** R406 e é o melhor caracterizado dentre os inibidores da Syk em doentes. Ambos desenvolvidos pela empresa Farmacêutica Rigel

Fonte adaptado de Thoma et al. (2014).

Blomgren et al. (2011) sintetizaram compostos derivados de imidazopiridinas que apresentam atividade inibitória para a enzima Syk. Então, é de grande interesse, neste trabalho, pesquisar como a atividade desses compostos

pode estar correlacionada com a estrutura, para que possamos analisar as características farmacofóricas importantes para realizarmos modificações estruturais de forma a aumentar o poder inibitório desses compostos.

2.5 Imidazopiridinas

Imidazopiridinas é uma classe de aminas heterocíclicas que são moléculas fundamentais para a vida. Pequenas alterações na estrutura química desses compostos são responsáveis por grandes variações nos seus efeitos biológicos. Esse grupo de moléculas possui uma extensa gama de bioatividades, como antivirais, antimicrobianas, antiprotozoárias, antibacterianas e anticarcinogênicas (EL-SAYED, 2013; ZENG et al., 2012). Por outro lado, alguns trabalhos (BENDALY et al., 2009; CHENG; CHEN; WANG, 2006) descrevem essa classe de composto como cancerígenas e mutagênicas (EL-SAYED, 2013). Imidazopiridinas, possivelmente interagem com o DNA e exibem atividade biológica e farmacêutica (SIDDIQUI et al., 2013). Na Figura 6, estão representados alguns compostos derivados de imidazopiridinas.

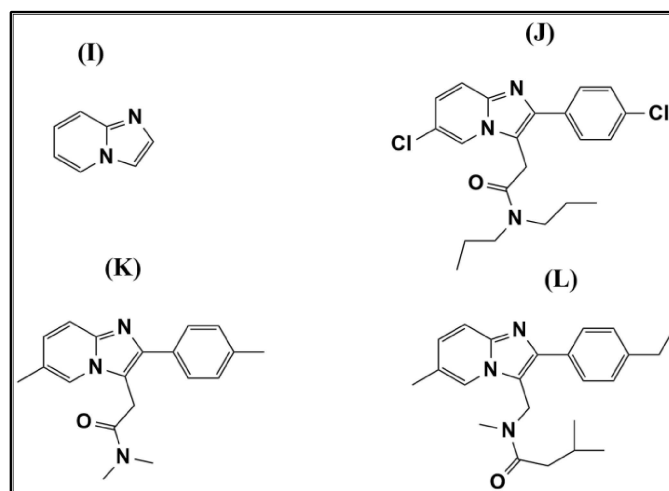


Figura 6 Drogas derivadas de imidazo [1,2- α] piridina. **(H)** Imidazol [1,2, α]piridina, **(I)** Alpidem (BERSON et al., 2001; SALETU; GRUNBERGER; LINZMAYER, 1986) e **(J)** Zolpidem (HARRISON; KEATING, 2005) ambos são usados para tratar a depressão e insônia, **(K)** Necopidem (BOERNER; MOLLER, 1997) essa droga possui efeitos sedativos

Fonte adaptado de Zeng et al. (2012).

2.6 Química medicinal computacional no desenvolvimento de fármacos

A química medicinal é uma área que envolve a descoberta de novos compostos biologicamente ativos, abrangendo a modificação da estrutura molecular das substâncias ou a síntese, o isolamento de substâncias que possuem princípios ativos naturais, presentes em plantas, animais, minerais; o reconhecimento ou elucidação da estrutura; a caracterização das moléculas, desde a sua constituição atômica, levando em consideração a relação entre a estrutura e propriedades, e, também, a proposição e validação de modelos matemáticos, por meio dos estudos de relações entre a estrutura química e a atividade biológica, permitindo propor compostos protótipos de interesse. Essa área de conhecimento, utilizando diferentes estratégias metodológicas

complementares, é tradicional e reconhecidamente utilizada no planejamento de fármacos (AMARAL; MONTANARI, 2002).

O desenho racional de moléculas candidatas a fármacos é frequentemente baseado na exploração estrutural de diferenças fisiológicas entre o patógeno e o hospedeiro. O ponto de partida para esse tipo de estudo é a estrutura tridimensional do alvo molecular escolhido. Este é representado por uma macromolécula que pode ter sua estrutura molecular conhecida ou não. Um terço dos fármacos atuais atuam como inibidores enzimáticos e diversas enzimas são alvos-terapêuticos, sendo que cerca de várias estruturas protéicas estão resolvidas por RMN ou raios-X e são disponíveis em bancos de dados. Em ambos os casos, o planejamento molecular de novo ligante/análogo-ativo pode ser realizado, seja partindo-se do substrato natural ou introduzindo-se modificações moleculares, guiadas pelas estratégias clássicas da Química Medicinal e/ou Química Medicinal Computacional (CARVALHO et al., 2003).

Essas estratégias envolvem métodos computacionais que são, atualmente, ferramentas fundamentais para a descoberta e planejamento de fármacos. Isso porque, considera-se que os métodos computacionais possuem uma grande vantagem frente aos métodos experimentais, por causa do seu baixo custo, uma vez que, evitam repetição de análises e reações, além de gastos com reagentes e materiais, visto que simulações podem ser feitas para otimizar e compreender processos, reações, interações intra e intermoleculares, entre outros (CARVALHO et al., 2003). A aplicação de métodos computacionais para solucionar problemas químicos estende-se às áreas onde a química desempenha um importante papel, direto ou indireto.

Dentre os métodos computacionais, destacam-se os estudos de QSAR, metodologia importante para o planejamento de compostos protótipos, uma vez que, permitem uma análise rápida da atividade biológica versus propriedades

físico-químicas de uma série de moléculas de interesse (CARVALHO et al., 2003).

2.7 Relação Quantitativa Estrutura-Atividade (QSAR)

Estudos sobre a metodologia da relação quantitativa entre a estrutura química e a atividade biológica (QSAR), tem sido uma das principais estratégias utilizadas para prever as propriedades físico-químicas e biológicas (incluindo toxicológico) de novas moléculas, correlacionando com características estruturais e propriedades dos compostos que são codificados dentro de descritores por meio de equações matemáticas (CORREA-BASURTO, 2014; PRATIM ROY et al., 2009).

Os descritores moleculares predizem as propriedades das substâncias, classificam as estruturas químicas ou procuram similaridades entre elas. (ARROIO; HONORIO; SILVA, 2010). Tais propriedades estão associadas com características físico-químicas, eletrônicas ou mecânico-quânticas a parâmetros geométricos e topológicos das moléculas, dessa forma, podem ser correlacionadas diretamente com a atividade (MARTINS, 2010).

A metodologia de QSAR economiza recursos e agiliza o processo de desenvolvimento e planejamento de novos compostos biologicamente ativos (KHAN et al., 2009). Os modelos de QSAR devem ser capazes de prever a correta atividade de uma série congênere. Além disso, eles têm sido cada vez mais utilizados para fazer previsões sobre novos derivados de 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).

Atualmente, várias metodologias de QSAR têm sido propostas, QSAR-1D a QSAR-7D, cada uma caracterizando-se por diferentes abordagens de

cálculos, descritores moleculares e algoritmos estatísticos para a construção dos modelos resultantes (SANTOS-FILHO et al., 2009).

2.8 Evolução da Metodologia de QSAR

Muitos pesquisadores contribuíram para a fundação da metodologia de QSAR, relatos da literatura indicam que suas raízes desenvolveram-se ao longo de um século, por volta dos anos de 1860 a 1960 (KUBINYI, 2002). Segundo Kubinyi (2002), a primeira descrição sobre a relação entre as propriedades moleculares e biológicas foi relatada pelo Afa Cros, em 1863, tese realizada na Universidade de Estrasburgo. Cros observou que existia uma relação entre atoxicidade de álcoois alifáticos primários e a sua solubilidade em água, nos mamíferos (SCHULTZ, 2003). Após três anos, Brown e Fraser, estudando sobre os efeitos biológicos de certos alcalóides, chegaram à conclusão de que a atividade fisiológica deveria ser uma função da constituição química (KUBINYI, 2002). Uma evidência, de que era possível codificar estruturas químicas de uma maneira quantitativa, e, assim, de maneira implícita, foram surgindo vários trabalhos usando essa relação.

Uma grande contribuição para o avanço da metodologia de QSAR foram os estudos de Hammett, em 1930, em que correlacionava propriedades eletrônicas dos ácidos orgânicos e bases com as suas constantes de equilíbrio e reatividade, podendo ser expressa pela Equação (1) conhecida como Equação de Hammett.

$$\text{Log}K_x = \rho\sigma + \text{Log}K_0 \quad (1)$$

Onde K_x é a constante de ionização do ácido benzóico substituído; K_0 é a constante de ionização do ácido benzóico não substituído; σ é a constante do

grupo, mede a influência eletrônica de um dado grupo substituinte e ρ é a constante de reação (PÂRVU, 2003; TAVARES, 2004).

A constante σ de Hammett tornou-se um descritor capaz de caracterizar a atividade de muitos conjuntos de moléculas, análogo a esse parâmetro, Hansch definiu o parâmetro π , como sendo a relação logarítmica entre o coeficiente de partição de um composto substituído e o coeficiente de partição de seu análogo não substituído, com a finalidade de representar a contribuição hidrofóbica de um determinado substituinte.

$$\pi_x = \text{Log}P_X - \text{Log}P_H \quad (2)$$

Onde π_x é o parâmetro que reflete a contribuição hidrofóbica do grupo substituinte, X; P_X o coeficiente de partição do composto X substituído e P_H o coeficiente de partição do composto não substituído (TAVARES, 2004).

A constante de Hammett σ , o parâmetro de lipofilicidade π e o coeficiente de partição octanol/água ($\log P$), foram os descritores mais utilizados nas primeiras análises de QSAR (MARTINS, 2010) mais tarde, esses estudos ficaram conhecidos como a metodologia de QSAR-1D (LILL, 2007).

A metodologia de QSAR deve ser atribuída aos trabalhos de Hansch e Fujita (1964) os fundamentos para o modelo de QSAR, propostos por esses pesquisadores, consideraram critérios que afere a influência de cada uma das propriedades físico-químicas envolvidas na determinação da atividade biológica, combinando diferentes descritores estruturais em uma equação linear, equação (3), conhecida como Análise de Hansch.

$$\text{Log } 1/C = a\text{Log}P + \rho\sigma + b \quad (3)$$

Onde C é a concentração molar que produz determinado efeito biológico, $\text{Log } 1/C$ a potência biológica e a , ρ e b são coeficientes calculados pela análise de regressão para o conjunto de dados (TAVARES, 2004).

Com a introdução de modelos multiparamétricos na evolução dos estudos da Análise de Hansch, tornou-se possível descrever a relação da atividade biológica em função das propriedades físico-químicas de compostos bioativos, conhecidos como QSAR 2D (TAVARES, 2004).

Na década de 1980, os estudos de QSAR sofreram grandes transformações em virtude da introdução dos chamados parâmetros moleculares tridimensionais, que consideram a influência de diferentes confôrmeros, estereoisômeros ou enantiomêros. Esse tipo de modelo ficou conhecido como QSAR 3D (MARTINS, 2010). O primeiro modelo publicado, possuindo essas características, foi o método de análise comparativa de campo molecular (CoMFA), proposta por Cramer et al. (1988). Esse método, que é baseado em campo molecular, constituiu o primeiro verdadeiro método de QSAR 3D (KUBINYI, 1997). Em contraste com a análise de Hansch, CoMFA é mais adequado para descrever as interações ligante-receptor, uma vez que considera as propriedades dos ligantes na sua suposta conformação bioativa (POTEMKIN; GRISHINA, 2008). Como o resultado de uma análise CoMFA, regiões do espaço são identificadas como sendo favoráveis ou desfavoráveis para a interação ligante-receptor (KUBINYI, 1997).

Outros métodos de QSAR-3D também são conhecidos como, por exemplo, o QSAR-4D desenvolvido em 1997, por Hopfinger e colaboradores (HOPFINGER et al., 1997). Essa metodologia incorpora a flexibilidade molecular e a liberdade de alinhamento (CUNHA et al., 2005). Além disso, a metodologia QSAR-4D é baseada no alinhamento de moléculas em uma caixa virtual que é utilizada para gerar um conjunto de descritores, entretanto, quando essa metodologia é aplicada a certos compostos sem o receptor, esse estudo é

denominado QSAR-4D independente do receptor. Ao ser levado em consideração o receptor, fala-se em QSAR-4D dependente do receptor. (YANMAZ et al., 2011).

A grande vantagem da metodologia de QSAR-4D frente aos demais métodos de QSAR-3D mais divulgado, por exemplo o CoMFA, é o emprego da amostragem conformacional (PITA, 2006). O formalismo do método de QSAR-4D será detalhado no item de Metodologia.

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; DOBLER; LILL, 2005). Posteriormente, com a necessidade de estudar diferentes cenários de solvatação, foi desenvolvida a metodologia do QSAR-6D (VEDANI et al., 2007). Finalmente, o método mais recentemente desenvolvido, o QSAR-7D, que inclui modelos virtuais de interação ligante-receptor (POLANSKI, 2009).

2.9 Simulação de Dinâmica Molecular

Uma das etapas da metodologia do QSAR-4D consiste no emprego de Simulações por Dinâmica Molecular. A DM é uma ferramenta computacional que permite simular os movimentos dos átomos individuais que compõem o sistema durante um determinado tempo de simulação. Ela é considerada uma ampliação da Mecânica Molecular, no qual as moléculas são tratadas como uma coleção de partículas mantidas unidas por forças newtonianas. Para realizar uma simulação de DM é fundamental conhecer as equações de movimento que governam a dinâmica das partículas e o potencial de interação entre essas partículas, descrito como “campo de força” (NAMBA; SILVA; SILVA, 2008).

O campo de força, representado na equação 4, é uma função de energia potencial, o qual permite que a energia potencial total do sistema, $V(r)$, seja medida a partir da estrutura tridimensional do sistema. $V(r)$ é aferido como a soma de inúmeros termos de energia, englobando os termos para átomos ligados, como os comprimentos e ângulos de ligação, ângulos diedros e os termos para átomos não-ligados, como as interações de van der Waals e de Coulomb (NAMBA; SILVA; SILVA, 2008).

$$V(r) = \sum V_L + \sum V_\theta + \sum V_\phi + \sum V_{vdW} + \sum V_{elect} \quad (4)$$

Onde V_L representa a energia de estiramento da ligação, V_θ representa a energia de deformação do ângulo de ligação, V_ϕ representa a energia, decorrente da torção em torno de uma ligação, V_{vdW} é a energia das interações de van der Waals e V_{elect} representa as energias de atração ou repulsão eletrostática entre duas cargas.

Geralmente, os campos de forças são desenvolvidos de modo independente e com todos os conjuntos de parâmetros específicos e bem definidos. Alguns campos de força, com o intuito de se obter uma melhor concordância com espectros vibracionais, incluem outros termos para descrever as ligações de hidrogênio ou para acoplar oscilações entre ângulos e comprimentos de ligação. No caso de sistemas como proteínas, os campos de força mais utilizados são CHARMM (Chemistry at Harvard Macromolecular Mechanics) (BROOKS et al., 2009), AMBER (Assisted model Building with Energy Refinement) (CASE et al., 2005), MM2 (WEINER et al., 1986) dentre outros (NAMBA; SILVA; SILVA, 2008).

Para iniciar as simulações de DM, o sistema de interesse deve sofrer uma otimização em sua geometria com o propósito de encontrar um conjunto de coordenadas que minimizam a energia potencial e, assim, eliminando quaisquer

maus contatos entre os átomos. Em seguida, é atribuída ao sistema uma temperatura de interesse, onde as velocidades iniciais de todas as partículas fundamentam-se na distribuição de Maxwell-Boltzman. Geralmente, dá-se por concluído quando o sistema em estudo alcança o equilíbrio termodinâmico. Assim, podendo então gerar as trajetórias da DM (NAMBA; SILVA; SILVA, 2008). A partir da trajetória, é possível extrair várias propriedades como, por exemplo, energia, número médio de ligações de hidrogênio formadas ao longo da simulação, distância média entre átomos, raio de giro, dentre outras (OLIVEIRA, 2012; SPOEL et al., 2005).

O emprego de DM com outras técnicas computacionais, como os cálculos de energia livre de ligação, Ancoramento Molecular e até mesmo QSAR, tem sido fundamentais para a compreensão das interações ligante-proteína, direcionando o planejamento racional de novos e potenciais inibidores (NAMBA; SILVA; SILVA, 2008).

3 OBJETIVOS

No presente trabalho, visa-se ao estudo entre a estrutura química e a atividade biológica de uma série de noventa e sete análogos de imidazopiridinas sintetizados e testados farmacologicamente por Blomgren e colaboradores (BLOMGREN et al., 2011), inibidores da enzima Syk.

A análise dos dados será realizada aplicando o método de QSAR-4D (*em inglês, Four-Dimensional Quantitative Structure Activity Relationship*) independente do receptor para cada um dos noventa e sete compostos estudados, serão averiguados os melhores modelos matemáticos resultantes, gerados pela frequência de ocupação dos átomos mais importantes para a interação, com o objetivo de identificar características que possam aumentar o poder inibitório desses compostos.

4 METODOLOGIA

O trabalho foi realizado no Laboratório de Modelagem Molecular do Departamento de Química da Universidade Federal de Lavras.

4.1 Programas

O presente trabalho foi desenvolvido, utilizando QSAR-4D (4D-QSAR..., 1997) e MOLSIM 3.0 (MOLSIM..., 1997), com campo de força MM2 (WEINER et al., 1986).

4.2 Dados Biológicos

No Quadro 1, estão relacionados os noventa e sete compostos de imidazopiridinas que foram utilizados neste trabalho e seus respectivos valores de constante de inibição (pIC_{50}). As atividades biológicas desses compostos são relatadas como a concentração capaz de inibir 50% da atividade da enzima (IC_{50}). Os valores de IC_{50} (μM) foram convertidos em unidades molar, e, posteriormente, expressas em unidades negativas de logaritmo (pIC_{50}). Esses compostos foram sintetizados e avaliados farmacologicamente por Blomgren e colaboradores (BLOMGREN et al., 2011), a estrutura química desses compostos comum a todos, está representada na Figura 7. Ademais, é importante ressaltar que todos os dados foram obtidos do mesmo laboratório, eliminando quaisquer erros caso fossem provenientes de fontes diferentes.

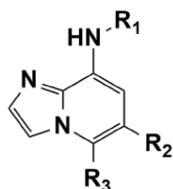
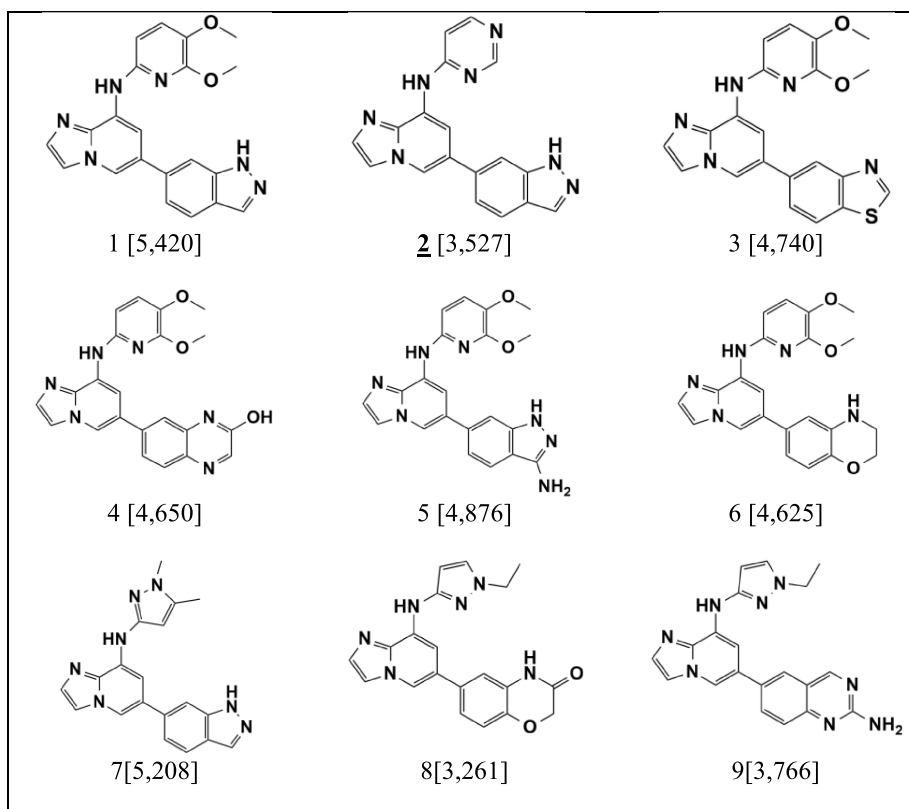


Figura 7 Estrutura geral dos compostos desenvolvidos por Blomgren e colaboradores

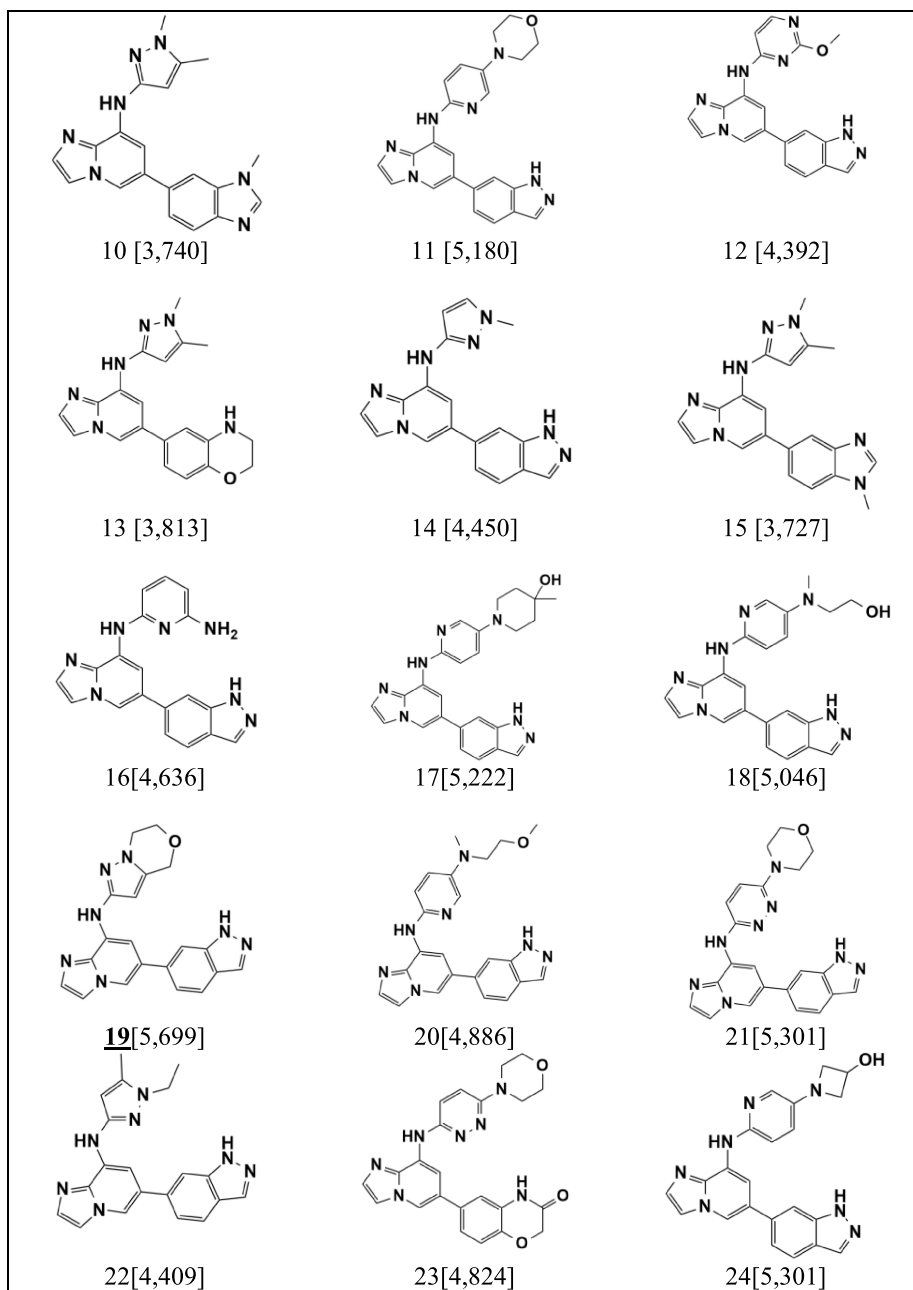
Fonte adaptado de Blomgren et al. (2011).



Quadro 1 Estrutura química e os respectivos valores de $pIC_{50}(M)$ dos 97 compostos imidazopiridínicos. Em **negrito e sublinhado** estão representadas as moléculas selecionadas para compor o grupo teste, as demais moléculas, participam do grupo de treinamento da metodologia QSAR-4D

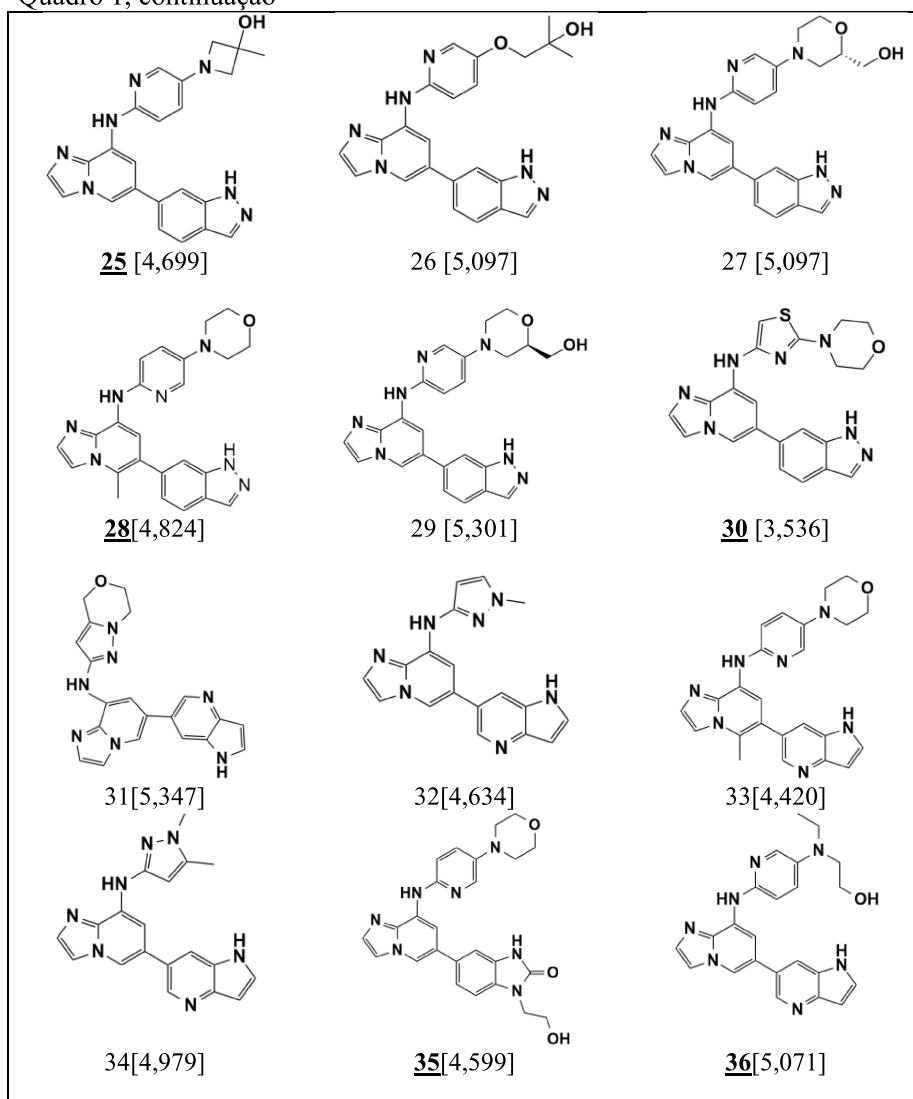
(...continua...)

“Quadro 1,continuação”



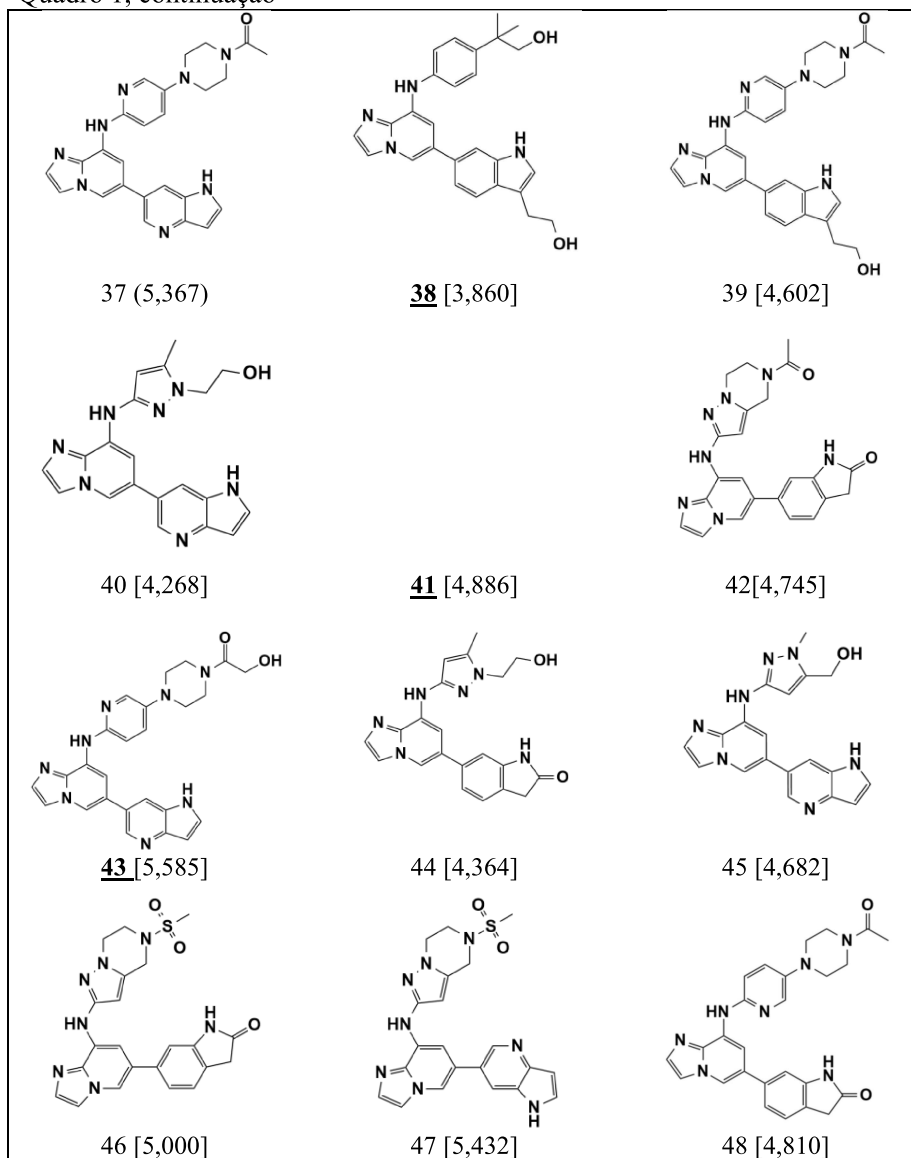
(...continua...)

“Quadro 1, continuação”



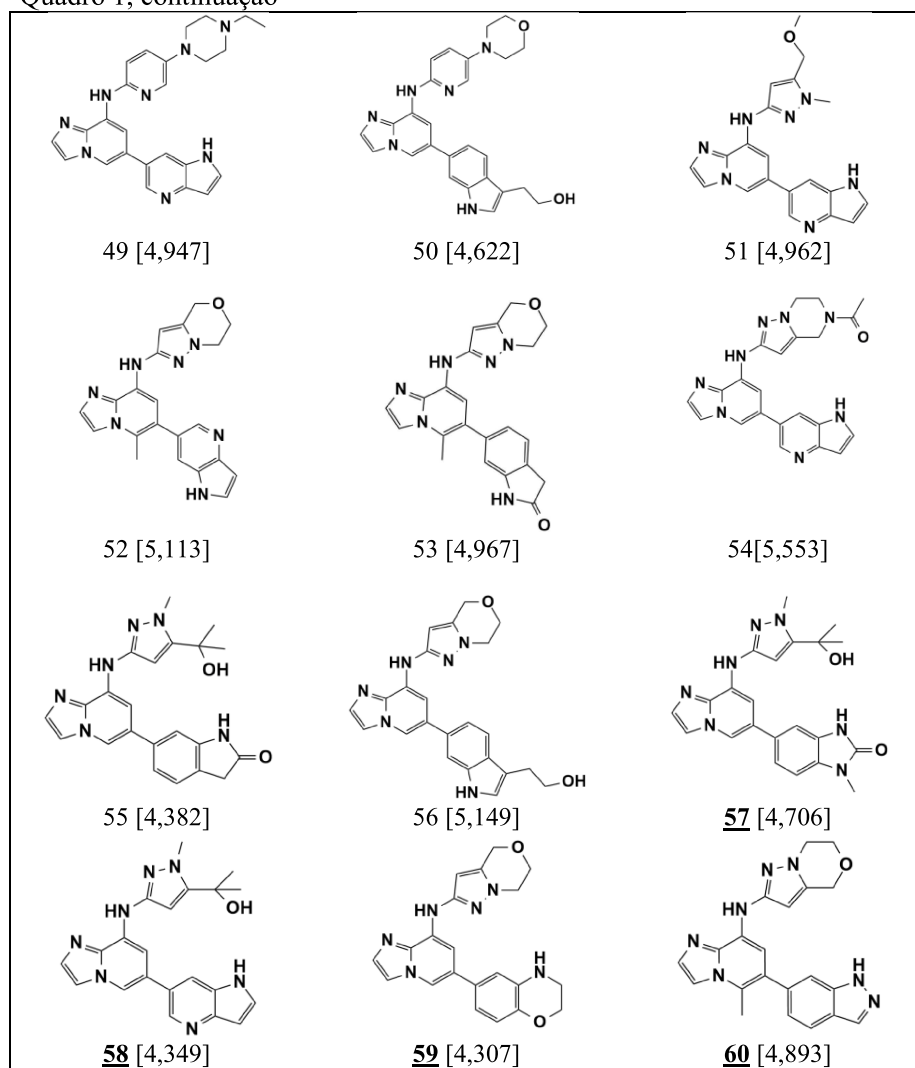
(...continua...)

“Quadro 1, continuação”



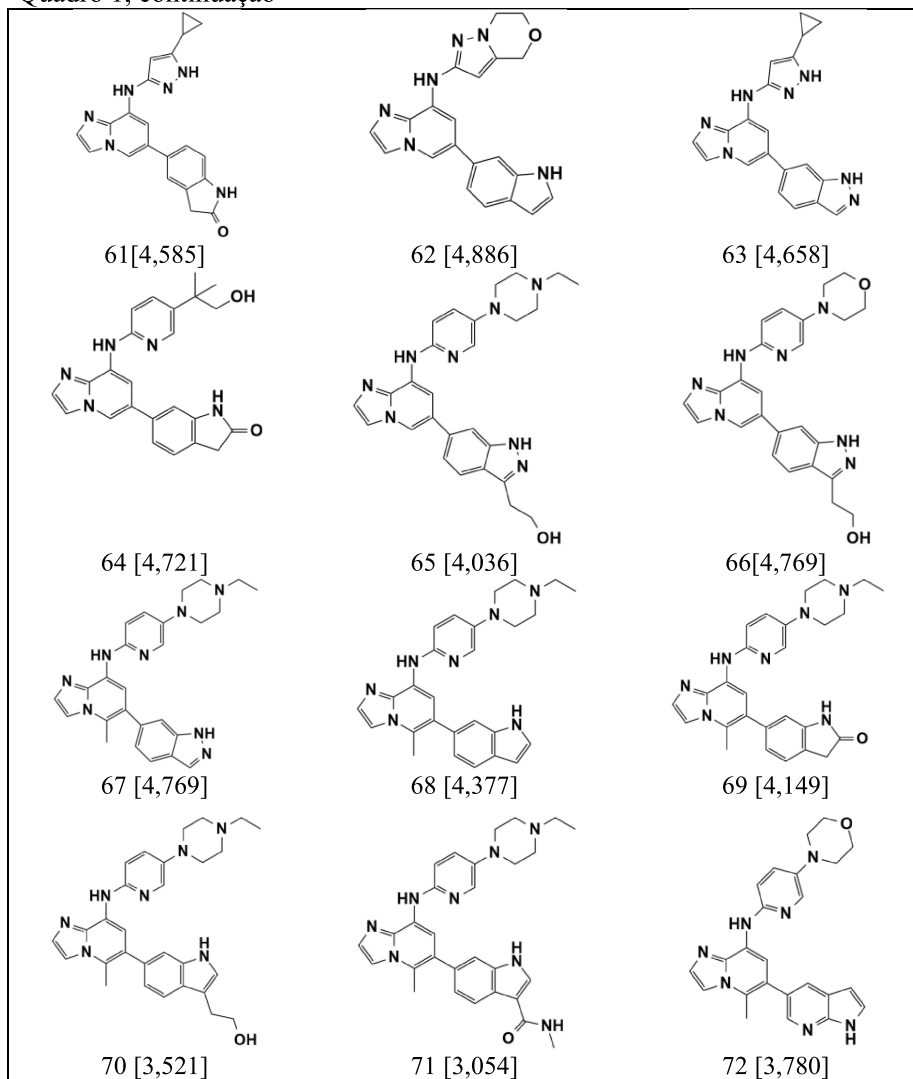
(...continua...)

“Quadro 1, continuação”



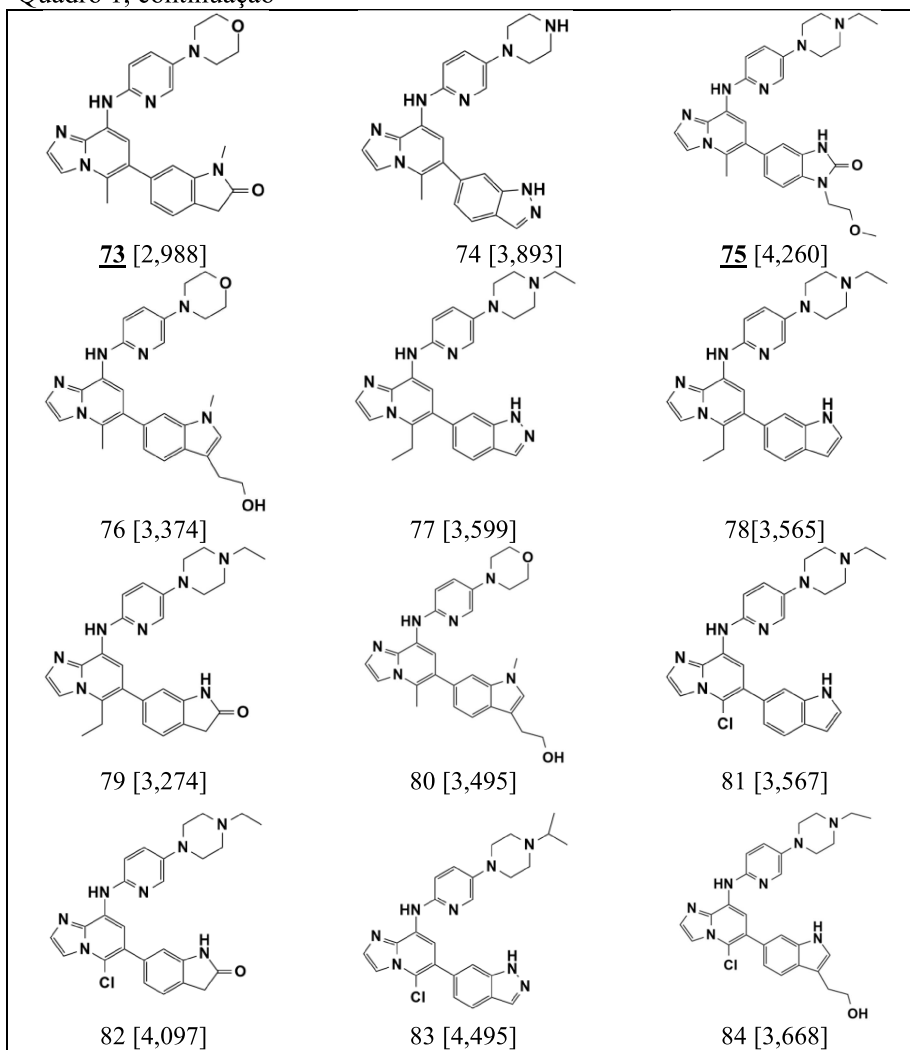
(...continua...)

“Quadro 1, continuação”



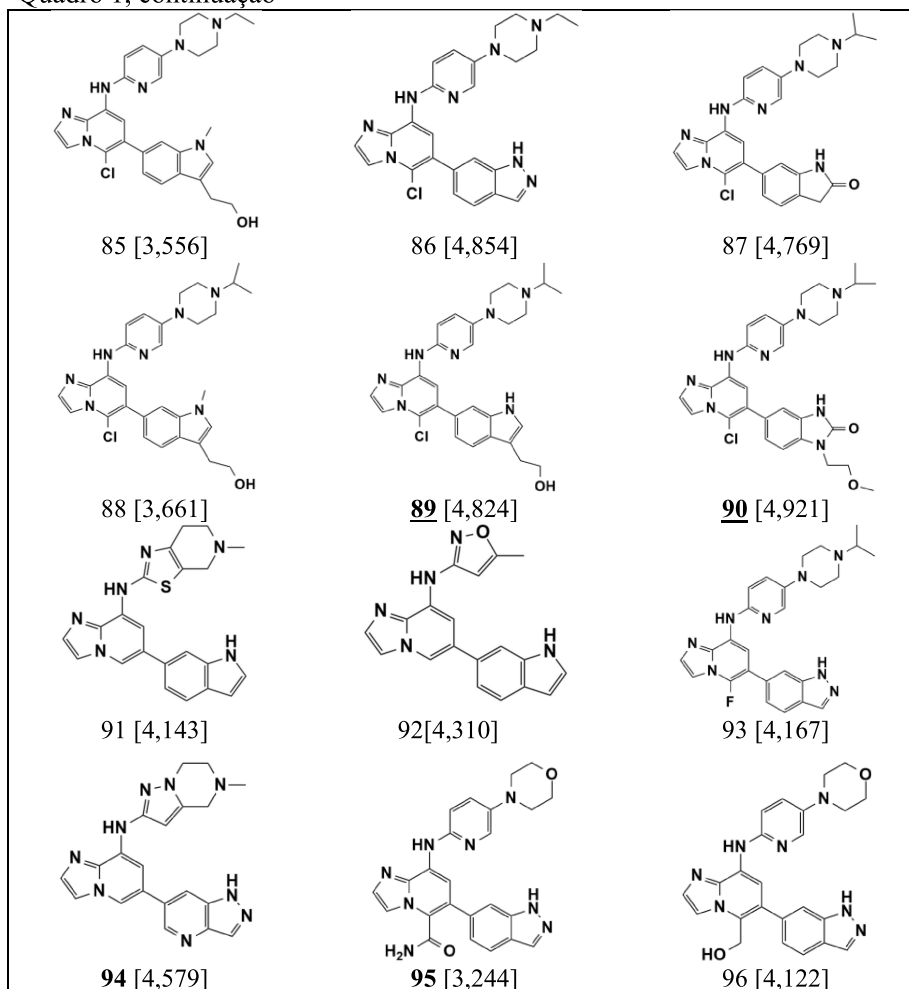
(...continua...)

“Quadro 1, continuação”



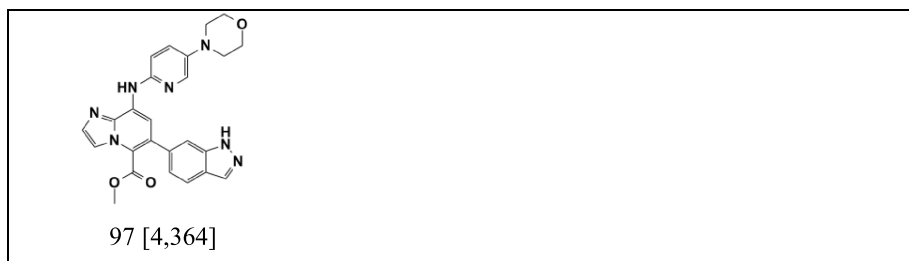
(...continua...)

“Quadro 1, continuação”



(...continua...)

“Quadro1, conclusão”



No gráfico 1, representa-se a distribuição de valores de pIC_{50} desses compostos, que abrange uma faixa em torno de 2,9 a 5,7, e os valores de atividade biológica apresentam uma distribuição regular em toda a extensão.

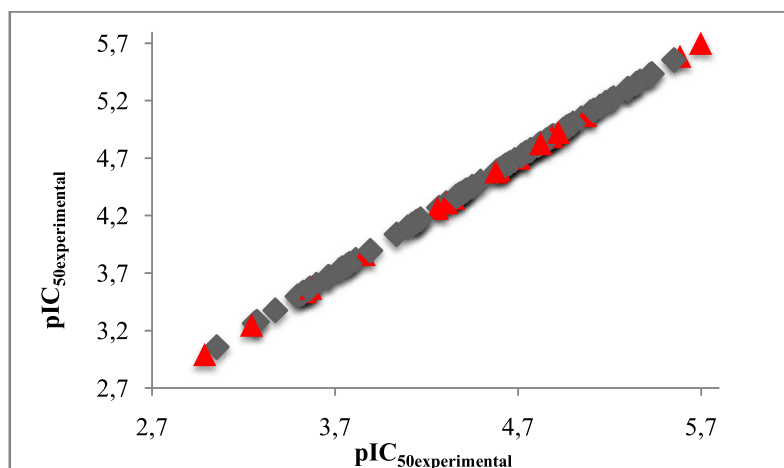


Gráfico 1 Distribuição dos valores de $pIC_{50\text{experimental}}$ para os compostos do conjunto de treinamento e de teste

Nota: Os pontos que estão na cor cinza representam os compostos do conjunto de treinamento e os pontos que estão em vermelho representam os compostos do conjunto de teste.

4.3 Construção e Otimização das Estruturas Tridimensionais

As estruturas tridimensionais de cada um dos noventa e sete compostos relatados no Quadro 1, foram construídas com base na estrutura do composto 0VG (Figura 8), cocrystalizado com a enzima Syk, código:4FZ6 com resolução 1,85Å (LUCAS et al., 2012), obtida a partir do PDB (em inglês, Protein Data Bank) (BERMAN et al., 2000), usando o programa Discovery Studio (ACCELRY'S SOFTWARE, 2012). É importante destacar que essa estrutura é a conformação ligada ao sítio ativo da enzima (CUNHA; MATOS; RAMALHO, 2013).

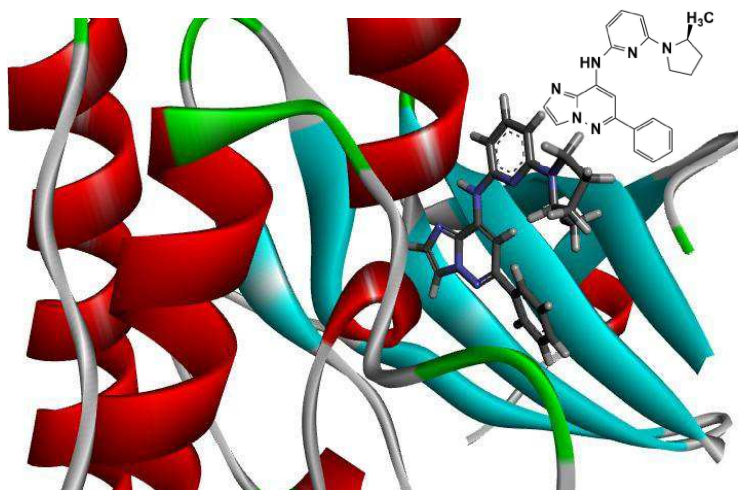


Figura 8 Representação da estrutura do ligante 0VG, N-{6 - [(2S)-2-metilpirrolidin-1-il] piridin-2-il}-6-fenilimidazo[1,2-b]piridazin-8-amina, cristalizado no sítio ativo da enzima Sky

Foram otimizadas as estruturas tridimensionais dos compostos e calculadas as cargas atômicas parciais, via mecânica molecular, utilizando o campo de força CHARMM+ (BROOKS et al., 2009), do programa Discovery Studio (ACCELRY'S SOFTWARE, 2012).

4.4 Metodologia do QSAR-4D: Divisão do conjunto de treinamento e de teste

Os modelos QSAR-4D foram desenvolvidos, utilizando 77 compostos para o conjunto treinamento dedicado à calibração do modelo, e validadas externamente, usando 20 compostos chamados de conjunto teste, selecionados a partir de uma série de compostos derivados de imidazopiridinas que foram sintetizados por Blomgren e colaboradores. Os 97 compostos (treinamento e teste) estão listados no Quadro 1. Uma vez que os inibidores da série são estruturalmente relacionados, o conjunto de treinamento e conjunto de teste foi dividido da seguinte forma: os 97 compostos foram distribuídos em três subgrupos de acordo com seus valores crescentes de pIC_{50} (Grupo I: $pIC_{50} = 2,5-3,5$; Grupo II: $3,6-4,5$; Grupo III: $pIC_{50} = 4,6-5,7$) e 20% dos compostos de cada subgrupo foi selecionado aleatoriamente como um composto representativo do conjunto teste. Assim, os 77 compostos restantes abrangem o conjunto de treinamento. Usando essa abordagem, temos compostos no conjunto de teste que compreendem toda a extensão dos valores das atividades biológicas (CUNHA et al., 2005).

4.5 Simulação por Dinâmica Molecular

Na metodologia QSAR-4D, é importante prevê o estado conformacional de mais baixa energia do conjunto de amostragem para cada composto (PITA, 2006). Com esse intuito, aplicou-se a Simulação por Dinâmica Molecular (SDM), fundamentada nos princípios da Mecânica Clássica, que fornece informações sobre o comportamento dinâmico de um sistema molecular.

Nesse contexto, os noventa e sete compostos tridimensionais neutros foram submetidos ao processo de SDM, a fim de construir um perfil de amostragem conformacional (CEP, do inglês Conformational Ensemble Profile) de cada ligante. Nesse contexto, todas as estruturas resultantes da SDM foram usadas para construir os modelos de QSAR-4D (HOPFINGER et al., 1997).

As 97 estruturas do conjunto amostral foram submetidas à SDM, usando o programa Molsim com o campo de força MM2 (HOPFINGER; PEARLSTEIN, 1984), incorporado ao programa QSAR-4D. A temperatura para cada SDM foi de 300K, próxima à temperatura usada nos ensaios farmacológicos e o tempo de simulação foi realizado em 100 ps, integrados em intervalos de 0,001ps, gerando um total de 1000 conformações para cada composto. Os cálculos de DM foram executados, aplicando uma constante dielétrica dependente da distância, $\frac{1}{4} D * r_{ij}$, que foi criada a $3 * r_{ij}$, com o intuito de tentar modelar o efeito do solvente (SILVA; RAMALHO; CUNHA, 2014).

Além disso, foram fixados três átomos de carbono (Figura 9), os átomos abc, comum a todos os compostos com o objetivo de evitar uma grande mudança conformacional dos ligantes na ausência da estrutura da proteína.

4.6 Definição dos elementos de interação farmacofórica

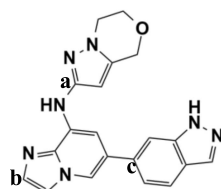
De acordo com a metodologia de QSAR-4D, os descritores são gerados a partir de diferentes tipos de interações dos átomos, que são classificados em sete tipos de elementos de interação farmacofórica (IPE), correspondentes aos tipos de átomos que podem ser classificados, dependendo do tipo de interação que ele é capaz de realizar (CUNHA et al., 2005). Neste trabalho, considerou-se os seguintes elementos de interação farmacofórica: não polares (np), polares com densidade de carga positiva (p+), polares com densidade de carga negativa

(p-), aceptores de ligação de hidrogênio (hba), doadores de ligação de hidrogênio (hbd) e aromáticos (ar) (SILVA; RAMALHO; CUNHA, 2014).

4.7 Definição do alinhamento e tamanho das células

O alinhamento é uma das etapas mais importantes da metodologia de QSAR-4D, pois é nessa fase que os compostos do conjunto de treinamento são comparados, assim a qualidade do modelo gerado depende do alinhamento testado (GARCIA, 2014). A escolha dos átomos para o alinhamento é fundamentada pela seleção de átomos conservados ou similares em grupos funcionais presentes em todas as estruturas selecionadas (PITA, 2006). Dessa forma, como os compostos imidazopiridínicos são congêneres, podemos considerar que esses compostos se ligam de maneira semelhante à Syk e, por isso, escolhemos três átomos de carbono comuns a todos os compostos para compor o alinhamento, apresentados na Figura 9.

Além disso, no alinhamento, é levada em consideração a ordem dos átomos escolhidos, visto que durante o processo o primeiro átomo selecionado é posicionado na origem dos três eixos cartesianos e a molécula é rodada de modo que o segundo átomo se posicione no eixo x, e o terceiro se posicione no eixo y (GARCIA, 2014).



Alinhamento	1 ^a átomo	2 ^a átomo	3 ^a átomo
1	a	b	c
2	a	c	b
3	b	a	c
4	b	c	a
5	c	a	b
6	c	b	a

Figura 9 Três letras (a,b,c) ordenando os átomos de carbonoutilizados nos estudos de QSAR-4D definindo seis alinhamentos testados. Composto 19 ($pIC_{50\text{experimental}}=5,69$), o inibidor mais potente da série de Blomgren e colaboradores, é usado para definir o código de letra do átomo

O CEP de cada ligante gerado na DM foi inserido em uma caixa virtual levando em consideração os distintos alinhamentos realizados. A caixa é formada por células cúbicas que podem ter dimensões variadas, sua função é registrar a distribuição de ocupação espacial de cada átomo de cada ligante do conjunto amostral. As medidas de ocupação de cada célula cúbica da caixa virtual, por diferentes tipos de átomos, forma o conjunto de descritores de ocupação das células, GCODs (em inglês, Grid Cell Occupancy Descriptors), que são as variáveis independentes usadas na construção dos modelos de QSAR-4D. No presente estudo, foram usadas células de 1,0 Å de aresta, visto que quanto menor o tamanho da célula tem-se modelos mais refinados em termos dos tipos de átomos (IPE) selecionados (SILVA; RAMALHO; CUNHA, 2014).

4.8 Construção do modelo QSAR-4D: Redução de dados e abordagem PLS-GA

Geralmente, a metodologia de QSAR-4D produz um número muito grande de descritores, em razão da quantidade de células cúbicas usadas e por causa dos IPEs (ROMEIRO et al., 2005). Ademais, pode haver descritores correlacionados o que complica a compreensão e a validação dos modelos resultantes (PITA et al., 2012). Com o objetivo de reduzir essas variáveis (descritores) que são prejudiciais para os modelos, identificando e selecionando apenas os GCODs altamente ponderados, necessários para a etapa subsequente do QSAR-4D, realizaram-se os seguintes procedimentos como pré - tratamento do banco de dados:

- a) Emprego da regressão dos mínimos quadrados parciais (PLS, em inglês *Partial Least Squares*); essa ferramenta seleciona os descritores que aparecem com maior frequência (de maior peso) para a variação da atividade biológica.
- b) Posteriormente, os GCODs que apresentavam variância igual a zero para todos os compostos foram excluídos.

Após esses procedimentos, os GCODs que sobreviveram foram otimizados usando uma abordagem que combina métodos de algoritmos genéticos (GA) e PLS, conhecidos como aproximação de funções genéticas GFA (em inglês *Genetic Function Approximation*), implementados no programa QSAR-4D (ROGERS; HOPFINGER, 1994). No GFA, os modelos iniciais são formados por acaso, selecionando algumas variáveis independentes, GCODs, para criação de uma população de modelos. Nessa etapa, são realizados processos de cruzamento e mutação nos modelos, que, combinando os melhores

termos, gera modelos melhores (ROMEIRO et al., 2005; SENESE; HOPFINGER, 2003). A grande vantagem do GFA é a construção e otimização de múltiplos modelos, ao invés da otimização de apenas um único modelo (ROGERS; HOPFINGER, 1994).

No presente estudo, os cálculos de GFA foram iniciados utilizando 100 modelos. Em seguida, foram aplicadas 50000 operações de cruzamento, com uma probabilidade de mutação de 10%. O *smoothing factor* de 3,0 foi ajustado para obter equações com, no máximo, 15 descritores, com o intuito de evitar um possível superajuste nos dados (MANNHOLD; KROGSGAARD-LARSEN; TIMMERMAN, 1993), resultando em menores valores de LOF (Equação 4), fator de desajuste de Friedman (do inglês, lack-of-fit) (ROGERS; HOPFINGER, 1994), que é a medida dos erros dos mínimos quadrados (LSE, do inglês least-squares error) penalizados, ou seja, quando dois modelos têm o valor de LSE iguais, aquele com menor número de variáveis independente tem melhor valor de LOF (PITTA, 2006; ROMEIRO et al., 2005).

$$\text{LOF} = \frac{\text{LSE}}{\left(1 - \left(\frac{c+d \cdot p}{M}\right)\right)^2} \quad (5)$$

Índice de falta de ajuste, onde **C** é o número de termos no modelo, **d** é *smoothing factor*, controla o numero de variáveis do modelo, **p** é o número total de variáveis, contido em cada termo e **M** é o número de compostos, LSE (*least squares error*) erro dos mínimos quadrados.

4.9 Validação dos modelos de QSAR-4D

Os dez modelos resultantes da otimização do GA-PLS, de acordo com seus valores de LOF, foram validados por três diferentes critérios: a validação cruzada ou interna, a validação externa e o teste *Y-randomization*.

No programa QSAR-4D, o método de validação cruzada é conhecido como 'leave one out'. Nesse método, cada composto do conjunto treinamento é eliminado por vez, então, é prevista a variável dependente pelo modelo construído para o composto que foi deixado de fora do modelo, e esse processo é repetido até que todos os compostos tenham ficado de fora da análise uma vez. Esse procedimento é bastante importante, pois gera um parâmetro estatístico denominado coeficiente de correlação quadrática da validação cruzada, q^2 (Tabela 1), que prediz a significância e a robustez do modelo construído (MARTINS, 2010). Um bom modelo apresentará valores de q^2 acima de 0,5 (FERREIRA; MONTANARI; GAUDIO, 2002). Neste presente estudo, o coeficiente q^2 foi convertido em $q^2_{ajustado}$ (Tabela 1), a fim de comparar a capacidade preditiva de modelos com números diferentes descritores. Para avaliar o grau de ajuste do modelo obtido, gerado pelo conjunto de treinamento, foi utilizado o cálculo do coeficiente determinação (r^2). Neste estudo, usamos o $r^2_{ajustado}$ que permite comparações entre modelos com número diferente de variáveis independentes. Um valor de r^2 acima de 0,7 indica que o modelo é bem correlacionado (CALDAS; RAMALHO; CUNHA, 2014).

A validação externa tem por finalidade enriquecer o modelo no sentido de comprovar a sua capacidade de prever valores de atividades biológicas de compostos não incluídos no conjunto de treinamento (LIMA NETO et al., 2006; ROMEIRO et al., 2005;). O parâmetro estatístico calculado neste processo é o coeficiente de correlação quadrático predito da validação externa, R^2_{pred} (Tabela 1) (PRATIM ROY et al., 2009).

Outra abordagem de validação é o *Y-randomization*, é uma técnica fundamentada em repetir o procedimento de construção do modelo com as atividades randomizadas e, posteriormente, realizar uma avaliação estatística dos modelos obtidos. Essa técnica tem por objetivo verificar se a atividade biológica está definitivamente correlacionada com as variáveis independentes dessa validação, dessa forma é estimado que o r^2 do modelo seja maior que o r^2 da randomização (GRAMATICA, 2007; WEBER, 2008).

Tabela 1 Parâmetros Estatísticos utilizados na validação de QSAR-4D

Parâmetro	Definição
$r^2 = 1 - \frac{\sum(Y - Y_{pred})^2}{\sum(Y - \langle Y \rangle)^2}$	Coefficiente de correlação
$r_{ajustado}^2 = r^2 - \left(\frac{k-1}{M-k}\right)(1-r^2)$	Coefficiente de correlação ajustado
$RMSE_c = \sqrt{\frac{\sum(Y - Y_{pred})^2}{M}}$	Erro da calibração
$Q_{cv}^2 = 1 - \frac{\sum(Y_{pred} - Y)^2}{\sum(Y - \langle Y \rangle)}$	Coefficiente de correlação da validação cruzada
$Q_{CVajustado}^2 = \frac{1 - (1 - Q_{cv}^2)}{\frac{M-1}{M-p}}$	Coefficiente de correlação da validação cruzada ajustado
$R_{pred}^2 = 1 - \frac{\sum(Y_{pred}(teste) - Y_{teste})^2}{\sum(Y_{teste} - \langle Y \rangle_{treinamento})}$	Coefficiente de correlação predito da validação externa
$RMSE_p = \sqrt{\frac{\sum(Y - Y_{pred})^2}{M}}$	Erro predito da validação externa
$r_{m(teste)} = r^2 \left(1 - \sqrt{r^2 - r_0^2}\right)$	Coefficiente de correlação do conjunto de teste.

Onde M é o número de compostos do conjunto de treinamento, LSE (*least squares error*) erro dos mínimos quadrados, Y é o valor experimental da atividade biológica e Y_{pred} é o valor calculado pelo modelo, k número de variáveis independentes, p variáveis latentes, r_0^2 representa o coeficiente de correlação ao quadrado, entre os valores de atividade predita e atividade experimental do conjunto de teste interceptando o eixo 0.

4.10 Identificação dos compostos *outliers*

Os compostos chamados como *outliers*, são compostos que não foram bem calculados pela equação e definidos como aqueles, cujos valores modulares da diferença entre pIC_{50} experimental e predito são superiores ao dobro do desvio padrão desses resíduos (HOPFINGER et al., 1997).

4.11 Seleção da conformação “representativa”

A etapa final da metodologia QSAR-4D consiste em uma conformação representativa que pode ser definida como o estado confôrmero de mínima energia do conjunto de amostragem para cada ligante, segundo o método de QSAR-4D essa conformação maximiza a potência do composto no modelo selecionado. Isso é alcançado pela primeira identificação de todos os estados confôrmero mostrados para cada composto, que estão dentro de um intervalo de energia (ΔE). Neste estudo, a conformação foi selecionada dentro de um valor de 5kcal.mol^{-1} da conformação de menor energia (GARCIA, 2014).

5 RESULTADOS E DISCUSSÃO

O conjunto de dados foi baseado em compostos derivados de imidazopiridinas, listados no Quadro 1, com suas respectivas atividades biológicas (BLOMGREN et al., 2011). As atividades dessas moléculas foram analisadas, utilizando a enzima Syk humana purificada em ensaios *in vitro*.

5.1 Análise de simulação por Dinâmica

No estudo de QSAR-4D, a Simulação de DM como descrito anteriormente, foi utilizada para avaliar o espaço conformacional de cada um dos noventa e sete compostos, resultantes dos processos de otimização de geometria. Nessa etapa, gerou-se um gráfico (Gráfico 2), que representa a variação de energia total em função do número de passos utilizados na dinâmica molecular referente ao composto 19, mais ativo da série estudada.

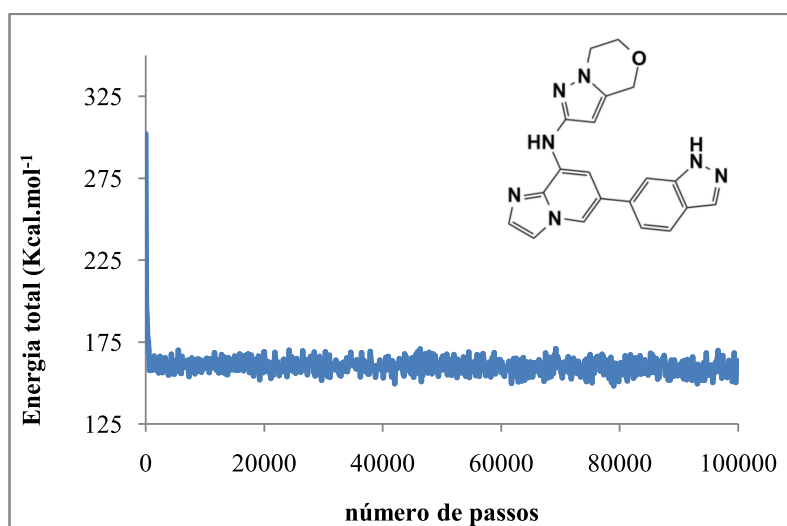


Gráfico 2 Representa os valores de energia total para DM de 100.000 ps referente ao composto 19

Observa-se, neste gráfico, que a energia total se mantém estável durante quase toda simulação, e que a maior energia encontrada foi no início da simulação.

5.2 Seleção dos melhores alinhamentos de QSAR-4D

Conforme descrito anteriormente, as 1000 conformações de cada molécula geradas na etapa de simulação por DM foram sobrepostas, utilizando seis alinhamentos. É importante destacar que realizaram vários alinhamentos, entretanto, somente os alinhamentos que obtiveram melhores valores dos parâmetros estatísticos são os que foram apresentados neste estudo. Como mostrado na Figura 9, os seis alinhamentos testados na metodologia de QSAR-4D-IR são: (1) átomos a-b-c, (2) átomos a-c-b, (3) átomos b-a-c, (4) átomos b-c-a, (5) átomos c-a-b e (6) átomos c-b-a.

Na análise do GFA, foram gerados 10 modelos ou equações para cada alinhamento, esses modelos são ordenados pelo valor de LOF, que é o valor do fator de desajuste de Friedman, contudo, será apresentada apenas a análise do melhor modelo decorrente de cada alinhamento. Os modelos QSAR-4D para alinhamento 1, 2, 3, 4, 5 e 6 (Modelo 1, Modelo 2, Modelo 3, ..., respectivamente) são definidos pelas equações:

Alinhamento 1 $pIC_{50} = 5,06 - 2,10 (0,6,-2, np) - 2,65 (0,3,-3,np) - 1,15 (0,10,3,np)$ (6)
 $- 1,25(2,1,-8,np) - 1,21(-1,9,0,np) - 3,29 (0,-3,-5,np)$
 $- 1,11 (0,0,-3,alh) - 2,44 (0,9,-2,np) + 0,87 (1,-1,-4,np)$
 $- 0,93 (-1,6,3,np) - 1,40 (0,1,-3,np)$

Modelo 1	N	r^2	r^2_{ajust}	RMSEC	q^2	q^2_{ajust}	LSE	LOF	r^2_p	RMSEP	r_m^2	Y-rand
	77	0,75	0,71	0,29	0,62	0,60	0,10	0,23	0,68	0,20	0,616	0,245

Alinhamento 2 $pIC_{50} = 4,44 - 4,06 (0,5,7,np) - 1,78 (0,-3,13,np) - 3,20 (2,9,3 np)$ (7)
 $- 1,49 (1,3,10,np) + 0,73(0,9,4,np) + 0,63(-1,9,5,np)$
 $- 0,72(1,10,6,np) + 1,53(1,10,5,np) - 1,00(0,3,3,alh)$
 $- 0,669(0,-2,5,np) + 0,39(0,-2,10)$

Modelo 2	N	r^2	r^2_{ajust}	RMSEC	q^2	q^2_{ajust}	LSE	LOF	r^2_p	RMSEP	r_m^2	Y-rand
	77	0,73	0,69	0,30	0,66	0,64	0,11	0,25	0,64	0,22	0,634	0,239

Alinhamento 3 $pIC_{50} = 4,61 - 1,58 (-1,6,7,1) - 3,58 (-1,3,8,np) - 1,99 (-2,1,14,np)$ (8)
 $3,10,np) - 2,30(0,8,5,np)$
 $- 1,13(0,-2,5,np) + 0,73(-1,-1,11,np) + 0,73(0,-2,10,np)$

Modelo 3	N	r^2	r^2_{ajust}	RMSEC	q^2	q^2_{ajust}	LSE	LOF	r^2_p	RMSEP	r_m^2	Y-rand
	77	0,70	0,66	0,31	0,64	0,62	0,12	0,23	0,59	0,23	0,575	0,195

Alinhamento 4 $pIC_{50} = 4,51 + 3,26 (2,-8,11,np) - 7,19 (2,-7,8,np) - 2,90(1,-4,4,np)$ (9)

$-2,16 (0,-6,-2,np) - 1,93 (0,-4,1,np) - 1,18 (0,0,6,ar)$
 $+0,64 (2,-4,8,np) + 0,95 (0,-4,9,np) + 0,78 (-1,-3,9,np)$
 $-0,65 (0,-4,-3,np) - 0,59 (1,-2,1,np)$

Modelo 4	N	r ²	r ² _{ajust}	RMSEC	q ²	q ² _{ajust}	LSE	LOF	r ² _p	RMSEP	r _m ²	Y-rand
	77	0,69	0,65	0,33	0,60	0,58	0,12	0,28	0,66	0,21	0,638	0,312

Alinhamento 5 $pIC_{50} = 3,68 + 3,48 (-1,-8,-6,np) - 2,85 (-1,-7,-4,np) - 2,49 (2,-3,0,np)$ (10)

$+ 1,52 (0,-4,-3,np) - 1,11(-1,-4,5,np) - 0,71(0,-1,0,alh)$
 $+0,39 (-1,-1,-3,np) + 1,04 (1,-3,-4,np) + 0,75 (0,5,5,np)$
 $+ 0,66 (0,-1,8,alh)$

Modelo 5	N	r ²	r ² _{ajust}	RMSEC	q ²	q ² _{ajust}	LSE	LOF	r ² _p	RMSEP	r _m ²	Y-rand
	77	0,68	0,64	0,33	0,63	0,61	0,13	0,27	0,63	0,22	0,568	0,397

Alinhamento 6 $pIC_{50} = 4,41 - 6,01 (0,5,-1,np) - 4,55 (1,11,-4,np) + 1,00 (-1,9,2,np)$ (11)

$- 2,08 (1,-3,2,np) - 6,91(0,0,-3,np) + 3,21 (0,-1,-2,np)$
 $- 0,84(1,8,0,np) - 0,57 (-1,3,-3,np) + 0,86 (-1,10,1,np)$
 $+ 0,73 (0,-2,-4,np)$

Modelo 6	N	r ²	r ² _{ajust}	RMSEC	q ²	q ² _{ajust}	LSE	LOF	r ² _p	RMSEP	r _m ²	Y-rand
	77	0,68	0,64	0,31	0,59	0,57	0,13	0,27	0,48	0,26	0,436	0,342

Na escolha do melhor alinhamento, essas equações foram avaliadas usando diferentes critérios, como: a) o valor do q^2_{ajustado} , que é considerado o índice mais adequado para comparar modelos de QSAR, quando estes apresentam números de termos diferentes. Como descrito anteriormente, esse parâmetro estatístico compara a capacidade preditiva do modelo e bons modelos que possuem valores de q^2 maior que 0,5. b) No intuito de evitar um superajuste nos dados, averiguou-se o número de descritores dos modelos obtidos. Considera-se que o número máximo de termos deve estar na razão de cinco compostos do banco de dados para cada termo da equação (MANNHOLD; KROGSGAARD-LARSEN; TIMMERMAN, 1993). Então, esse número máximo de termos é obtido, dividindo o número total de compostos do conjunto de treinamento ($N=77$) por cinco, resultando, no máximo, um valor de 15 descritores para cada equação. c) Os índices de significância do modelo, como r^2_{ajustado} , erro padrão (SE), *lack-of-fit* (LOF), e a validação externa usando o conjunto de teste.

Analisando-se os valores de q^2_{ajustado} nas equações acima, pode-se observar que todos os modelos apresentam uma boa preditividade. Quanto ao número de variáveis independentes, todos os modelos apresentaram no máximo onze descritores. Entretanto, analisando outros índices estatísticos, observamos que os alinhamentos 1 e 2 se destacaram entre os demais, pois os seus respectivos modelos foram os únicos que apresentaram um coeficiente de correlação linear quadrático ajustado (r^2_{ajustado}) maior ou igual a 0,7 e menores valores de erro padrão em comparação aos demais alinhamentos, isso indica que os alinhamentos 1 e 2 são correlacionados e podem ser usados para calcular a atividade do conjunto de teste (CALDAS; RAMALHO; CUNHA, 2014). Contudo, analisando os descritores de cada um dos modelos representativos dos alinhamentos 1 e 2, observou-se que o modelo representativo do alinhamento 2, é um modelo que apresenta descritores que possuem informações similares, e

inclusive, possui descritor em uma região comum a todos os compostos do conjunto de treinamento, o que não tem sentido, portanto, sendo um modelo não interessante para explicação, como podemos observar na Figura 10.

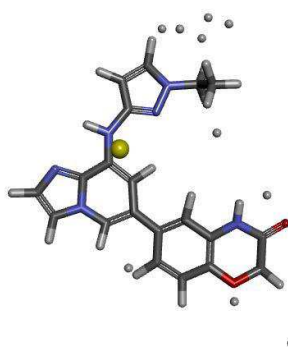


Figura 10 Molécula **8** da série estudada, representando o descritor em amarelo que está localizado na parte comum de todos os compostos do conjunto de treinamento

Por essas razões, o alinhamento 1 foi escolhido como o melhor alinhamento deste estudo QSAR-4D. Apresentando uma significância estatística de 60% da relação entre a resposta biológica e os descritores estruturais químicos para o modelo 1, por meio da análise “*leave-one-out cross validation*” (LOO-cv). Além disso, foi realizado um teste de robustez, baseado no randomization, e o valor obtido foi de 0,245. Quanto à validação externa, o alinhamento 1 se destacou entre os demais alinhamentos, apresentando uma boa capacidade preditiva (r^2_{predito}) com um valor igual a 0,680.

5.3 Análise do Alinhamento 1

Como descrito anteriormente, o banco de dados inicial sofreu pré-tratamento para reduzir o número de termos, para esse propósito, foi utilizado o corte de variância que determina o grau de diversidade de uma determinada

variável. Neste trabalho, foram realizados cortes de variância igual a 0; 0,01; 0,04 e 0,1. Entretanto os modelos provenientes do banco de dados com a maioria desses valores de corte apresentaram valores de r^2_{predito} da validação externa, muito inferior a 0,5 e, até mesmo, valores negativos, esses resultados indicam que esses modelos não conseguiam prever a potência de um composto. Isso pode ocorrer, pois, ao eliminarem células com variância menor, estas podem estar correlacionadas com a atividade e, portanto, terem informações importantes que contribuem na robustez e na confiança do modelo obtido, reduzindo de modo significativo a possibilidade de se obter um bom modelo. O corte por variância igual a zero, no entanto, foi superior aos demais cortes, pois apresentou maiores valores de r^2_{predito} . Portanto, os dez melhores modelos do alinhamento 1, foram obtidos do banco de dados com valor de corte por variância igual a zero e estão apresentados, resumidamente, na Tabela 2.

Tabela 2 Representação dos dez modelos de QSAR-4D obtidos do Alinhamento 1, e célula de 1Å.

Modelos	Q ²	R ²	LSE	LOF	Equações
M1	0,626	0,752	0,100	0,230	$pIC_{50} = 5,06 - 2,10(0,6,-2,np) - 2,65(0,3,-3,np) - 1,15(0,10,3,np) - 1,25(2,1,-8,np) - 1,21(-1,9,0,np) - 3,29(0,-3,-5,np) - 1,11(0,0,-3,alh) - 2,44(0,9,-2,np) + 0,87(1,-1,-4,np) - 0,93(-1,6,3,np) - 1,40(0,1,-3,np)$
M2	0,625	0,747	0,102	0,235	$pIC_{50} = 4,92 - 2,24(0,6,-2,np) - 2,88(0,3,-3,np) - 1,00(0,10,3,np) + 1,63(2,1,-8,np) - 1,21(-1,9,0,np) - 3,62(0,-3,-5,np) - 0,97(0,0,-3,alh) - 2,52(0,9,-2,np) + 0,79(1,-1,-4,np) - 1,10(0,1,-3,np) + 0,19(0,6,3,np)$

“Tabela 2, conclusão”

Modelos	Q ²	R ²	LSE	LOF	Equações
M3	0,622	0,750	0,101	0,232	$pIC_{50} = 5,05 - 2,37(0,6,-2,np) - 1,80(1,-1,-7,alh) - 2,61(0,3,-3,np) - 1,13(0,10,3,np) - 1,19(2,1,-8,np) - 1,12(-1,9,0,np) - 3,26(0,-3,-5,np) - 1,09(0,0,-3,4) - 2,41(0,9,-2,1) + 0,83(1,-1,-4,1) - 1,38(0,1,-3,1)$
M4	0,617	0,737	0,106	0,223	$pIC_{50} = 5,04 - 2,25(0,6,-2,np) - 2,63(0,3,-3,np) - 1,06(0,10,3,np) - 1,57(2,1,-8,np) - 1,21(-1,9,0,np) - 3,58(0,-3,-5,np) - 1,09(0,0,-3,alh) - 2,45(0,9,-2,np) + 0,93(1,-1,-4,np) - 1,36(0,1,-3,np)$
M5	0,615	0,722	0,112	0,236	$pIC_{50} = 5,01 - 2,08(0,6,-2,np) - 2,76(0,3,-3,np) - 1,40(2,1,-8,np) - 1,14(-1,9,0,np) - 3,17(0,-3,-5,np) - 1,05(0,0,-3,alh) - 2,32(0,9,-2,np) + 0,84(1,-1,-4,np) - 0,77(-1,6,3,np) - 1,29(0,1,-3,np)$
M6	0,610	0,712	0,117	0,224	$pIC_{50} = 4,99 - 2,207(0,6,-2,np) - 2,74(0,3,-3,np) - 1,66(2,1,-8,np) - 1,15(-1,9,0,np) - 3,43(0,-3,-5,np) - 1,04(0,0,-3,alh) - 2,33(0,9,-2,np) + 0,90(1,-1,-4,np) - 1,27(0,1,-3,np)$
M7	0,599	0,729	0,109	0,229	$pIC_{50} = 5,109 - 2,66(0,6,-2,np) - 2,23(1,-1,-7,alh) - 2,66(0,3,-3,np) - 1,16(0,10,3,np) - 0,86(-1,9,0,np) - 3,30(0,-3,-5,np) - 1,15(0,0,-3,alh) - 2,43(0,9,-2,np) - 0,89(1,0,-8,np) - 1,47(0,1,-3,np)$
M8	0,595	0,732	0,108	0,227	$pIC_{50} = 5,03 - 2,44(0,6,-2,np) - 2,62(1,-1,-7,alh) - 2,53(0,3,-3,np) - 1,22(0,10,3,np) - 1,05(-1,9,0,np) - 3,09(0,-3,-5,np) - 1,07(0,0,-3,alh) - 2,74(0,9,-2,np) + 0,93(1,-1,-4,np) - 1,35(0,1,-3,np)$
M9	0,593	0,730	0,109	0,250	$pIC_{50} = 5,13 - 2,69(0,6,-2,np) - 2,24(1,-1,-7,alh) - 2,58(0,3,-3,np) - 1,17(0,10,3,np) - 0,86(-1,9,0,np) - 3,35(0,-3,-5,np) - 1,18(0,0,-3,alh) - 2,46(0,9,-2,np) - 0,95(1,0,-8,np) - 1,42(0,1,-3,np) - 0,05(0,0,0,np)$
M10	0,592	0,701	0,121	0,232	$pIC_{50} = 5,06 - 2,13(0,6,-2,np) - 2,82(0,3,-3,np) - 1,46(2,1,-8,np) - 1,15(-1,9,0,np) - 3,30(0,-3,-5,np) - 1,10(0,0,-3,alh) - 2,36(0,9,-2,np) - 0,87(-1,6,3,np) - 1,35(0,1,-3,np)$

LSE (*Least Squares Error*) é o erro dos mínimos quadrados, e LOF (*lack-of-fit*) é o valor do fator de desajuste de Friedman

Analisando a tabela 2, observa-se que todos os modelos obtidos do alinhamento 1 possuem boa capacidade preditiva e apresentam valores do coeficiente de correlação ao quadrado acima de 0,7.

Após a seleção dos dez modelos, foram calculados os valores residuais de atividade entre os pares dos modelos ($pIC_{50calc}-pIC_{50exp}$) apresentados na tabela 3, com o propósito de analisar a matriz de correlação cruzada. Esse tipo de matriz verifica se os modelos fornecem o mesmo tipo de informação ou não (ROGERS; HOPFINGER, 1994).

Tabela 3 Matriz de correlação cruzada dos resíduos de atividade entre pares de modelos para os dez modelos obtidos na análise de QSAR-4D usando o alinhamento 1 com valor de corte por variância igual a zero e célula de 1Å

Modelos	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
M1	1									
M2	0,959	1								
M3	0,973	0,959	1							
M4	0,971	0,98	0,975	1						
M5	0,945	0,914	0,923	0,926	1					
M6	0,927	0,936	0,931	0,955	0,982	1				
M7	0,907	0,888	0,941	0,909	0,864	0,87	1			
M8	0,928	0,905	0,965	0,918	0,873	0,871	0,919	1		
M9	0,896	0,882	0,932	0,901	0,853	0,862	0,998	0,91	1	
M10	0,911	0,877	0,888	0,888	0,964	0,942	0,903	0,836	0,898	1

Analisando os dados da Tabela 3, observa-se que todos os modelos, estão correlacionados (r próximos ou iguais a 1). Isso significa que todos os modelos podem representar o conjunto de treinamento de forma similar. Analisando os parâmetros estatísticos entre os modelos (**M1** a **M10**), percebe-se que o modelo **M1** apresenta melhores resultados, tais como $q^2=0,626$ e $r^2=0,752$.

Considerando esses resultados, selecionou-se o modelo (**M1**), a partir do banco de dados, como sendo o melhor modelo que representa o Alinhamento **1**, o qual será discutido a seguir.

5.4 Análise do melhor modelo do Alinhamento 1

A fim de averiguar se o modelo **M1** apresenta o mesmo tipo de informação estrutura-atividade, foi calculada a matriz de correlação cruzada entre os valores de ocupação das células (GCODs) apresentados na Tabela 4.

Tabela 4 Matriz de correlação cruzada dos GCODs do melhor modelo do Alinhamento 1

GCODs	0,6,-2 (np)	0,3,-3 (np)	0,10,3 (np)	2,1,-8 (np)	-1,9,0 (np)	0,-3,-5 (np)	0,0,-3 (alh)	0,9,-2 (np)	1,-1,-4 (np)	-1,6,3 (np)	0,1,-3 (np)
0,6,-2 (np)	1										
0,3,-3 (np)	-0,07	1									
0,10,3 (np)	-0,047	-0,089	1								
2,1,-8 (np)	0,162	-0,102	0,061	1							
-1,9,0 (np)	-0,072	-0,071	-0,089	-0,065	1						
0,-3,-5 (np)	-0,009	-0,223	-0,161	0,088	0,071	1					
0,0,-3 (alh)	-0,027	-0,034	-0,034	-0,039	-0,038	-0,086	1				
0,9,-2 (np)	0,352	-0,113	-0,084	0,375	-0,076	0,199	-0,043	1			
1,-1,-4 (np)	-0,073	0,016	0,094	-0,099	0,011	-0,07	-0,04	-0,101	1		
-1,6,3 (np)	0,174	-0,13	-0,115	0,334	-0,004	0,306	-0,052	0,207	-0,143	1	
0,1,-3 (np)	-0,028	0,699	-0,064	-0,127	-0,007	-0,259	-0,044	-0,132	-0,032	-0,166	1

np: não polar e alh: acceptor de ligação de hidrogênio.

Pode-se verificar que, com exceção de um par, os demais pares de descritores do modelo **M1** selecionado não estão correlacionados, pois apresentam $r < 0,5$. Ou seja, esses descritores mostram pouca tendência de uma variável estar influenciando na outra. Assim, pode-se dizer que cada variável contribui de maneira única para a construção de modelos confiáveis (MANNHOLD; KROGSGAARD-LARSEN; TIMMERMAN, 1993).

O único par de GCODs que estão correlacionados são as células (0, 1, -3, np) e (0,3,-3,np) com um valor de $r = 0,699$. A correlação encontrada, pode ser justificada, em razão da proximidade espacial, pois a distância medida entre as células é igual a 2,0Å (Figura 11). Além disso, pode-se verificar que essas células apresentam o mesmo tipo de IPE correspondente ao mesmo tipo de interação não polar (np) e, também, coeficientes negativos, demonstrando que uma ocupação idêntica das duas células resultará em efeitos similares. Segundo Cunha et al. (2005), variáveis independentes relacionadas ocorrem, tanto em um mesmo modelo ou em modelos diferentes, pois a maioria dos modelos estão altamente correlacionados uns com os outros.

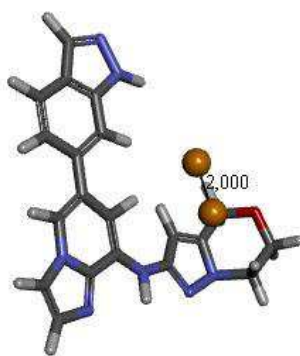


Figura 11 Molécula **19** mais ativa da série, representando a distância do par de GCODs correlacionados

5.5 Análise dos Valores Residuais e Identificação dos *Outliers*

Os valores da atividade biológica predita ($pIC_{50Predito}$) computados a partir da equação (6), juntamente com os respectivos valores residuais do conjunto de treinamento são apresentados na Tabela 5. Os valores residuais foram calculados como a diferença entre os valores de $pIC_{50Predito}$ e $pIC_{50Experimental}$. No Gráfico 2, exibem-se os valores de pIC_{50} experimentais *versus* pIC_{50} predito do conjunto de treinamento.

Os *outliers* são considerados como compostos que apresentem valor residual duas vezes superiores ao valor de desvio padrão dos resíduos ($pIC_{50Exp} - pIC_{50Pred}$) (CUNHA et al., 2006; HOPFINGER et al., 1997; MANNHOLD; KROGSGAARD-LARSEN; TIMMERMAN, 1993). O desvio padrão (DP), para o conjunto de treinamento, foi de **0,294**, portanto, o valor residual máximo permitido é igual a \pm **0,588**. Assim, o modelo **M1** possui três *outliers*, compostos **15** (0,677), **54** (-0,878) e **80** (-0,674).

Tabela 5 Valores de $pIC_{50Experimental}$ e $pIC_{50predito}$ e valores residuais ($pIC_{50predito} - pIC_{50experimental}$) para os compostos do conjunto de treinamento

Composto^a	pIC_{50} Experimental	pIC_{50} Predito	Valores residuais
3	4,740	4,917	0,177
4	4,650	4,638	-0,011
5	4,876	5,020	0,144
6	4,625	4,829	0,204
7	5,208	4,871	-0,337
8	3,261	3,398	0,137
9	3,766	3,444	-0,322
10	3,740	4,194	0,454
11	5,180	5,017	-0,163
12	4,392	3,957	-0,434
13	3,813	4,105	0,292
14	4,450	4,605	0,155
15	3,727	4,404	0,677
16	4,636	5,065	0,429

(...continua...)

“Tabela 5, continuação”

Composto^a	pIC₅₀ Experimental	pIC₅₀ Predito	Valores residuais
17	5,222	5,051	-0,171
18	5,046	5,069	0,023
20	4,886	4,613	-0,273
21	5,301	5,013	-0,288
22	4,409	4,813	0,404
23	4,824	4,699	-0,124
24	5,301	4,946	-0,355
26	5,097	5,015	-0,082
27	5,097	5,055	-0,041
29	5,301	5,061	-0,240
31	5,347	5,095	-0,251
32	4,634	4,440	-0,193
33	4,42	4,492	0,072
34	4,979	4,643	-0,336
37	5,367	4,948	-0,419
39	4,602	4,554	-0,047
40	4,268	4,850	0,583
42	4,745	4,752	0,007
44	4,364	4,112	-0,252
45	4,682	5,191	0,509
46	5,000	5,137	0,137
47	5,432	5,135	-0,296
48	4,81	4,975	0,165
49	4,947	4,968	0,022
50	4,622	4,624	0,002
51	4,962	4,942	-0,019
52	5,113	5,299	0,186
53	4,967	5,019	0,052
54	5,553	4,766	-0,787
55	4,382	4,588	0,206
56	5,149	5,127	-0,022
61	4,585	4,647	0,062
62	4,886	4,944	0,058
63	4,658	4,490	-0,168
64	4,721	4,818	0,097
65	4,036	4,316	0,280
66	4,769	4,524	-0,245
67	4,769	4,437	-0,332
68	4,377	4,238	-0,139
69	4,149	4,485	0,336
70	3,521	3,439	-0,082
71	3,054	3,215	0,161

(... continua ...)

“Tabela 5, conclusão”

Composto ^a	pIC ₅₀ Experimental	pIC ₅₀ Predito	Valores residuais
72	3,78	3,755	-0,025
74	3,893	4,077	0,184
76	3,374	3,449	0,075
77	3,599	3,815	0,216
78	3,565	3,989	0,423
79	3,274	3,439	0,165
80	3,495	2,821	-0,674
81	3,567	4,075	0,508
82	4,097	3,802	-0,295
83	4,495	4,632	0,137
84	3,668	3,897	0,229
85	3,556	3,300	-0,256
86	4,854	4,703	-0,151
87	4,769	4,429	-0,339
88	3,661	3,834	0,173
91	4,143	3,893	-0,250
92	4,31	4,876	0,566
93	4,167	4,499	0,332
96	4,122	4,024	-0,098
97	4,364	4,181	-0,183

a. Compostos *outliers* estão em negritos e sublinhados.

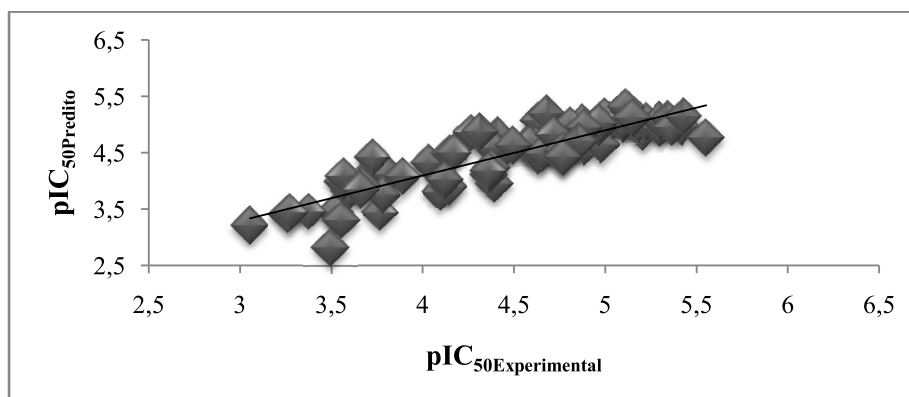


Gráfico 3 Valores de pIC₅₀Experimental *versus* pIC₅₀Predito do conjunto de treinamento usando o melhor modelo de QSAR-4D representativo do alinhamento 1

Em se tratando do conjunto de teste os compostos *outliers*, foram calculados da mesma maneira que os compostos do conjunto de treinamento ou seja, os compostos com valor residual duas vezes superiores ao valor do DP. O valor do DP encontrado para o grupo de teste foi de 0,408, dessa forma, o valor máximo permitido é $\pm 0,816$. Os modelos foram capazes de prever bem os valores de atividade e não apresentaram nenhum composto do conjunto de teste como *outlier*. Como descrito anteriormente, o conjunto de teste, formado por 20 compostos (Quadro 1), não participam da construção dos modelos, portanto o cálculo para prever a atividade teórica para esses compostos, utilizando o modelo M1, corresponde a um processo de validação externa.

Na Tabela 6, estão descritos os valores da atividade biológica experimental, juntamente com os valores da atividade biológica predita ($pIC_{50Predito}$), calculados a partir da equação (6), com os respectivos valores residuais do conjunto de teste.

Tabela 6 Valores de $pIC_{50Experimental}$ e $pIC_{50predito}$ e valores residuais ($pIC_{50predito} - pIC_{50experimental}$) para os compostos do conjunto de teste

Composto	$pIC_{50 Experimental}$	$pIC_{50 Predito}$	Valores residuais
19	5,699	5,622	-0,077
25	4,699	4,971	0,272
28	4,824	4,557	-0,267
2	3,527	3,964	0,437
30	3,536	4,207	0,671
35	4,599	4,563	-0,036
36	5,071	4,623	-0,447
38	3,86	4,094	0,234
41	4,886	4,911	0,025
43	5,585	4,908	-0,677
57	4,706	4,011	-0,695
58	4,349	4,298	-0,050
59	4,307	4,818	0,511
60	4,893	4,858	-0,034

“Tabela 6, conclusão”

Composto	pIC ₅₀ Experimental	pIC ₅₀ Predito	Valores residuais
73	2,988	3,302	0,314
75	4,26	4,294	0,034
89	4,824	4,111	-0,714
90	4,921	4,869	-0,052
94	4,579	4,736	0,157
95	3,244	3,765	0,521

O composto **15**, **54** e **80** do conjunto de treinamento foram considerados *outliers* para esse modelo. Considera-se que o motivo do composto **15** ser um *outlier*, pode estar relacionado com a sua estrutura química na região do anel metil-benzimidazol. Observa-se, no Quadro 1, que esta região é similar para as moléculas **7**, **10** e **34** do grupo de treinamento e, dessa forma, o modelo ao prever a atividade da molécula **15**, comparou-a com os compostos **7**, **10** e **34** que obtiveram atividade predita maior (Quadro 1). Levando, assim, a uma superestimação no valor de atividade do composto **15**, o valor calculado para esse composto foi de **4,404** (Tabela 5).

Em se tratando do composto **54**, certamente foi considerado um *outlier*, em razão de um grupamento pirazolo-pirazina que está presente somente na molécula **42** do conjunto de treinamento (Quadro 1). Como essa molécula apresenta um valor de pIC₅₀Experimental inferior, referindo-se ao composto **54**, o modelo, ao prever a atividade da molécula **54**, comparou-a com a molécula **42**, levando a um valor subestimado de atividade predita para o composto **54**, cujo valor predito de atividade é igual a **4,766** (Tabela 5).

Já, o composto **80**, obteve um comportamento diferente dos demais *outliers*, esta molécula apresenta o mesmo substituinte indol-etanol-metil apenas para as moléculas **85** e **88** do conjunto de treinamento. Entretanto, a atividade predita para a molécula **80** foi de **2,821** (Tabela 5) um valor estimado baixo, em comparação aos valores de atividade predita para os compostos **85** (pIC₅₀ predito =

3,300) e **88** ($pIC_{50 \text{ predito}} = 3,834$). Era esperado que o modelo previsse uma atividade maior para o composto **80**, já que os compostos que possuem esse substituinte (**85** e **88**) apresentaram atividade biológica experimental maior que o composto **80**, o que não ocorreu.

5.6 Análise dos Descritores

Na Figura 12, estão representados os descritores em 3D obtidos pelo modelo **M1** (Equação 6), usando o composto **19** mais potente da série como referência. Neste trabalho, os GCODs estão representados por esferas. A esfera verde representa o GCOD com coeficiente positivo, correspondente à região no composto onde a presença de certos substituintes favorece no aumento de sua potência inibitória. Já, as esferas laranja representam os GCODs com coeficientes negativos, que correspondem às regiões na molécula onde a presença de determinado substituinte irá diminuir a potência dos compostos.

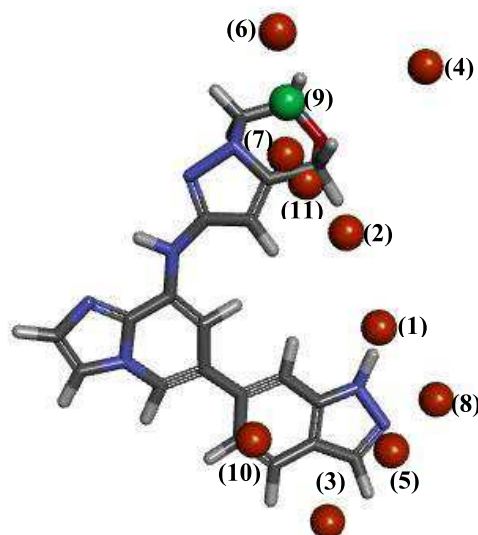


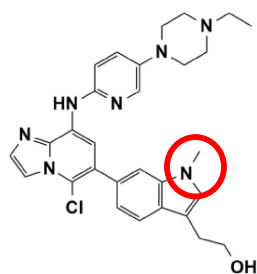
Figura 12 Composto **19** com os descritores gerados pelo modelo **M1**. Cada GCOD é descrito como “x, y, z, IPE”, que representam as coordenadas cartesianas da célula de grade selecionada (x, y, z) e o respectivo tipo de átomo (IPE). Os GCODs descritos são: **(1)** (0, 6, -2, np), **(2)** (0, 3, -3, np), **(3)** (0, 10,3, np), **(4)** (2, 1, -8, np), **(5)** (-1, 9, 0, np), **(6)** (0, -3, -5,np) **(7)** (0, 0,-3,alh), **(8)** (0, 9,-2, np), **(9)** (1,-1,-4, np), **(10)** (-1, 6, 3, np), **(11)** (0,1,-3, np)

O GCOD (0, 6, -2, np) com coeficiente negativo, localiza-se no anel indol dos compostos **80**, **85** e **88**, onde apresentam alta frequência de ocupação para esse descritor. A célula (0, 6, -2, np) é do tipo não polar (Figura 13), isso significa que o grupo metil, nessa posição, é desfavorável para a atividade biológica. Essa interpretação é apoiada quando o composto **85**, com um grupo metil ligado ao átomo do nitrogênio do anel indol é comparado ao composto **84**, o qual possui um hidrogênio (Figura 14). Acredita-se que esse inibidor **84** está mais propício para realizar interações que favorecem a atividade, por exemplo, ligação de hidrogênio e, conseqüentemente, tem uma maior potência de inibição do que o composto **85**. Esse argumento é justificado pelos estudos de Currie et al. (2014), nos quais havia mencionado a importância do NH livre, ligado ao

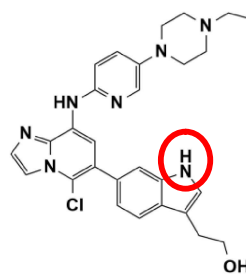
anel indol indicando que o NH pode formar ligações de hidrogênio com a Syk (Figura 15), interações favoráveis que contribuem com o aumento da potência.



Figura 13 Modelo **M1** de QSAR-4D-IR representativo do alinhamento 1, mostrando o GCOD com coeficiente negativo (0,6,-2,np) (esfera laranja) localizado no anel indol do composto **80**, no qual apresentou maior frequência de ocupação para esse GCOD



Composto **85** ($pIC_{50} = 3,556$)



Composto **84** ($pIC_{50} = 3,668$)

Figura 14 Estruturas químicas dos compostos de imidazopiridinas e suas respectivas atividades biológicas experimentais.

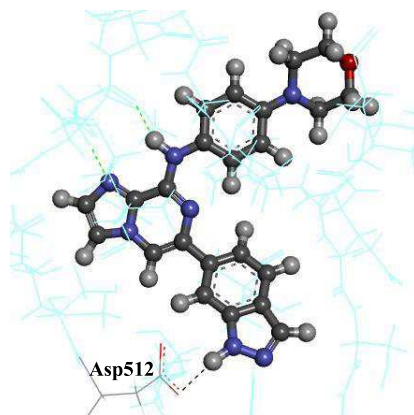


Figura 15 Representação da interação do NH livre, do ligante CG9, cocristalizado no sítio da Syk obtida do PDB, cujo código: 4PUZ, fazendo interação de hidrogênio com o resíduo de aminoácido Asp512

Os GCODs $(0,10,3,np)$, $(0,-3,-5,np)$, $(0,3,-3,np)$ e $(0,1,-3,np)$ representados na Figura 16, apresentam coeficientes negativos e representam elementos de interação não polar, ou seja, indicam que substituintes hidrofóbicos nessas posições contribuem com o decréscimo da potência dos compostos. O GCOD $(0,10,3,np)$ está localizado no átomo CH_2 do grupo etanol ligado ao anel indazol do composto **66** (Figura 17). O GCOD $(0,-3,-5,np)$ está localizado no grupo tiazolpiridina do composto **91** que possui maior contribuição para esse descritor (Figura 15). Já, os GCODs $(0,3,-3,np)$ e $(0,1,-3,np)$ estão localizados no grupo etil ligado ao anel indol (Figura 16), dos compostos **8** e **9**, respectivamente, moléculas estas que possuem maior frequência ocupacional par esses GCODs. Certamente, esses casos referem-se a um impedimento espacial. Pois, em se tratando do GCOD $(0,10,3,np)$, observa-se que o grupo etanol causa um efeito estérico no composto **66**, pois quando comparado com o composto **11**, composto este que não possui nenhuma frequência ocupacional para esse descritor e também não apresenta o grupo etanol (Figura 18), o que se observa é um aumento na atividade dessa molécula. Já, o GCOD $(0,-3,-5,np)$, observa-se

que todos os compostos que apresentam substituintes menores, nesta região, não apresentam frequência de ocupação para esse descritor. Portanto, cadeias mais compactas, na região R_1 (Figura 7), são fundamentais para o aumento do potencial inibidor desses compostos. Como exemplo, tem-se o composto **92** ($pIC_{50\text{experimental}} = 4,310$), que possui atividade biológica maior, em comparação ao composto **91** ($pIC_{50} = 4,143$), representado na Figura 19. Esse argumento também se aplica aos GCODs (0,3,-3,np) e (0,1,-3,np), indicando que aumento da cadeia na região R_1 (Figura 7), como a presença do grupo etil interfere, significativamente, na atividade dos compostos **8** e **9** (Figura 16), o que pode ser comprovado pela baixa atividade dessas moléculas (molécula **8**, $pIC_{50\text{experimental}} = 3,261$ e molécula **9**, $pIC_{50\text{experimental}} = 3,766$), em comparação com composto mais ativo (molécula **19**, $pIC_{50\text{experimental}} = 5,699$).

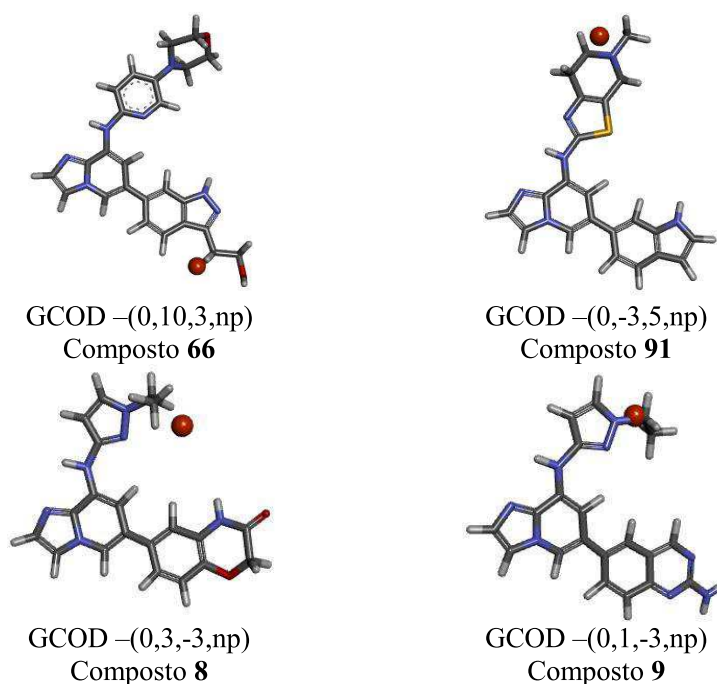


Figura 16 Modelo **M1** de QSAR4D-IR representativo do alinhamento 1, mostrando os GCODs com coeficiente negativo (esferas laranjas)

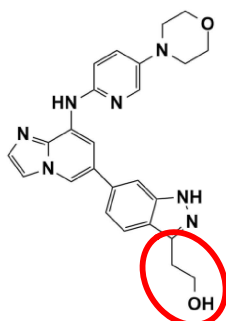
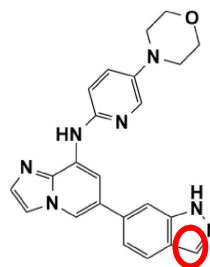
Composto **66** ($pIC_{50} = 4.769$)Composto **11** ($pIC_{50} = 5.180$)

Figura 17 Estruturas químicas dos compostos de imidazopiridinas e suas respectivas atividades biológicas experimentais

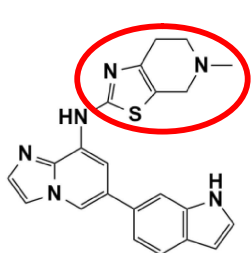
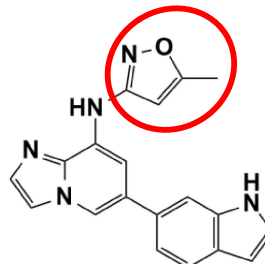
Composto **91** ($pIC_{50}=4,143$)Composto **92** ($pIC_{50}=4,310$)

Figura 18 Estruturas químicas dos compostos de imidazopiridinas e suas respectivas atividades biológicas experimentais

O GCOD com coeficiente negativo (2,1,-8,np) representa uma interação não polar e está localizado próximo ao grupo etil ligado ao anel piperazinil do composto 79 (Figura 19). Indicando que substituintes polares, nessa região, aumentam a potência do inibidor. Isso pode ser explicado, baseado no composto 48, essa molécula é diferente do composto 79, nessa região, possuindo um grupo carbonila, essa diferença é bem significativa na atividade, como demonstrado na Figura 20. Esse argumento pode ser sustentado, pois ao observar esse composto no sítio de ligação da enzima Syk (Código: 4FZ6), observa-se que o grupo etil posiciona-se próximo ao resíduo de aminoácido Lys, resíduo este polar básico, indicando que grupos polares, nessa posição, favorecem interações com esse

aminoácido, aumentando, assim, a atividade do composto representado na Figura 21.

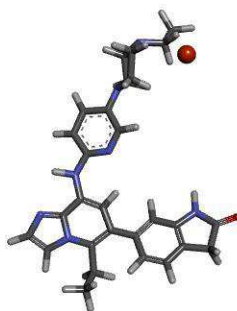
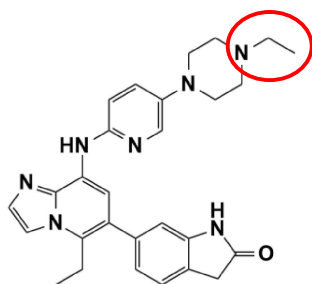
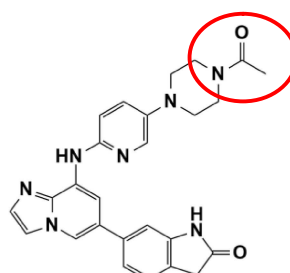


Figura 19 Modelo **M1** de QSAR-4D-IR representativo do alinhamento 1, mostrando o GCOD com coeficiente negativo (2,1,-8, np) (esfera laranja) localizado no substituinte etil do composto **79** que apresentou maior frequência de ocupação para esse GCOD



Composto **79** ($pIC_{50} = 3.274$)



Composto **48** ($pIC_{50} = 4.810$)

Figura 20 Estruturas químicas dos compostos de imidazopiridinas e suas respectivas atividades biológicas experimentais

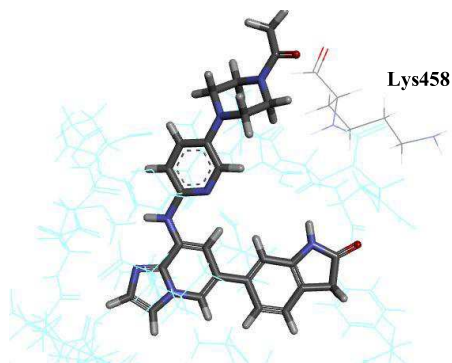


Figura 21 Representação do composto 48 no sítio da enzima Syk, observa-se que o resíduo de aminoácido lisina se encontra próximo a essa região, reforçando o argumento de que compostos polares, nessa região, contribuem com a atividade desse inibidor

O GCOD com coeficiente negativo (-1,9,0,np), representa um elemento de interação não polar e está localizado no grupo indol do composto **79** (Figura 22). Nesse caso, trata-se, mais especificamente, da orientação relativa desse grupo, pois essa região do composto tem grande liberdade conformacional, sendo modificada, principalmente, durante a Simulação por Dinâmica Molecular. É importante observar que essa região interfere, consideravelmente, para o aumento ou decréscimo da potência, assim sendo, a orientação dessa conformação do composto **79** para com esse descritor, desfavorece a atividade.



Figura 22 Modelo **M1** de QSAR-4D-IR representativo do alinhamento 1, mostrando o GCOD com coeficiente negativo (0,10,3,np) (esfera laranja) localizado no grupo indol do composto **79**,no qual apresentou maior frequência de ocupação para esse GCOD

O GCOD com coeficiente negativo (0,0,-3,alh) é do tipo acceptor de ligação de hidrogênio (Figura 23). Está localizado no átomo de nitrogênio do grupo pirimidina, isso indica que esse átomo, nessa região do inibidor, decresce com o aumento da potência. O composto **12** é um exemplo, é a única molécula que possui contribuição para esse descritor, isso pode ser explicado, em razão do composto **12** apresentar uma região diferente de todas as outras do conjunto de treinamento, contudo, ele é menos potente (composto **12**, $pIC_{50\text{experimental}} = 4,392$) com relação ao composto mais ativo (composto **19**, $pIC_{50\text{experimental}} = 5,699$), corroborando a presença desse descritor.



Figura 23 Modelo **M1** de QSAR-4D-IR representativo do alinhamento 1, mostrando o GCOD com coeficiente negativo (0,0,-3,alh) (esfera laranja), localizado no átomo de nitrogênio do grupo pirimidina do composto **12**, no qual apresentou maior frequência de ocupação para esse GCOD

O GCOD com coeficiente negativo (0,9,-2,np), representa um elemento de interação não polar e está localizado no grupo indol, indicando que grupos hidrofóbicos, nessa região, desfavorecem a atividade biológica (Figura 24), ou seja, é importante, nessa região, grupos hidrofílicos. Como pode ser observado na Figura 25, o composto **76** é o que possui a maior contribuição desse descritor, aproximadamente 0,472, possui atividade significativamente menor em comparação com o composto **68** que apresenta uma irrelevante frequência de ocupação de 0,001 para esse descritor. Portanto, é necessária a presença de grupos polares nessa região do metil. Isso pode ser explicado, pois, ao analisar o composto **76** dentro do sítio ativo da enzima Syk (Código: 4FZ6), observou-se que existem, próximo a essa região, dois resíduos polares que, possivelmente, podem realizar interações com um substituinte hidrofílico, nessa posição, representado na Figura 26.

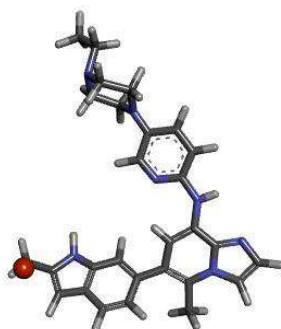
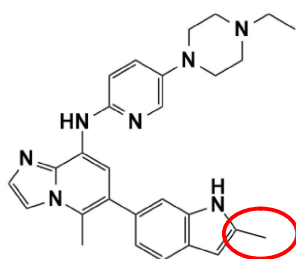
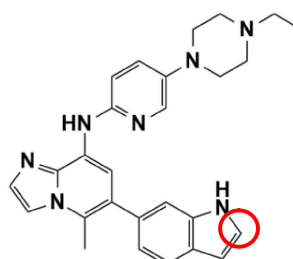


Figura 24 Modelo **M1** de QSAR-4D-IR representativo do alinhamento 1, mostrando o GCOD com coeficiente negativo (0,9,-2, np) (esfera laranja), localizado no grupo indol do composto **76**, no qual apresentou maior frequência de ocupação para esse GCOD



Composto **76** ($pIC_{50} = 3,374$)



Composto **68** ($pIC_{50} = 4,377$)

Figura 25 Estruturas químicas dos compostos de imidazopiridinas e suas respectivas atividades biológicas experimentais

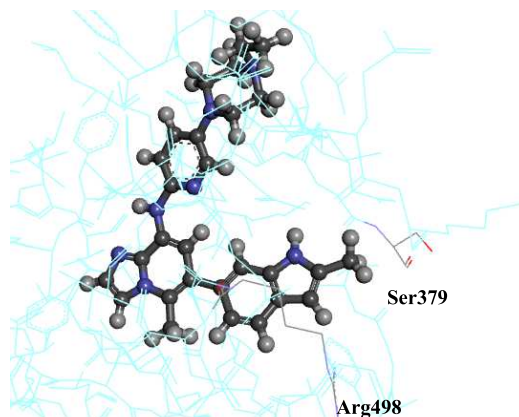
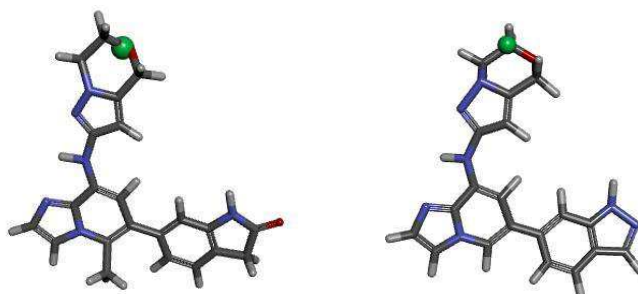


Figura 26 Representação do composto **76** dentro do sítio da enzima Syk obtida do PDB. Observa-se que os resíduos de aminoácidos Serina e Arginina se encontram próximos a essa região, reforçando o argumento de que, compostos polares, nessa região, contribuem com a atividade desse inibidor

O GCOD (1,-1,4,np) com coeficiente positivo é único descritor que contribui para o aumento na potência dos compostos nesse modelo (Equação 6), já que possui um coeficiente positivo de 0,872. Representa um elemento de interação não polar e está localizado no átomo de oxigênio do anel oxazina do composto **53** (Figura 27). Nesse caso, trata-se mais especificamente da orientação desse grupo durante a simulação por DM, pois esse composto é um dos mais potentes da série, e durante a DM o anel oxazina, que também possui átomo não polar ficou localizado próximo a essa região, como podemos observar na Figura 27. Em resumo, a molécula mais ativa (composto **19**) e todos os compostos da série que possuem esse substituinte apresentam uma elevada atividade, portanto pode-se dizer que o anel oxazina contribui favoravelmente com a atividade.



Composto **53** ($pIC_{50\text{experimental}} = 4,967$) Composto **19** ($pIC_{50\text{Experimental}} = 5,699$)

Figura 27 Modelo **M1** de QSAR-4D-IR representativo do alinhamento 1, mostrando o GCOD com coeficiente positivo (1,-1,-4,np) (esfera verde) localizado no átomo de oxigênio do anel oxazina do composto **53**, composto este que apresentou maior frequência de ocupação para esse GCOD. A direita o composto **19**, mostrando que durante a DM esse descritor se posicionou no átomo CH_2 , próximo a essa região, esse fato explica a presença desse GCOD

O GCOD, com coeficiente negativo (-1,6,3,np), representa um elemento de interação não polar e está localizado no grupo indazol do composto **77** (Figura 28). Isso indica que substituintes polares, nessa região, podem aumentar a potência de inibição. A informação que esse GCOD traz é bem interessante, pois nenhum dos compostos dessa série possui algum grupo não hidrofóbico nessa exata região. Outra observação é que o composto **77** que apresenta maior contribuição para esse GCOD possui um grupo etil próximo a essa posição (-1,6,3), que pode ser o responsável pela presença desse descritor, já que muitas moléculas da série que apresentam substituinte não polar, nessa região, possuem valor de atividade baixo. Como exemplo, tem-se o composto **67**, que apresenta uma estrutura química semelhante ao do composto **77**, o que diferencia é a ausência do grupo etil. O composto **67** apresenta um grupo metil, nessa região, e a diminuição do grupo apolar interfere, significativamente, na atividade (Figura 29).

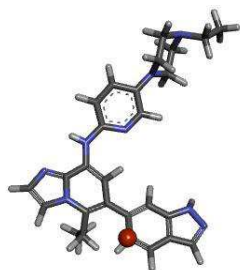
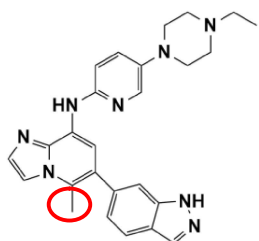
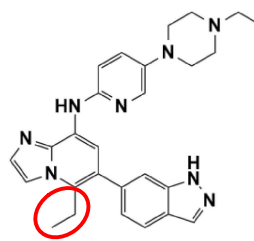


Figura 28 Modelo **M1** de QSAR-4D-IR representativo do alinhamento 1, mostrando o GCOD com coeficiente negativo (-1,6,3,np) (esfera laranja), localizado no grupo indazol do composto **77**, onde apresentou maior frequência de ocupação para esse GCOD



Composto **67** ($pIC_{50} = 4,769$)



Composto **77** ($pIC_{50} = 4,260$)

Figura 29 Estruturas químicas dos compostos de imidazopiridinas e suas respectivas atividades biológicas experimentais

Em resumo, pode-se observar que a maioria dos GCODS que apresentam coeficiente negativo é do tipo de interação não polar, sugerindo que grupos hidrofílicos nas posições R_1 , R_2 e R_3 (Figura 7) são importantes para a potência dos inibidores. Na Figura 30, observam-se as regiões com os possíveis substituintes preferidos.

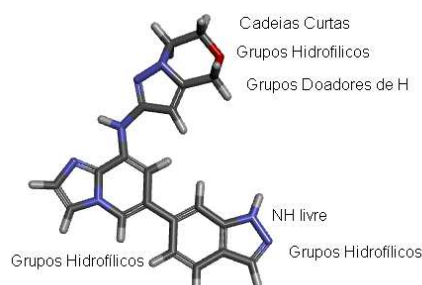


Figura 30 Resumo geral das informações retiradas do Modelo M1, resultante da metodologia QSAR-4D

Analisando os resultados obtidos, observamos que o modelo M1 foi capaz de fornecer características farmacofóricas importantes que contribuem para a inibição da Syk. Mas, infelizmente, essas características não foram suficientes para propor estruturas hipotéticas que apresentassem atividade maior que o composto **19**, sendo o composto mais ativo da série, cujo valor da atividade predita pelo modelo para essa molécula foi de $pIC_{50predito} = 5,622$.

O que se observou foi que o modelo teve uma dificuldade para estimar a atividade calculada bem próxima à da experimental, em se tratando dos compostos que tem atividade experimental próxima ao composto mais ativo. Isso ocorreu, pois essa série de moléculas possui substituintes nas posições R_1 , R_2 e R_3 (Figura 7) muito heterogêneos, e o modelo, geralmente, funciona muito melhor para os compostos que estão dentro de seu domínio de aplicabilidade, que é definido pela sua similaridade estrutural.

Contudo, algumas estruturas hipotéticas propostas, obtiveram atividade próxima ao valor esperado, 5,62, como podemos observar nas Figuras 31, 32 e 33. A partir das atividades calculadas, observa-se que essas estruturas propostas reforçam as informações transmitidas pelos GCODs, que colaboram para a inibição da Syk, nota-se que todos os compostos que foram propostos referem-se a uma miscelânea entre a molécula mais ativa da série (Composto **19**) e características importantes fornecidas pelo modelo.

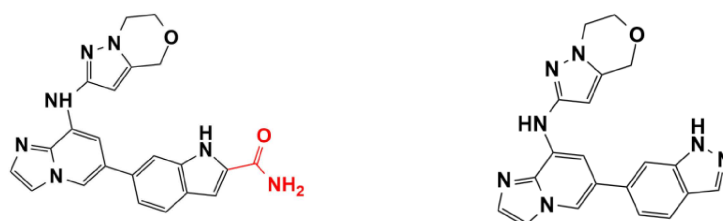
Composto **P1** $pIC_{50\text{predito}} = 5,540$ Composto **19** $pIC_{50\text{predito}} = 5,622$

Figura 31 Estrutura hipotética (**P1**), proposta para estudos posteriores, baseados nos resultados obtidos no estudo de QSAR4D-IR de inibidores imidazopiridínicos. À direita, composto 19 mais ativo da série

Observa-se, na Figura 31, que a região R_1 (Figura 7), manteve-se o grupamento oxazina, ou seja, cadeia curta. Na região R_2 (Figura 7), o NH permaneceu livre, podendo vira realizar interações de hidrogênio com o receptor e acrescentou um grupo hidrofílico a essa porção (destacado em vermelho). Na região R_3 (Figura 7) não ocorreu mudanças.

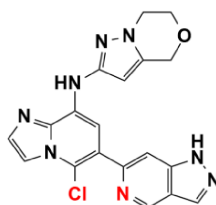
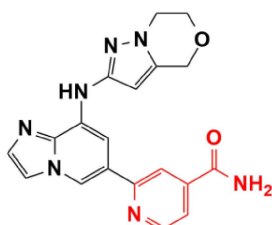
Composto **P2** $pIC_{50\text{predito}} = 5,453$

Figura 32 Estrutura hipotética (**P2**), proposta para estudos posteriores, baseados nos resultados obtidos no estudo de QSAR4D-IR de inibidores imidazopiridínicos

Na figura 32, observa-se que, na região R_1 (Figura 7), manteve-se o grupamento oxazina, mesmo substituinte do composto mais ativo. Na região R_2 (Figura 7), o NH permaneceu livre, podendo vira realizar interações de hidrogênio com o receptor e acrescentou um átomo de nitrogênio no anel indazol

(destacado em vermelho) a essa porção. Na região R₃ (Figura 7), adicionou um átomo de cloro.



Composto **P3** $pIC_{50\text{predito}} = 5,469$

Figura 33 Estrutura hipotética (P3), proposta para estudos posteriores baseados nos resultados obtidos no estudo de QSAR4D-IR de inibidores imidazopiridínicos

Com respeito à Figura 33, nota-se que na região R₁ (Figura 7) manteve-se o grupamento oxazina, mesmo substituído do composto mais ativo. Na região R₂ (Figura 7), acrescentou-se o grupamento piridina-carboxamida (destacado em vermelho). Na região R₃ (Figura 7), não houve mudanças.


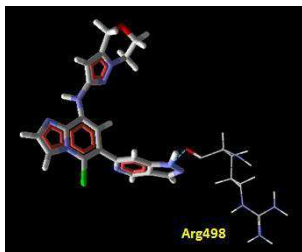
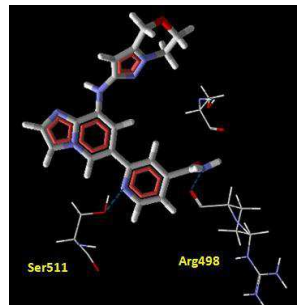
No estudo de ancoramento molecular, o inibidor P1, que foi apontado como o melhor inibidor proposto (com base nos estudos teóricos), apresentou um menor valor de energia de interação intermolecular (Tabela 7) do que os outros compostos propostos (P2 e P3) e, também, menor energia de interação do que o composto 19, mais ativo da série.

Tabela 7 Valores de energia de interação intermolecular proteína/inibidores

Moléculas	Energia de Interação (Kcal. Mol ⁻¹)
P1	-120,953
P2	-113,533
P3	-111,573
19	-114,156

Além disso, a molécula **P1** apresentou um maior número de ligações de hidrogênio com os resíduos de aminoácidos do sítio ativo da Syk (Tabela 8), ou seja, esse Inibidor (P1) apresenta uma maior estabilidade no interior do sítio ativo, quando comparado às estruturas propostas **P2** e **P3**, validando, assim, o modelo.

Tabela 8 Valores de energia de interação intermolecular proteína/inibidores, energias de ligações de hidrogênio e número de ligações de hidrogênio

Moléculas	Energia do comprimento das ligações de H (Kcal MoL ⁻¹)	Número de ligações de H
P1	-2,12 -0,53 -0,69	3
		
P2	-1,96	1
		
P3	-1,50 -2,12	2
		

6 CONCLUSÃO

Após a análise da metodologia de QSAR-4D, aplicada aos compostos derivados de imidazopiridina, conclui-se que o modelo gerado pelo melhor alinhamento (Alinhamento 1) apresentou valores de r^2 e q^2 de 0.752 e 0.626, respectivamente. Além disso, o resultado da validação externa r^2_{predito} igual a 0,680, comprova que esse modelo é capaz de fornecer previsões confiáveis para novos compostos. Quanto à posição dos descritores, nos sugere que uma cadeia curta na região R_1 é importante para o aumento da potência dos inibidores e, ademais, observamos que nas regiões R_2 e R_3 , substituintes hidrofílicos aumentam o poder inibitório.

REFERÊNCIAS

4D-QSAR user's manual. Version 1.00. Lake Forest: The Chem21 Group, 1997. 142 p.

ACCELRY'S SOFTWARE. **Discovery studio modeling environment**. Version 3.5. San Diego, 2012.

AMARAL, A. T. do; MONTANARI, C. A. 25 years of medicinal chemistry in Brazil. **Química Nova**, São Paulo, v. 25, p. 39-44, jul. 2002.

ANDRADE, C. H. et al. 4D-QSAR: perspectives in drug design. **Molecules**, Basel, v. 15, n. 5, p. 3281-3294, May 2010.

ARROIO, A.; HONÓRIO, K. M.; SILVA, A. B. F. da. Propriedades químico-quânticas empregadas em estudos das relações estrutura-atividade. **Química Nova**, São Paulo, v. 33, n. 3, p. 694-699, 2010.

ASSIS, L. C. et al. Interactions of pyrimidine derivatives with dihydrofolate reductase and thymidylate synthase: directions toward combating toxoplasmosis. **Current Bioactive Compounds**, Sharjah, v. 9, n. 2, p. 153-166, 2013.

ASSIS, T. M. et al. In silico study of *Leishmania donovani* α - β tubulin and inhibitors. **Journal of Chemistry**, Melville, v. 2014, p. 1-8, July 2014.

AVILA, C. M.; ROMEIRO, N. C. Proteínas tirosinas quinases: desafios do desenvolvimento de fármacos para a terapia do câncer. **Revista Virtual de Química**, Niterói, v. 2, n. 1, p. 59-82, ago. 2010.

BAHIA, D. et al. *Schistosoma mansoni*: expression of es-like tyrosine kinase SmFes in the tegument and terebratorium suggests its involvement in host penetration. **Experimental Parasitology**, San Diego, v. 116, p. 225-232, Jan. 2007.

BENDALY, J. et al. Role of human CYP1A1 and NAT2 in 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine-induced mutagenicity and DNA adducts. **Xenobiotica**, London, v.39, n. 5, p.399-406, May 2009.

BERMAN, H. M. et al. The protein data bank. **Nucleic Acids Research**, Oxford, v. 28, n. 1, p. 235-242, Oct. 2000.

BERSON, A. et al. Toxicity of alpidem, a peripheral benzodiazepine receptor ligand, but not zolpidem, in rat hepatocytes: role of mitochondrial permeability transition and metabolic activation. **Journal of Pharmacology and Experimental Therapeutics**, Baltimore, v. 299, n. 2, p. 793-800, 2001.

BOERNER, R. J.; MOLLER, H. J. Saripidem-a new treatment for panic disorders. **Psycho pharmako therapie**, Berlin, v. 4, n. 4, p. 145-148, 1997.

BLOMGREN, P. et al. **Imidazopyridininessykinhibitors**. US n. 112995 A1, 15 Sept. 2011.

BROOKS, B. R. et al. CHARMM: the biomolecular simulation program. **Journal of Computational Chemistry**, New York, v. 30, n. 10, p. 1545-1614, 2009.

BYTH, K. F. et al. The cellular phenotype of AZ703, a novel selective imidazo 1,2-a pyridine cyclin-dependent kinase inhibitor. **Molecular Cancer Therapeutics**, Birmingham, v. 5, n. 3, p. 655-664, 2006.

CALDAS, G. B.; RAMALHO, T. C.; CUNHA, E. F. F. da. Application of 4D-QSAR studies to a series of benzothiophene analogs. **Journal of Molecular Modeling**, Washington, v. 20, n. 10, p. 1-10, 2014.

CARVALHO, I. et al. Introduction to molecular modeling of drugs in the medicinal chemistry experimental course. **Química Nova**, São Paulo, v. 26, n. 3, p. 428-438, maio/jun. 2003.

CASE, D. A. et al. The Amber biomolecular simulation programs. **Journal of Computational Chemistry**, New York, v. 26, n. 16, p. 1668-1688, 2005.

CASTILLO, M. et al. Highly potent aminopyridines as Syk kinase inhibitors. **Bioorganic & Medicinal Chemistry Letters**, New York, v.22, n. 17, p. 5419-5423, Sept. 2012.

CASTRO, R. O. Regulation and function of syk tyrosine kinase in mast cell signaling and beyond. **Journal of Signal Transduction**, New York, v.2011, p. 1-9, 2011.

CHAKRAVARTY, S. D. et al. Kinase inhibitors: a new tool for the treatment of rheumatoid arthritis. **Clinical Immunology**, Orlando, v.148, n. 1, p.66-78, July 2013.

CHENG, K.W.;CHEN,F.;WANG,M. Heterocyclicamines: chemistry and health. **Molecular Nutrition & Food Research**, Weinheim, v.50,n. 12, p.1150-1170, Dec. 2006.

CORREA-BASURTO, J. et al. QSAR, docking, dynamic simulation and quantum mechanics studies to explore the recognition properties of cholinesterase binding sites. **Chemico-Biological Interactions**, Limerick, v.209, p.1-13, Feb. 2014.

CRUZ,O. J. d'; UCKUN, M. F. Protein kinase inhibitors against malignant lymphoma. **Expert Opinion Pharmacotherapy**, London, v.14, n.6, p.707-721, 2013.

CUNHA, E. F. F. da et al. 4D-QSAR models of HOE/BAY-793 analogues as HIV-1 protease inhibitors. **QSAR & Combinatorial Science**, Weinheim, v. 24, n. 2, p. 240-253, Mar. 2005.

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, 2006.

CUNHA, E.F.F. da; MATOS, K. S.; RAMALHO, T.C. QSAR and docking studies of HCV NS3 serine protease inhibitors. **Medicinal Chemistry**, Washington, v. 9, n. 6, p. 774-805, Sept. 2013.

CURRIE, S.K. et al. Discovery of GS-9973, a selective and orally efficacious inhibitor of spleen tyrosine kinase. **Journal of Medicinal Chemistry**, Washington, v. 57, n. 9, p.3856-3873, Apr. 2014.

CYWIN, C. L. et al. Discovery and SAR of novel Naphthyridines as potent inhibitors of spleen tyrosine kinase (SYK). **Bioorganic & Medicinal Chemistry Letters**, New York, v. 13, n. 8, p. 1415-1418, 2003.

EL-SAYED et al. The position of imidazopyridine and metabolic activation are pivotal factors in the antimutagenic activity of novel imidazo[1,2-a] pyridine derivatives. **European Journal of Pharmacology**, Amsterdam, v.715, n. 1, p.212-218, Sept. 2013.

FANG, C. et al. Microfluidic and imaging platform for a kinase activity radioassay to analyze minute patient cancer samples. **Cancer Research**, Baltimore, v.70, n. 21, p. 8299-8308, Aug. 2010.

FERREIRA, M. M. C.; MONTANARI, C. A.; GAUDIO, A. C. Seleção de variáveis em QSAR. **Química Nova**, São Paulo, v.25, n.3, p.439-448, 2002.

GARCIA, L. S. **Estudos de potenciais inibidores da enzima tirosina quinase de Bruton**. 2014. 100p. Dissertação (Mestrado em Agroquímica) - Universidade Federal de Lavras, Lavras, 2014.

GRAMATICA, P. Principles of QSAR models validation: internal and external. **QSAR & Combinatorial Science**, Weinheim, v. 26, n. 5, p. 694-701, 2007.

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

HANSCH, C.; FUJITA, T. ρ - σ - π analysis: a method for the correlation of biological activity and chemical structure. **Journal of the American Chemical Society**, Easton, v. 86, p. 1616-1626, 1964.

HARRISON, T. S.; KEATING, G. M. Zolpidem. **CNS Drugs**, New York, v. 19, n. 1, p. 65-89, Aug. 2005.

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, 1997.

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

KAUR, M.; BAHIA, M. S. SILAKARI, O. Exploring the role of water molecules for docking and receptor guided 3D-QSAR analysis of naphthyridine derivatives as spleen tyrosine kinase (Syk) inhibitors. **Journal of Chemical Information and Modeling**, Washington, v. 52, n. 10, p. 2619-2630, Oct. 2012.

KAUR, M. et al. Designing of new multi-targeted inhibitors of spleen tyrosine kinase (Syk) and zeta-associated protein of 70 kDa (ZAP-70) using hierarchical virtual screening protocol. **Journal of Chemical Information and Modeling**, Washington, v. 39, p. 165-175, Feb. 2013.

KAUR, M.; SINGH, M.; SILAKARI, O. Inhibitors of switch kinase 'Spleen Tyrosine Kinase' in inflammation and immune-mediated disorders: a review. **European Journal of Medicinal Chemistry**, Paris, v. 67, p. 434-446, Sept. 2013.

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.

KUBINYI, H. From narcosis to hyperspace: the history of QSAR. **Quantitative Structure Activity Relationships**, Weinheim, v. 21, n. 4, p. 348-356, Oct. 2002.

MANNHOLD, R.; KROGSGAARD-LARSEN, P.; TIMMERMAN, H. **QSAR: hansch analysis and related approaches**. Weinheim: J. Wiley, 1993.252 p.

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

KULATHU, Y.; GROTHE, G.; RETH, M. Autoinhibition and adapter function of Syk. **Immunological Reviews**, Copenhagen, v. 232, n. 1, p. 286-299, Oct. 2009.

KYTTARIS, V. C. et al. Kinase inhibitors: a new class of antirheumatic drugs. **Journal of Drug Design, Development and Therapy**, Auckland, v.6, p. 245-250, Sept. 2012.

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, 2009.

LAI, J. Y. Q. et al. Potent small molecule inhibitors of spleen tyrosine kinase (Syk). **Bioorganic & Medicinal Chemistry Letters**, New York, v. 13, n. 18, p. 3111-3114, 2003.

LIDDLE, J. et al. Discovery of GSK143, a highly potent, selective and orally efficacious spleen tyrosine kinase inhibitor. **Bioorganic & Medicinal Chemistry Letters**, New York, v. 21, n. 20, p. 6188-6194, 2011.

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

LIMA NETO, Q. A. da et al. Estudo da relação estrutura-atividade quantitativa de dibenzoilmetanos α -substituídos quanto à atividade anticâncer da mama (linhagem celular MCF7). **Revista Brasileira de Cancerologia**, São Paulo, v. 52, n. 4, p. 345-357, 2006.

LUCAS, C. M. et al. Rational design of highly selective spleen tyrosine kinase Inhibitors. **Journal of Medicinal Chemistry**, Washington, v.55, n. 23, p. 10414-10423, Nov. 2012.

MANCINI, D. T. et al. Toxoplasmose: perspectivas no estudo de novos alvos terapêuticos. **Revista Virtual de Química**, Niterói, v. 4, n. 4, p. 434-455, 2012.

MARTINS, J. P. A. **Química computacional aplicada a QSAR**. São Paulo: CENAPAD-SP, 2010. 52 p.

MARZARO, G. et al. Approach to fit Multi-target QSAR Models for tyrosine kinases inhibitors. **European Journal of Medicinal Chemistry**, Paris, v.46, n. 6, p.2185-2192, June 2011.

MÓCSAI, A. et al. The immunomodulatory adapter proteins DAP12 and Fc receptor γ -chain (FCR γ) regulate development of functional osteoclasts through the Syk tyrosine kinase. **PNAS**, Washington, v. 101, n. 16, p.6158-6163, 2004.

MÓCSAI, A.; RULAND, J.; TYBULEWICZ, V.L.J. The SYK tyrosine kinase: a crucial player in diverse biological functions. **Nature Reviews Immunology**, London, v. 10, n. 6, p. 387-402, June 2010.

MOLSIM users guide. Version 3.0. Lake Forest: Doherty DC & The Chem 21 Group, 1997.1 CD-ROM.

NAMBA, A. M.; SILVA, V. B. D.; SILVA, C. H. T. P. D. Dinâmica molecular: teoria e aplicações em planejamento de fármacos. **Eclética Química**, Araraquara, v.4, n. 4, p.13-24, dez. 2008.

NIJJAR, J. S. et al. New molecules for the treatment of rheumatoid arthritis, inhibition of spleen tyrosine kinase in the treatment of rheumatoid arthritis. **Rheumatology**, Montreal, v.52, n. 9, p.1556-1562, May 2013.

OLIVEIRA, A. A. **Estudos de ancoramento e dinâmica molecular de derivados do ácido oxâmico inibidores da enzima lactato desidrogenase de *Plasmodium falciparum***. 2012. 142 p. Dissertação (Mestrado em Química) - Instituto Militar de Engenharia, Rio de Janeiro, 2012.

PAMUK, O.N.; TSOKOS, G. C. Spleen tyrosine kinase inhibition in the treatment of autoimmune, allergic and autoinflammatory diseases. **Arthritis Research & Therapy**, London, v.12, n.6, p. 2-11, Dec. 2010.

PÂRVU, L. QSAR-a piece of drug design. **Journal of Cellular and Molecular Medicine**, New York, v.7, n. 3, p. 333-335, 2003.

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.

PITA, S. S. R. **Modelos de QSAR-4D dependente do receptor de inibidores peptídicos da tripanotona redutase**. 2006. 94 p. Dissertação (Mestrado em Ciências) - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2006.

POLANSKI, J. Receptor dependent multidimensional QSAR for modeling drug: receptor interactions. **Current Medicinal Chemistry**, Wageningen, v. 16, n. 25, p. 3243-3257, Jan. 2009.

POTEMKIN, V.; GRISHINA, M. Principles for 3D/4D QSAR classification of drugs. **Drug Discovery Today**, Kidlington, v. 13, n. 21/22, p. 952-959, Nov. 2008.

PUERTA, J.A.G.;MÓCSAI,A. Tyrosine kinase inhibitors for the treatment of rheumatoid arthritis. **Current Topics in Medicinal Chemistry**, Hilversum, v.13, n.6, p.760-770, 2013.

PRATIM ROY, P. et al. On two novel parameters for validation of predictive QSAR models. **Molecules**, Basel, v. 14, n. 5, p. 1660-1701, 2009.

RANIERI, G. et al. Tyrosine kinase inhibitors (TKIs) in human and pet tumours with special reference to breast cancer: a comparative review. **Critical Reviews in Oncology/Hematology**, New York, v. 88, n. 2, p. 293-308, 2013.

RICCABONI, M.; BIANCHI, I.; PETRILHO, P. Spleen tyrosine kinases: biology, therapeutic targets and drugs. **Drug Discovery Today**, Kidlington, v.15, n. 13/14, p. 517-530, July 2010.

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, June 2005.

SALETU, B.; GRÜNBERGER, J.; LINZMAYER, L. Pharmacokinetic and dynamic studies with a new anxiolytic imidazo-pyridine alpidem utilizing pharmaco-EEG and psychometry. **International Clinical Psychopharmacology**, Baltimore, v. 1, n. 2, p. 145-164, 1986.

SANTOS-FILHO, O. et al. The receptor-dependent QSAR paradigm: an overview of the current state of the art. **Medicinal Chemistry**, Washington, v.5, n. 4, p. 359-366, July 2009.

SCHULTZ, T.W. Quantitative structure-activity relationships (QSARs) in toxicology: a historical perspective. **Journal of Molecular Structure**, Amsterdam, v. 622, n. 1, p. 1-22, Mar. 2003.

SENESE, C. L.; HOPFINGER, A. J. A simple clustering technique to improve QSAR model selection and predictivity: application to a receptor independent 4D-QSAR analysis of cyclic urea derived inhibitors of HIV-1 protease. **Journal of Chemical Information and Computer Sciences**, Washington, v.43, p.2180-2193,2003.

SIDDIQUI, I. R. et al. Basic ionic liquid promoted heterocyclization to access fused imidazopyridines. **Tetrahedron Letters**, Oxford, v. 54, n. 37, p. 5083-5086, Sept. 2013.

SILVA, D. R.; RAMALHO, T. C.; CUNHA, E. F. F.da.4DSAR model for compounds with binding affinity towards dopamine D2 receptors. **Letters in Drug Design & Discovery**, Cambridge, v.11, n. 5, p. 649-664, June 2014.

SPOEL, D. van der et al. GROMACS: fast, flexible, and free. **Journal of Computacional Chemistry**, New York, v.26, p.1701-1718, Oct. 2005.

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: abordagem de Hansch. **Química Nova**, São Paulo, v. 27, n. 4, p. 631-639, 2004.

THOMA, G. et al. Syk inhibitors with high potency in presence of blood. **Bioorganic & Medicinal Chemistry Letters**, New York,v.24,n. 10, p.2278-2282, May 2014.

VEDANI, A.; DOBLER, M.; LILL, M. A. Combining protein modeling and 6D-QSAR: simulating the binding of structurally diverse ligands to the estrogen receptor. **Journal of Medicinal Chemistry**, Washington, v. 48, n. 11, p. 3700-3703, June 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.

WEBER, K.C. **Modelagem molecular de compostos arilpiperazínicos e suas interações com o receptor 5-HT**.2008. 138 p. Tese (Doutorado em Físico-Química) - Instituto de Química de São Carlos, São Carlos, 2008.

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, n. 2, p. 230-252, 1986.

YANMAZ, E. et al. 4D-QSAR analysis and pharmacophore modeling: electron conformation algenetic 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.

ZENG, J. et al. Copper(ii)/iron(iii) co-catalyzed intermolecular diamination of alkynes: facile synthesis of imidazopyridines. **Organic Letters**, Washington, v.17, n.14, p. 4386-4389, July 2012.

438

Mol.	GCOD_1	GCOD_2	GCOD_3	GCOD_4	GCOD_5	GCOD_6	GCOD_7	GCOD_8	GCOD_9	GCOD_10	GCOD_11
31	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,037	0,000	0,001
32	0,000	0,000	0,000	0,000	0,181	0,000	0,000	0,000	0,026	0,000	0,305
33	0,000	0,000	0,000	0,000	0,062	0,150	0,000	0,000	0,003	0,000	0,005
34	0,000	0,000	0,000	0,000	0,197	0,000	0,000	0,000	0,002	0,000	0,132
35	0,000	0,000	0,054	0,000	0,000	0,142	0,000	0,000	0,033	0,000	0,001
36	0,000	0,000	0,000	0,000	0,019	0,129	0,000	0,000	0,007	0,000	0,000
37	0,000	0,000	0,000	0,006	0,001	0,037	0,000	0,000	0,016	0,000	0,000
38	0,000	0,000	0,243	0,000	0,001	0,210	0,000	0,000	0,002	0,000	0,000
39	0,000	0,000	0,340	0,034	0,013	0,019	0,000	0,000	0,003	0,000	0,000
40	0,000	0,001	0,000	0,000	0,012	0,002	0,000	0,000	0,004	0,000	0,138
41	0,000	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,009	0,000	0,114
42	0,000	0,000	0,000	0,000	0,005	0,098	0,000	0,000	0,029	0,000	0,007
43	0,000	0,000	0,000	0,016	0,004	0,041	0,000	0,000	0,005	0,000	0,001
44	0,000	0,121	0,000	0,000	0,000	0,000	0,000	0,000	0,139	0,000	0,537
45	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,153	0,000	0,003
46	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,086	0,000	0,000
47	0,000	0,004	0,000	0,000	0,000	0,000	0,000	0,013	0,130	0,000	0,000
48	0,000	0,000	0,000	0,023	0,000	0,021	0,000	0,000	0,010	0,000	0,000
49	0,000	0,000	0,000	0,011	0,005	0,023	0,000	0,000	0,001	0,000	0,001
50	0,000	0,000	0,329	0,000	0,007	0,016	0,000	0,000	0,000	0,000	0,000
51	0,000	0,027	0,000	0,000	0,001	0,000	0,000	0,000	0,019	0,000	0,047
52	0,000	0,000	0,000	0,000	0,032	0,029	0,000	0,012	0,467	0,000	0,007
53	0,000	0,000	0,000	0,000	0,190	0,092	0,000	0,000	0,572	0,000	0,009
54	0,000	0,000	0,000	0,001	0,004	0,095	0,000	0,000	0,032	0,000	0,006
55	0,000	0,174	0,000	0,000	0,007	0,000	0,000	0,000	0,049	0,000	0,036
56	0,000	0,000	0,258	0,000	0,016	0,017	0,000	0,000	0,504	0,000	0,003
57	0,000	0,222	0,404	0,000	0,001	0,000	0,000	0,000	0,049	0,000	0,030
58	0,000	0,297	0,000	0,000	0,000	0,000	0,000	0,000	0,034	0,000	0,007
59	0,000	0,000	0,000	0,000	0,001	0,001	0,000	0,100	0,017	0,000	0,009
60	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,147
61	0,000	0,105	0,000	0,000	0,002	0,000	0,000	0,000	0,031	0,000	0,117
62	0,000	0,000	0,000	0,000	0,006	0,025	0,000	0,000	0,011	0,000	0,029
63	0,000	0,164	0,000	0,000	0,000	0,000	0,000	0,000	0,056	0,000	0,135
64	0,000	0,000	0,000	0,000	0,000	0,080	0,000	0,000	0,019	0,000	0,000

Mol.	GCOD_1	GCOD_2	GCOD_3	GCOD_4	GCOD_5	GCOD_6	GCOD_7	GCOD_8	GCOD_9	GCOD_10	GCOD_11
65	0,000	0,000	0,404	0,217	0,000	0,009	0,000	0,000	0,020	0,000	0,000
66	0,000	0,000	0,431	0,000	0,000	0,019	0,000	0,000	0,023	0,000	0,001
67	0,000	0,000	0,000	0,136	0,000	0,139	0,000	0,000	0,013	0,000	0,009
68	0,000	0,000	0,000	0,142	0,198	0,124	0,000	0,001	0,005	0,000	0,003
69	0,000	0,000	0,000	0,103	0,011	0,132	0,000	0,000	0,002	0,000	0,004
70	0,000	0,000	0,028	0,221	0,004	0,086	0,000	0,362	0,003	0,161	0,000
71	0,000	0,000	0,000	0,517	0,002	0,102	0,000	0,289	0,001	0,176	0,000
72	0,033	0,000	0,000	0,000	0,000	0,108	0,000	0,346	0,001	0,045	0,000
73	0,431	0,000	0,000	0,000	0,393	0,086	0,000	0,000	0,000	0,107	0,000
74	0,000	0,000	0,000	0,000	0,000	0,239	0,000	0,000	0,003	0,221	0,000
75	0,000	0,000	0,001	0,105	0,000	0,133	0,000	0,000	0,000	0,214	0,002
76	0,000	0,000	0,000	0,000	0,009	0,142	0,000	0,472	0,019	0,000	0,002
77	0,000	0,000	0,000	0,132	0,036	0,169	0,000	0,000	0,001	0,523	0,002
78	0,000	0,000	0,000	0,019	0,024	0,088	0,000	0,284	0,009	0,053	0,000
79	0,000	0,000	0,000	0,012	0,933	0,136	0,000	0,000	0,004	0,039	0,002
80	0,381	0,000	0,014	0,000	0,006	0,113	0,000	0,386	0,001	0,113	0,002
81	0,000	0,000	0,000	0,002	0,008	0,075	0,000	0,252	0,004	0,126	0,002
82	0,000	0,000	0,000	0,000	0,803	0,055	0,000	0,000	0,006	0,124	0,001
83	0,000	0,000	0,000	0,000	0,003	0,083	0,000	0,000	0,004	0,169	0,002
84	0,000	0,000	0,020	0,371	0,000	0,042	0,000	0,167	0,011	0,154	0,002
85	0,243	0,000	0,010	0,450	0,000	0,015	0,000	0,174	0,005	0,229	0,000
86	0,000	0,000	0,000	0,005	0,003	0,058	0,000	0,000	0,005	0,173	0,004
87	0,000	0,000	0,000	0,025	0,281	0,062	0,000	0,000	0,003	0,068	0,000
88	0,235	0,000	0,017	0,038	0,001	0,042	0,000	0,140	0,000	0,204	0,000
89	0,000	0,000	0,040	0,141	0,000	0,038	0,000	0,183	0,003	0,177	0,000
90	0,000	0,000	0,049	0,020	0,028	0,025	0,000	0,000	0,007	0,000	0,003
91	0,000	0,000	0,000	0,000	0,007	0,357	0,000	0,000	0,012	0,000	0,000
92	0,000	0,000	0,000	0,000	0,027	0,000	0,000	0,000	0,000	0,000	0,111
93	0,000	0,000	0,000	0,001	0,000	0,173	0,000	0,000	0,010	0,000	0,003
94	0,000	0,000	0,000	0,000	0,000	0,093	0,000	0,000	0,000	0,000	0,016
95	0,000	0,000	0,000	0,000	0,000	0,352	0,000	0,000	0,000	0,119	0,023
96	0,000	0,000	0,000	0,000	0,000	0,271	0,000	0,000	0,004	0,102	0,042
97	0,000	0,000	0,000	0,000	0,001	0,149	0,000	0,000	0,005	0,423	0,004