



GIOVANNA CARDOSO GAJO

**ESTUDOS QUANTITATIVOS DE
CORRELAÇÃO ESTRUTURA-ATIVIDADE EM
4D DE POTENCIAIS INIBIDORES DA ENZIMA**

Rho QUINASE

LAVRAS - MG

2016

GIOVANNA CARDOSO GAJO

**ESTUDOS QUANTITATIVOS DE CORRELAÇÃO ESTRUTURA-
ATIVIDADE EM 4D DE POTENCIAIS INIBIDORES DA ENZIMA Rho
QUINASE**

Tese 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 Doutor.

Orientador

Dra. Elaine Fontes Ferreira da Cunha

LAVRAS - MG

2016

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

Gajo, Giovanna Cardoso.

Estudos quantitativos de correlação estrutura-atividade em 4D de potenciais inibidores da enzima Rho quinase / Giovanna Cardoso Gajo. – Lavras: UFLA, 2016.

86 p.

Tese (doutorado) – Universidade Federal de Lavras, 2016.

Orientador(a): Elaine Fontes Ferreira da Cunha.

Bibliografia.

1. 4D-QSAR. 2. Dinâmica molecular. 3. Piridiltiazol. 4. ROCK1. 5. *Docking*. I. Universidade Federal de Lavras. II. Título.

GIOVANNA CARDOSO GAJO

**ESTUDOS QUANTITATIVOS DE CORRELAÇÃO ESTRUTURA-
ATIVIDADE EM 4D DE POTENCIAIS INIBIDORES DA ENZIMA Rho
QUINASE**

Tese 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 Doutor.

APROVADA em 14 de março de 2016.

Dr. Cleber Paulo Andrada Anconi	UFLA
Dra. Daiana Mancini Teixeira	UFLA
Dra. Melissa Soares Caetano	UFOP
Dr. Teodorico de Castro Ramalho	UFLA
Dr. Stephen Jones Barigye	UFLA

Dra. Elaine Fontes Ferreira da Cunha
Orientadora

LAVRAS – MG

2016

AGRADECIMENTOS

À Universidade Federal de Lavras e ao Departamento de Química, pela oportunidade de realizar este trabalho.

À professora Elaine, pela orientação e ensinamentos.

Aos professores Teodorico e Matheus, pela colaboração.

Aos colegas do Laboratório de Modelagem Molecular “Era do Gelo”, por tornarem essa experiência mais agradável e por compartilharem o aprendizado.

Aos amigos Daiana e Stephen, por me ajudarem a expandir o conhecimento; ao Josué, por me divertir com as provas de orgânica; ao Aquino, por me oferecer sempre um café; aos meus “vizinhos”, Telles e Maíra, por sempre me animarem; à Silvis, por me espiritualizar, ao Lucas, pelos debates musicais e à Fátima, pelo companheirismo.

As colegas do grupo QSAR, Dani, Lets A, Lets G, Mari e Tamiris, por serem minhas coorientadoras, minhas irmãzinhas de ofício que levo no coração.

A meus pais, Cilaine e José Márcio, pelo apoio.

Às minhas irmãs, Janaína e Jordana, pelo amor incondicional.

Aos meus tios, Paulo, Sandra, Regina e Abedi, por sempre me incentivarem.

À Capes, pela bolsa de estudos.

*Nunca deixe que alguém te diga que não pode fazer algo.
Nem mesmo eu. Se você tem um sonho, tem que protegê-lo.
As pessoas que não podem fazer por si mesmas dirão que
você não consegue. Se quer alguma coisa, vá e lute por ela.*

Ponto final.

CHRIS GARDNER

RESUMO

A série de inibidores da enzima ROCK quinase, desenvolvida por Lawrence, Pireddu e Sebti (2013) a partir de derivados de piridiltiazol, foi submetida ao estudo quantitativo de relação estrutura-atividade em quatro dimensões (QSAR-4D). Os modelos foram construídos aplicando-se otimização de algoritmo genético (GA) combinada com regressão de mínimos quadrados parciais (PLS). O melhor modelo apresentou os seguintes valores de otimização: $r^2 = 0,773$; $q^2_{cv} = 0,672$; $r^2_{pred} = 0,503$; $rm^2_{test} = 0,520$; $r^2_{Y-rand} = 0,19$; $Rp^2 = 0,590$. Ademais, por meio da análise dos descritores, foi possível otimizar novos compostos que apresentaram valores de concentração inibitória predita maior que a do composto mais ativo da série. Estes compostos foram submetidos ao estudo de *docking* e *drug-likeness* e se mostraram promissores candidatos a inibidores da ROCK1.

Palavras-chave: 4D-QSAR. Dinâmica molecular. Piridiltiazol. ROCK1. *Docking*.

ABSTRACT

A series of pyridylthiazole derivatives developed by Lawrence, Pireddu and Sebti (2013) as Rho-associated protein kinase inhibitors were subjected to four-dimensional quantitative structure-activity relationship (4D-QSAR) analysis. The models were generated applying genetic algorithm (GA) optimization combined with partial least squares (PLS) regression. The best model presented validation values of $r^2 = 0.773$; $q^2_{cv} = 0.672$; $r^2_{pred} = 0.503$; $rm^2_{test} = 0.520$; $r^2_{Y-rand} = 0.19$; $Rp^2 = 0.590$. Furthermore, analysing the descriptors it was possible to propose new compounds that predicted higher inhibitory concentration values than the most active compound of the series. These compounds were submitted to docking and drug-likeness study and are promising candidates to Rock1 inhibitors.

Keywords: 4D-QSAR. Molecular dynamics. Pyridylthiazole. ROCK1. Docking.

LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS

ALA	Alanina
ASP	Aspartato
GLU	Glutamina
LYS	Lisina
MET	Metionina
PHE	Fenilalanina
VAL	Valina
ATP	Trifosfato de adenosina
CEP	<i>Conformational ensemble profile</i>
GA	<i>Genetic algorithm</i>
GCOD	<i>Grid cell occupancy descriptor</i>
GFA	<i>Genetic function approximation</i>
IPE	<i>Interaction pharmacophore elements</i>
PDB	<i>Protein Data Bank</i> (Banco de Dados de Proteínas)
PLS	<i>Partial least squares</i>
QSAR	Relação quantitativa estrutura-atividade
QSAR-4D DR	Quarta dimensão da relação quantitativa estrutura-atividade dependente do receptor
QSAR-4D IR	Quarta dimensão da relação quantitativa estrutura-atividade independente do receptor
ROCK	Rho Quinase
SDM	Simulação da dinâmica molecular
Å	Angstrom
IC ₅₀	Concentração capaz de inibir 50% da atividade enzimática
K	Kelvin
µM	Micromolar
M	Molar
kcal mol ⁻¹	Quilocaloria por mol
pIC ₅₀	Log de IC ₅₀

LISTA DE FIGURAS

Figura 1	Ativação da enzima ROCK mediada pela RhoA. “A” representa a estrutura inativa da enzima na qual a porção C-terminal inibe o domínio quinase e “B” representa a exposição do domínio quinase proporcionado pela ligação da RhoA-GTP ou de moléculas como ácido araquidônico e enzimas da cascata de apoptose 15
Figura 2	Exemplos de inibidores da ROCK: Y-27632, Y-30141, fasudil e CID5056270 20
Figura 3	Enzima ROCK disponível no PDB (Código 3TV7). Em destaque, o ligante 1-[(1R)-1-(3-metoxifenil)etil]-3-(4-piridin-4-il-1,3-tiazol-2-il)urea da cadeia A e suas interações com os resíduos próximos... 35
Figura 4	Representação dos átomos fixados nos alinhamentos destacados.... 38
Figura 5	Gráfico de Williams para o modelo Ali5_001_3 56
Figura 6	Representação dos resíduos de aminoácidos com a molécula (29) menos ativa. As esferas em verde representam os GCODs positivos e as em vermelho, os negativos..... 57
Figura 7	Representação dos resíduos de aminoácidos. Em destaque, Gly85 com a molécula 30. A esfera em destaque representa o GCOD(1) .. 59
Figura 8	Representação dos resíduos de aminoácidos. Em destaque, Phe87 com a molécula 33. A esfera em destaque representa o GCOD(2) .. 60
Figura 9	Representação dos resíduos de aminoácidos em destaque Val90, com a molécula 19. A esfera em destaque representa o GCOD(3) .. 61
Figura 10	Representação dos resíduos de aminoácidos. Em destaque, Lys105 com a molécula 2. A esfera em destaque representa o GCOD(4) 62
Figura 11	Interações entre os resíduos de aminoácidos Met 156, Lys105 e Asp 216 com um ligante da série congênere. As interações pontilhadas representam as ligações de hidrogênio. Adaptada de Pireddu et al. (2012) 63
Figura 12	Representação dos resíduos de aminoácidos. Em destaque, Met156, Leu205, Ala215 com a molécula 25. A esfera em destaque representa o GCOD(5)..... 64
Figura 13	Representação dos resíduos de aminoácidos. Em destaque, Arg84 com a molécula 25. A esfera em destaque representa o GCOD(6) .. 65
Figura 14	Representação dos resíduos de aminoácidos com a molécula 35. A esfera em destaque representa o GCOD(7) 66

LISTA DE TABELAS

Tabela 1	Avaliação de inibidores da ROCK em fase de testes farmacológicos para diferentes patologias. Adaptado de Pan et al. (2013).....	19
Tabela 2	Estrutura dos 51 compostos análogos ao piridiltiazol e seus valores de pIC ₅₀ (LAWRENCE; PIREDDU; SEBTI, 2013).....	28
Tabela 3	Resumo dos resultados dos melhores modelos de QSAR-4D, obtidos para os dez alinhamentos com o grupo teste I e seus respectivos cortes.....	49
Tabela 4	Resumo dos resultados dos melhores modelos de QSAR-4D obtidos para os dez alinhamentos com o grupo teste II e seus respectivos cortes.....	51
Tabela 5	Resumo dos resultados dos melhores modelos de QSAR-4D obtidos para os dez alinhamentos com o grupo teste III e seus respectivos cortes.....	53
Tabela 6	Resumo dos resultados dos melhores alinhamentos de RD-4D-QSAR obtidos após validação.....	55
Tabela 7	Moléculas propostas e seus respectivos valores de pIC ₅₀ predito.....	67
Tabela 8	Propriedades moleculares.....	72
Tabela 9	Energias de <i>docking</i> dos compostos analisados.....	73

SUMÁRIO

1	INTRODUÇÃO	12
2	REFERENCIAL TEÓRICO	14
2.1	Via RhoA/Rho quinase	14
2.2	Principais inibidores da Rho quinase	17
2.3	Relação quantitativa estrutura-atividade	21
2.4	QSAR 4D	24
3	OBJETIVOS	26
4	METODOLOGIA	27
4.1	Coleta de dados biológicos	27
4.1.2	Obtenção da estrutura cristalizada de Rho Quinase	34
4.1.3	Construção e minimização dos modelos tridimensionais dos inibidores	35
4.2	QSAR	36
4.2.1	Simulação de dinâmica molecular	36
4.2.2	Definição do alinhamento	37
4.2.3	Definição do tamanho das células	38
4.2.4	Definição dos elementos de interação farmacofórica	39
4.2.5	Obtenção dos modelos de QSAR-4D	39
4.2.6	Validação interna dos modelos	41
4.2.7	Validação externa dos modelos	42
4.2.8	Definição do domínio de aplicabilidade	43
4.3	Otimização estrutural e <i>drug-likeness</i>	44
4.4	<i>Docking</i>	45
5	RESULTADOS E DISCUSSÃO	48
5.1	Alinhamentos para grupos teste I, II e III	48
5.2	Análise dos descritores	57
5.3	Regra de Lipinski	71
5.4	<i>Docking</i> molecular	73
6	CONCLUSÃO	75
7	DIFUSÃO DE TECNOLOGIA	76
	REFERÊNCIAS	77
	APÊNDICE	85

1 INTRODUÇÃO

Rho quinase, ou ROCK (E.C 2.7.11.1), pertence à família de Ser/Thr proteína-quinases, que é ativada, inicialmente, por meio da interação entre guanosina trifosfato (GTP) ativa e a proteína *Ras homolog family member A*, ou RhoA. Evidências crescentes sugerem que RhoA e ROCK participam de uma variedade de funções fisiológicas importantes na vasculatura, incluindo contração do músculo liso, proliferação celular, adesão celular, migração e muitos aspectos de respostas inflamatórias (RIENTO; RIDLEY, 2003). Duas isoformas, ROCK1 e ROCK2, foram identificadas (MATSUI et al., 1996). Elas partilham uma sequência homóloga de 90% no domínio quinase. Já nos domínios de regulação, no carbono C-terminal, mostram um desvio significativo (RIKITAKE; LIAO, 2005). Ambas estão ubiquamente expressas em vários tecidos humanos e de roedores.

As enzimas ROCK estão implicadas em uma variedade de áreas terapêuticas, incluindo doenças cardiovasculares (SHIMOKAWA; RASHID, 2007), distúrbios do sistema nervoso central (KUBO et al., 2008), inflamação (LOGRASSO; FENG, 2009) e câncer (OGATA et al., 2009), além de relatos no desenvolvimento de diabetes, doenças neurodegenerativas e hipertensão pulmonar (SAMUEL et al., 2011). A expressão de proteínas ROCK em células cancerosas tem sido relatada em câncer de ovário, de pâncreas, testicular e de bexiga (KAMAI et al., 2004). A inibição farmacológica das ROCKs é uma estratégia promissora para a prevenção da invasão das células malignas, um evento central no processo de metástase (ISHIZAKI et al., 2000).

À luz destas pesquisas, é possível observar que a ROCK é uma via crucial na mediação de várias funções celulares e está envolvida em vários processos patológicos. O planejamento e o desenvolvimento de inibidores dessa

enzima certamente serão benéficos para o tratamento de diversas doenças (PAN et al., 2013).

Uma ferramenta que pode ser útil no desenvolvimento de novos compostos protótipos para a inibição/ativação da via Rho/ROCK quinase é a química computacional. Ferramentas como o ancoramento molecular (*docking*) e o estudo das relações quantitativas entre a estrutura química e a atividade biológica (do inglês *quantitative structure activity relationship*, QSAR) destacam-se na química computacional. Por exemplo, na área de química medicinal teórica, a predição da atividade biológica de novos compostos usando relações matemáticas baseadas em propriedades estruturais, físico-químicas e conformacionais de potenciais agentes previamente testados é um campo de pesquisa extremamente ativo e promissor, amplamente reportado na literatura (ANDRADE et al., 2010; ASSIS et al., 2016; BOLAND et al., 2015; CHERKASOV et al., 2014). A metodologia de QSAR é útil para entender e explicar as forças motrizes por trás da ação de fármacos em âmbito molecular e permitir o projeto e o desenvolvimento de novos compostos com propriedades biológicas desejáveis (FERREIRA, 2002), reduzindo custos, tempo e experimentação animal (ORGANIZATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT - OECD, 2010).

O uso de QSAR no estudo de possíveis inibidores da ROCK é uma ferramenta de grande valia na otimização de novas estruturas e melhoria da potência e seletividade dos mesmos, além de possibilitar a análise do papel da enzima em várias funções biológicas.

2 REFERENCIAL TEÓRICO

2.1 Via RhoA/Rho quinase

As primeiras proteínas da superfamília Ras foram identificadas em 1985 (MADAULE; AXEL, 1985). Mais de 100 membros já foram descritos. Cerca de 20 tipos de Rho GTPases foram descritas em mamíferos: Rho (três isoformas: A, B, C); Rac (1, 2, 3); Cdc42; TC10; TCL; Chp (1, 2); RhoG; Rnd (1, 2, 3); RhoBTB (1, 2), RhoD; Rif e TTF. Elas alternam entre um estado ativo (GTP) e inativo (GDP) e em seu ciclo ativo interagem com uma das 60 proteínas alvo chamadas de efetoras (ETIENNE-MANNEVILLE; HALL, 2002; RIDLEY, 2001). Esta família de pequenas GTPases tem sido alvo de muitos estudos, tanto no envolvimento dos processos fisiológicos do organismo como na regulação do tônus muscular liso, na migração e na proliferação celular, quanto na sua participação em diversos processos patológicos, como na hipertensão, na aterosclerose e na disfunção erétil.

Em 1996, a ROCK foi descrita como a proteína alvo (efetora) da RhoA, ou seja, atua facilitando a ativação de seu estado inativo ligado à guanosina difosfato (GDP) para seu estado ativo ligado à guanosina trifosfato (GTP) (Figura 1) (TAKI et al., 2007). Esta descoberta forneceu informações sobre a ligação entre a Rho, a formação de fibras de estresse e de adesão focal e a contração de células da musculatura lisa (LEUNG et al., 1996).

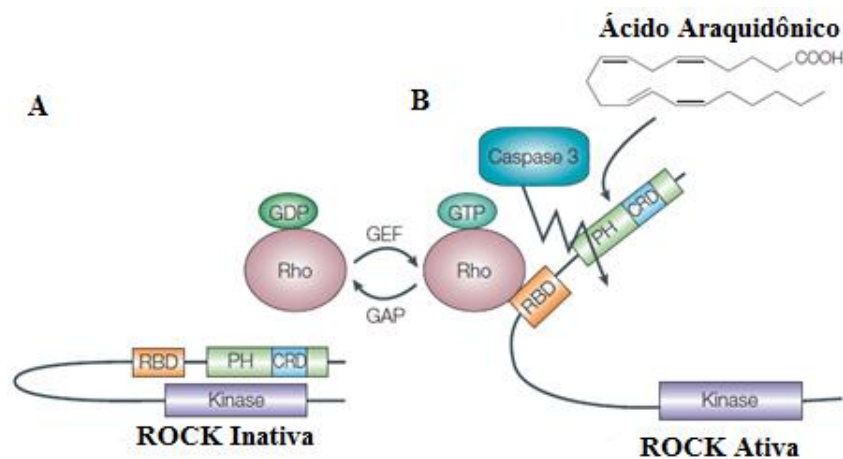


Figura 1 Ativação da enzima ROCK mediada pela RhoA. “A” representa a estrutura inativa da enzima na qual a porção C-terminal inibe o domínio quinase e “B” representa a exposição do domínio quinase proporcionado pela ligação da RhoA-GTP ou de moléculas como ácido araquidônico e enzimas da cascata de apoptose
Adaptado de Riento e Ridley (2003)

A Rho quinase (EC:2.7.11.1) é responsável por afetar a reorganização da actina, a regulação da cadeia leve da miosina (CLM) fosfatase e a fosforilação da CLM. Adicionalmente, parece estar envolvida na regulação da força e da velocidade do cruzamento das pontes de actina e miosina no músculo liso e em células não musculares, por inibir a miosina fosfatase mediada pela desfosforilação da cadeia regulatória da miosina II (WETTSCHURECK; OFFERMANN, 2002). É uma proteína quinase serina/treonina que contém um domínio N-terminal quinase catalítico e pode ser expressa por meio de duas isoformas: a ROCK1 (também conhecida como ROCK β ou p160 ROCK) e a ROCK2 (também chamada de ROCK α) que apresentam 64% e 79% de identidade em relação à sequência e à similaridade global, respectivamente. Em relação a seus domínios quinase, apresentam 92% de identidade e 97% de semelhança. As isoformas são distribuídas em diferentes tecidos; ROCK1 é expressa no pulmão, no fígado, no estômago, no baço, nos rins e nos testículos,

enquanto ROCK2 é altamente expressa em tecidos do cérebro, do coração e dos músculos (NAKAGAWA et al., 1996). Apesar da distribuição do tecido diferenciado, pouco se sabe sobre as diferenças funcionais entre as duas isoformas ROCK (SAPET et al., 2014).

Além da sua ativação pela RhoA, a Rho quinase pode ser ativada no músculo liso pelo ácido araquidônico, em resposta a vários agonistas (Figura 1) (WETTSCHURECK; OFFERMANNNS, 2002).

Seu ciclo, ou a ativação e a inativação da Rho, é regulado em grande parte por três classes de proteínas que são:

- a) fatores de troca guanina nucleotídeo que é específica para Rho (Rho GEFs), nos quais facilitam a ativação da Rho por promover a dissociação de GDP e, subsequentemente, a ligação de GTP. As Rho GEFs são consideradas as principais reguladoras da atividade da Rho;
- b) as proteínas ativadoras de GTPase (GAPs), que facilitam a inativação da forma GTP ativa pelo aumento da atividade intrínseca da GTPase na Rho;
- c) as inibidoras da dissociação de GTPase (GDIs), que inibem algumas GTPases da família Rho por ligação a membranas e previnem a dissociação e, portanto, a ativação de nucleotídeos.

A atividade da Rho também pode ser regulada por numerosos receptores de proteína G, a qual atua na transdução de sinais celulares (SCHAAFSMA et al., 2008; WETTSCHURECK; OFFERMANNNS, 2002). Adicionalmente, a ativação da RhoA é fortemente dependente de sua modificação pós-translacional (isoprenilação), mediada pela atividade de geranylgeranyl transferase (CASEY, 1995).

A ativação anormal da via Rho/Rho quinase tem se mostrado importante no acionamento de vias que contribuem para a fisiopatologia de algumas doenças, tais como hipertensão arterial e pulmonar, glaucoma, diabetes, disfunção erétil, neurodegeneração e câncer (NOSSAMAN et al., 2010; NUNES; RIGSBY; WEBB, 2010). No sistema respiratório, sua participação vem sendo descrita em doenças como a asma e a doença pulmonar obstrutiva crônica (DPOC) (FERNANDES; HENRY; GOLDIE, 2007).

2.2 Principais inibidores da Rho quinase

O envolvimento da via Rho quinase em diversas doenças tem despertado crescente interesse pelo potencial farmacológico dos inibidores da ROCK para o tratamento dessas patologias. Diversos estudos já foram realizados *in vitro* e *in vivo* utilizando estes inibidores químicos (Tabela 1), principalmente o (+)-HA-trans-4-(1-aminoetil)-N-(4-piridil)cicloexano-carboxamidadi-hidro-cloreto (Y-27632) e o fasudil (HA-1077, AT877 ou hidro-1-(5-isoquinolinasulfonil)-1H-1,4-diazepina cloridrato) (Figura 2) (IIZUKA et al., 2000; PIREDDU et al., 2012).

Y-27632 e o Y-30141 (Figura 2) são conhecidos como inibidores seletivos da Rho quinase, quando comparados a outras quinases (IIZUKA et al., 2000). Eles são permeáveis na célula e agem como um competidor com o local de ativação ATP na Rho quinase, tendo ação inibitória nas suas duas isoformas, ROCK1 e ROCK2 (DARENFED et al., 2007). O Y-27632 inibe a contração do músculo liso e mostra eficácia na normalização da pressão arterial em modelos de hipertensão.

O composto CID5056270 (Figura 2) também é relatado como um potente inibidor da Rho quinase (AUSTIN, 2004).

Vale ressaltar que Fasudil (Figura 2) tem afinidade pela Rho semelhante ao Y-27632, porém, com menor seletividade (SCHAAFSMA et al., 2008). Curiosamente, o hidroxifasudil é metabolizado *in vivo* a partir de quinolonas. Este é, atualmente, o único inibidor de Rho quinase disponível para uso clínico. Desde 1995 ele é utilizado, no Japão, para a prevenção de vasoespasmos em pacientes com hemorragia subaracnóidea (MONG; WANG, 2009). Também demonstrou eficácia no tratamento de hipertensão pulmonar (DOGGRELL, 2005) e Alzheimer (HUENTELMAN et al., 2010).

Tabela 1 Avaliação de inibidores da ROCK em fase de testes farmacológicos para diferentes patologias. Adaptado de Pan et al. (2013)

Patologias	Compostos	Efeitos	Animais
Hipertensão	Y27637	Regulação da pressão sanguínea	Rato
	Fasudil Sar407899		Humano Rato
Diabetes	Fasudil	Prevenção no desenvolvimento da diabetes	Rato
Arteriosclerose	Fasudil	Regressão das lesões coronárias ateroscleróticas	Porco
Infarto do miocárdio	Fasudil	Supressão do ventrículo esquerdo, dilatação da cavidade e disfunção	Camundongo
Falência cardíaca	Fasudil	Aumento da resistência e vasodilatação debilitada da vasculatura periférica	Humano
Glaucoma	Fasudil AMA0076	Redução da pressão intraocular	Humano
Lesão na coluna espinhal	Y27637	Estimulação da regeneração do axônio e da função dos membros posteriores	Camundongo
	Fasudil Y27637	Recuperação neurológica Redução da prevalência	Humano Camundongo
Alzheimer	Hidroxfasudil	Melhoria da aprendizagem espacial e memória de trabalho	Rato
		Proteção contra neurodegeneração no hipocampo	Rato
Acidente vascular cerebral	Fasudil	Melhoria das funções neurológicas	Humano
Câncer	Fasudil	Inibição do crescimento tumoral, progressão e metástase	Rato, camundongo
	Y27637		
	RKI1447 Wf536		

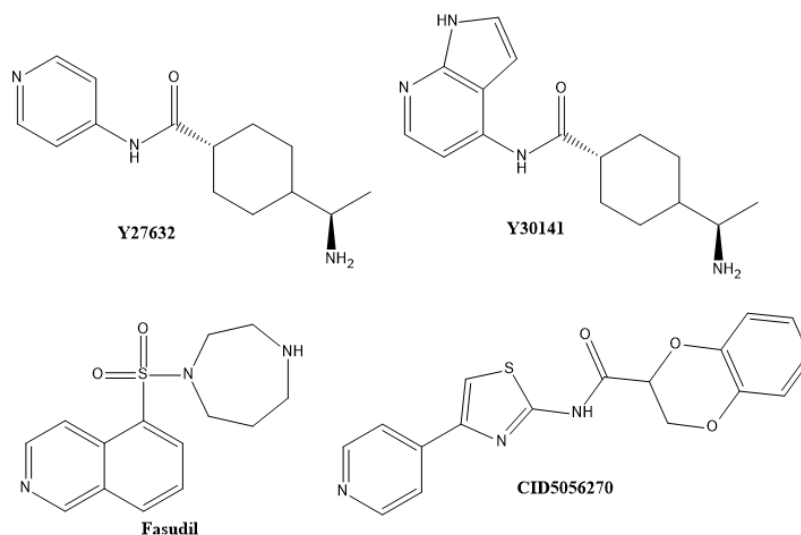


Figura 2 Exemplos de inibidores da ROCK: Y-27632, Y-30141, fasudil e CID5056270

Relatos encontrados na literatura destacam o risco da toxicidade fora do alvo para inibidores quinase em geral (HARMSEN; KOK, 2012). De fato, o composto Y-27632, que é um padrão largamente utilizado nos testes de inibição da ROCK, ainda não foi aprovado para utilização clínica, provavelmente devido ao seu elevado perfil de toxicidade aguda (categoria 4) (SIGMA-ALDRICH, 2016). Além disso, este inibidor imita a apoptose induzida por estatina hidrofóbica, causando miotoxicidade em fibroblastos em modelos de ratos (ITAGAKI et al., 2009).

Devido às potenciais aplicações terapêuticas, o grupo de pesquisa de Pireddu e colaboradores (PIREDDU et al., 2012) tomou como ponto de partida o composto CID5056270 (Figura 2) para a concepção de uma série de compostos inibidores à base de piridiltiazol, buscando desenvolver compostos mais potentes, seletivos e com menor toxicidade. Os resultados preliminares foram positivos e, por este motivo, esta série de inibidores foi a escolhida para dar desenvolvimento a este projeto.

A química computacional é muito utilizada para o planejamento de fármacos. A possibilidade de projetar compostos com propriedades definidas, evitando os custos da síntese experimental exploratória de um grande número de substâncias, tem impulsionado as pesquisas nessa área. Os fundamentos necessários para um projeto efetivo nessa área estão na relação quantitativa estrutura-atividade (QSAR) (MARTINS et al., 2009).

2.3 Relação quantitativa estrutura-atividade

Nas técnicas utilizadas em QSAR considera-se que existe relação entre as propriedades de uma molécula e sua estrutura e tenta-se estabelecer relações matemáticas simples para descrever e, em seguida, prever uma dada propriedade para um conjunto de compostos, geralmente pertencentes a uma mesma família química (MILANI, 2011). O estudo de QSAR compreende também a definição dos descritores moleculares capazes de caracterizar satisfatoriamente conjuntos moleculares diferentes e o tratamento estatístico que pode ser aplicado a esses descritores, no intuito de melhorar sua capacidade preditiva (MARTINS et al., 2009). Portanto, a atividade biológica de um fármaco está relacionada com as características moleculares específicas, que são denominadas descritores (ANDRADE et al., 2010).

Em 1960, os estudos de QSAR baseavam-se em dados experimentais considerando-se uma única propriedade físico-química, como a solubilidade ou ponto de fusão, para explicar o efeito biológico da molécula (QSAR-1D). Hansch e Fujita (1964) passaram a considerar que a atividade biológica observada é resultado da contribuição de diversas propriedades físico-químicas (QSAR-2D) e seu modelo tenta identificar a contribuição de cada fator no estabelecimento da resposta biológica (TAVARES, 2004). Informações significativas sobre um conjunto de dados podem ser extraídas utilizando-se o

QSAR-2D, entretanto, esses descritores não têm informações relacionadas com as características tridimensionais dos compostos (OPREA, 2002).

No ano de 1980, a estrutura tridimensional do ligante começou a ser considerada, dando origem ao QSAR-3D. Os modelos quantitativos utilizando a terceira dimensão foram desenvolvidos e relacionam a atividade biológica com suas propriedades calculadas no espaço tridimensional (3D). Este modelo fornece uma informação estrutural completa, analisando as propriedades dos compostos por campos moleculares que descrevem interações estéricas, eletrostáticas, hidrofóbicas e por ligação de hidrogênio, calculadas a partir de sua suposta conformação bioativa (KUZ'MIN; ARTEMENKO; MURATOV, 2008). Em 1988, o método de análise comparativa de campo molecular (CoMFA, do inglês *comparative molecular field analysis*) foi publicado por Cramer, Patterson e Bunce (1988). Este método, que é baseado em campo molecular, constituiu o primeiro verdadeiro método de QSAR-3D (KUBINYI, 1997). Uma grande dificuldade da metodologia de QSAR-3D é encontrar um alinhamento apropriado para os compostos, além de os estudos produzirem uma enorme quantidade de variáveis, em que nem sempre todas são significativas para o resultado final (OOMS, 2000).

Nos anos 1990, o QSAR-4D foi proposto por Hopfinger e colaboradores (HOPFINGER et al., 1997), considerando o espaço conformacional dos ligantes. Este modelo é semelhante ao modelo 3D, porém, a informação estrutural é obtida por meio de um conjunto de conformações, orientações e estados de protonação, obtidos por cálculos de dinâmica molecular (YANMAZ et al., 2011).

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). Com o QSAR-6D, a estimativa simultânea

de diferentes cenários de solvatação pode ser considerada (VEDANI et al., 2007). Finalmente, com o método mais recentemente desenvolvido, o QSAR-7D, emprega-se o real receptor alvo ou o modelo do receptor baseado no alvo, obtidos por modelagem de proteínas homólogas (POLANSKI, 2009).

Os modelos QSAR são construídos a partir de três componentes: (1) um conjunto de dados de atividade relacionada a um grupo de substâncias, geralmente obtido experimentalmente; (2) um conjunto de dados de propriedades relacionadas à estrutura ou critérios estruturais de um mesmo grupo de compostos químicos e (3) a medida da relação entre as duas anteriores, obtida por tratamento estatístico (OECD, 2010).

Devido ao fato de serem consideradas, em geral, muitas variáveis para serem correlacionadas com a atividade biológica, é necessário um tratamento estatístico dos dados para se selecionar aquela com importância estatística significativa. Os métodos mais usados são: regressão linear múltipla (MLR, sigla do inglês *multiple linear regression*) e regressão pelo método dos mínimos quadrados parciais (PLS, do inglês *partial least squares*) (OOMS, 2000).

Além dos métodos de QSAR citados, diversas outras abordagens vêm sendo utilizadas para esses estudos, dentre as quais se destacam CoMSIA (KLEBE; ABRAHAM; MIETZNER, 1994), HQSAR (LOWIS, 1997), MIA-QSAR (NUNES et al., 2012) e LQTA-QSAR (MARTINS et al., 2009), objetivando a obtenção de modelos preditivos.

A *Organization for Economic Co-operation and Development* (OECD) estabeleceu cinco princípios para validação de modelos de QSAR para uso prático de agências de regulamentação. Esses princípios são: (1) atividade biológica ou propriedade definida; (2) algoritmo claro; (3) domínio de aplicabilidade (DA) definido; (4) avaliação apropriada da robustez e preditividade e (5) interpretação mecanística, que significa encontrar relações entre os descritores e a atividade biológica ou propriedade, em via de se

compreender melhor o mecanismo de ação de uma estrutura química ou aprofundar o conhecimento biológico sobre a propriedade em estudo (OECD, 2010).

2.4 QSAR 4D

Hopfinger et al. (1997) propuseram uma nova metodologia de QSAR chamada de QSAR-4D. A quarta-dimensão do QSAR-4D corresponde à amostragem do conjunto das conformações (*ensemble sampling*), ou seja, o 4D incorpora liberdade conformacional ao desenvolvimento de modelos de QSAR 3D, fazendo com que a mudança de estado molecular constitua a quarta dimensão.

A metodologia de QSAR-4D é semelhante à metodologia do CoMFA, baseando-se no alinhamento de moléculas em uma caixa cúbica virtual que é utilizada para gerar um conjunto de descritores. Contudo, quando esta metodologia é aplicada a certos compostos e seus derivados, cujas atividades são medidas e a geometria do receptor não é avaliada, este estudo é denominado QSAR-4D, independente do receptor (QSAR-4D-IR) (HOPFINGER et al., 1997). Ao serem levadas em consideração a geometria do receptor e suas modificações necessárias para interação com os ligantes, tem-se a análise de QSAR-4D dependente do receptor (QSAR-4D-DR) (SENESE et al., 2004).

Os descritores em QSAR-4D são representados pelas medidas de ocupação de cada célula da caixa virtual pelos átomos que formam as moléculas do conjunto de treinamento. Os descritores de ocupação das células da caixa, *grid cell occupancy descriptors* (GCODs), podem ser gerados a partir de diferentes tipos de átomos, chamados de *interaction pharmacophore elements* (IPes).

Recentemente, Martins et al. (2009) desenvolveram uma nova metodologia de QSAR-4D, chamada de LQTA-QSAR, que incorpora as principais vantagens observadas nos métodos CoMFA e QSAR-4D. A nova metodologia calcula os descritores de campo, a exemplo do que acontece em CoMFA, para um perfil de amostragem conformacional, assim como o utilizado em QSAR-4D.

A grande vantagem da metodologia de QSAR-4D em relação aos métodos de QSAR-3D mais difundidos (e.g., COMFA) é o emprego da amostragem conformacional e o campo de interação dos descritores.

3 OBJETIVOS

Este trabalho foi realizado com os objetivos de

- a) realizar estudos quantitativos de estrutura-atividade em 4D de compostos piridiltiazólicos, por meio do programa 4D-QSAR;
- b) propor novas moléculas candidatas a inibidores da enzima ROCK.

4 METODOLOGIA

4.1 Coleta de dados biológicos

Por meio de uma pesquisa no banco de dados de artigos e patentes *Web of Knowledge*, selecionou-se uma publicação com potenciais inibidores da enzima Rho quinase que apresentassem os respectivos valores de IC_{50} (concentração necessária para inibir 50% da atividade enzimática). Uma série de 51 compostos análogos ao piridiltiazol, desenvolvidos por Lawrence, Pireddu e Sebti (2013), foi selecionada por apresentar inibição para a isoforma ROCK1 (Tabela 2). Os valores de IC_{50} expressos em μM foram convertidos para M e, então, transformados para $-\log IC_{50}(pIC_{50})$.

Tabela 2 Estrutura dos 51 compostos análogos ao piridiltiazol e seus valores de IC_{50} (LAWRENCE; PIREDDU; SEBTI, 2013)

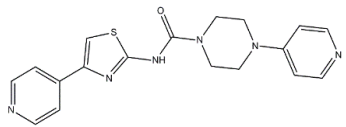
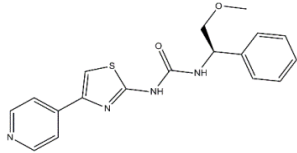
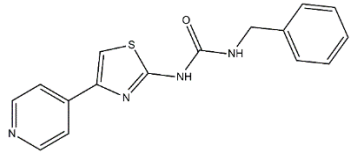
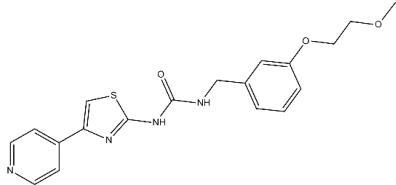
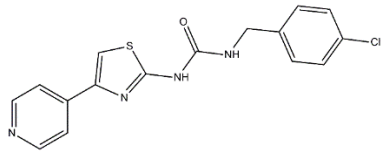
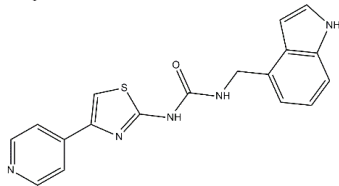
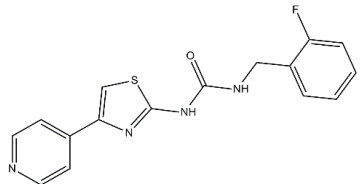
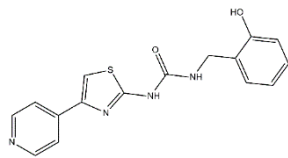
N°	Estrutura	IC_{50}	pIC_{50}	N°	Estrutura	IC_{50}	pIC_{50}
1		16,89	4,772	27		29,48	4,530
2		0,15	6,824	28		0,50	6,301
3		2,39	5,620	29		37,79	4,423
4		0,40	6,398	30		0,62	6,208

Tabela 3, continuação

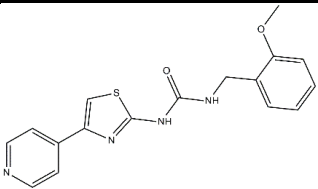
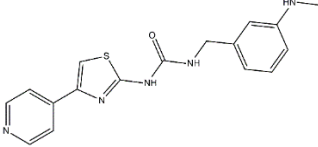
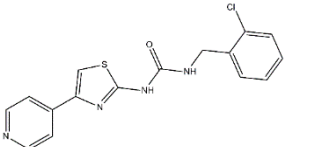
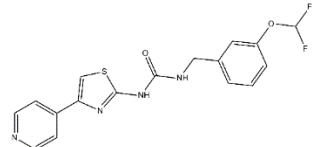
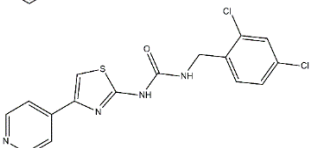
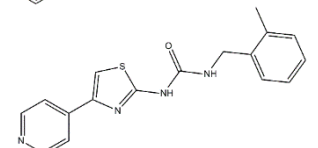
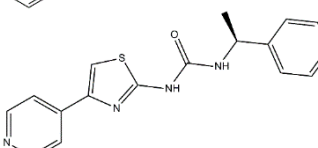
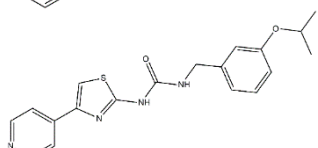
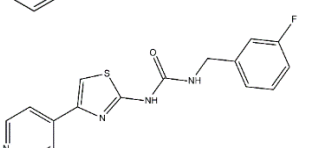
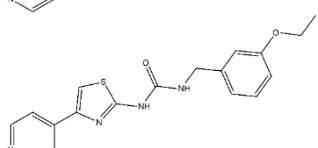
Nº	Estrutura	IC ₅₀	pIC ₅₀	Nº	Estrutura	IC ₅₀	pIC ₅₀
5		1,40	5,854	31		0,47	6,327
6		0,34	6,469	32		0,92	6,036
7		1,88	5,726	33		0,15	6,824
8		3,56	5,449	34		13,00	4,886
9		0,14	6,854	35		2,14	5,670

Tabela 4, continuação

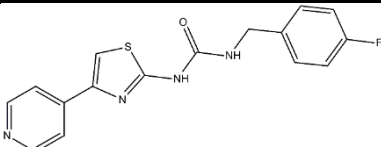
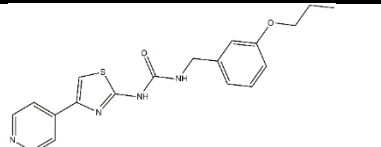
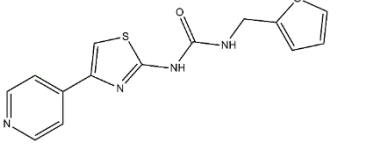
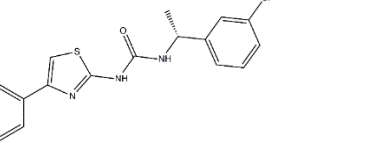
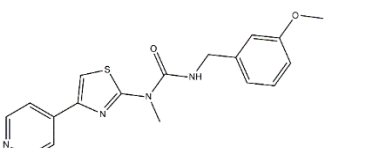
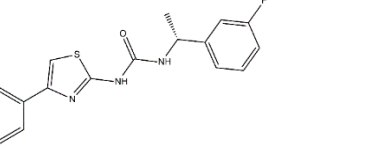
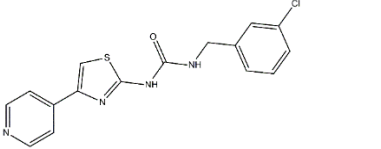
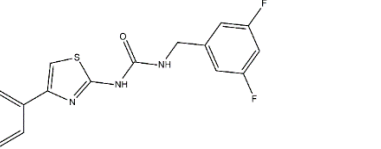
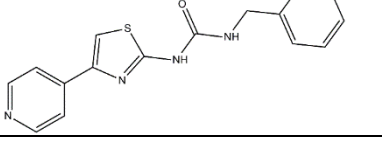
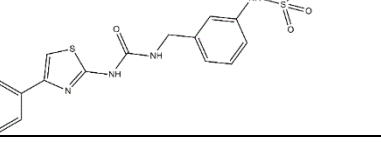
Nº	Estrutura	IC ₅₀	pIC ₅₀	Nº	Estrutura	IC ₅₀	pIC ₅₀
10		0,57	6,244	36		3,36	5,474
11		1,17	5,931	37		0,55	6,260
12		0,06	7,220	38		0,11	6,959
13		0,14	6,854	39		4,07	5,390
14		0,91	6,041	40		0,56	6,252

Tabela 5, continuação

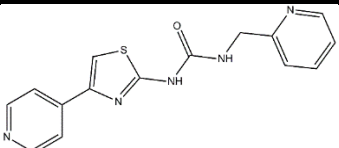
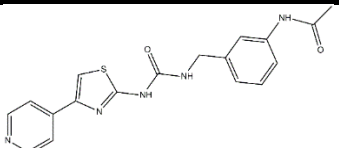
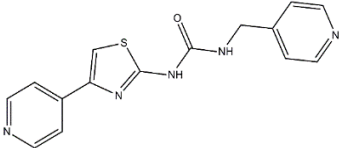
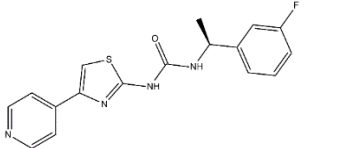
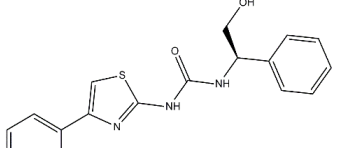
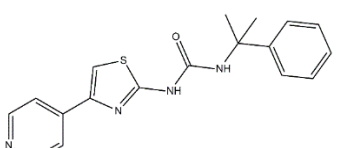
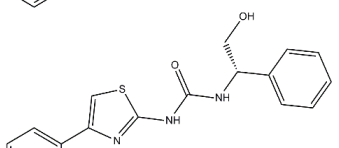
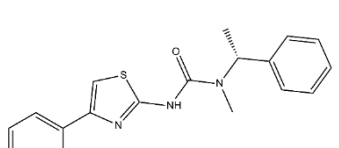
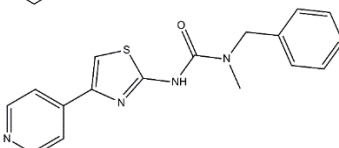
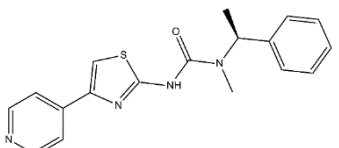
Nº	Estrutura	IC ₅₀	pIC ₅₀	Nº	Estrutura	IC ₅₀	pIC ₅₀
15		1,82	5,740	41		1,25	5,903
16		2,07	5,680	42		1,25	5,940
17		5,20	5,284	43		0,19	6,720
18		0,03	7,523	44		0,12	6,921
19		0,11	6,959	45		6,44	5,191

Tabela 6, continuação

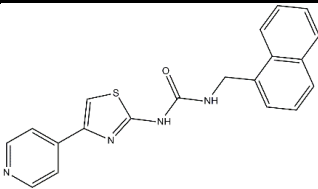
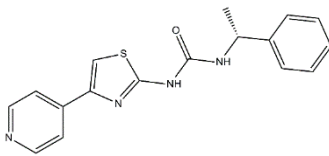
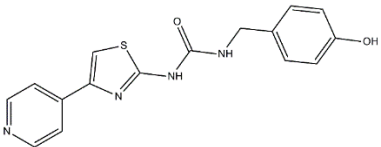
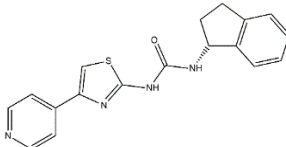
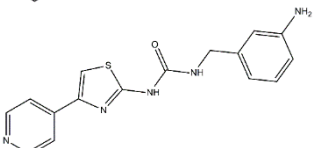
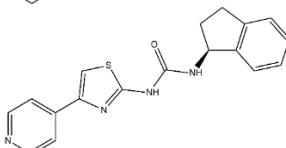
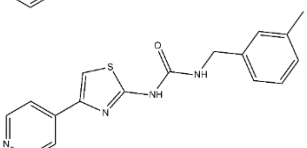
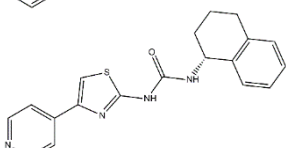
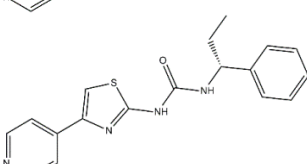
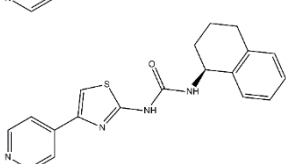
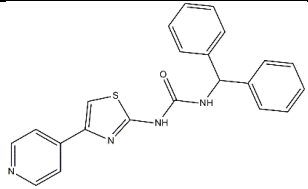
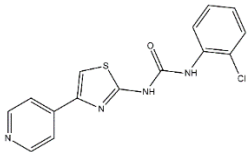
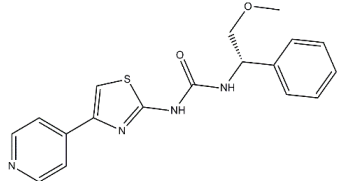
Nº	Estrutura	IC ₅₀	pIC ₅₀	Nº	Estrutura	IC ₅₀	pIC ₅₀
20		1,10	5,959	46		3,87	5,412
21		0,38	6,420	47		28,24	4,549
22		0,05	7,300	48		3,21	5,493
23		0,46	6,337	49		6,99	5,160
24		0,09	7,045	50		6,25	5,204

Tabela 7, conclusão

Nº	Estrutura	IC ₅₀	pIC ₅₀	Nº	Estrutura	IC ₅₀	pIC ₅₀
25		1,87	5,728	51		28,24	4,549
26		0,07	7,155				

Os modelos de QSAR-4D foram construídos utilizando-se um conjunto de 42 compostos denominados de conjunto de treinamento e validados externamente empregando-se um conjunto de nove compostos, denominados de conjunto teste.

A divisão entre compostos do conjunto de treinamento e do conjunto teste foi feita separando-se os 51 compostos, estruturalmente relacionados, em três subgrupos, de acordo com faixas de valores de atividade biológica e selecionando-se, aleatoriamente, aproximadamente 25% dos compostos de cada subgrupo para compor o conjunto teste e os remanescentes correspondem ao conjunto de treinamento. Foram realizadas três seleções aleatórias de grupos testes, a fim de obter maior credibilidade ao modelo. O grupo teste I contém as moléculas **1, 22, 30, 31, 35, 38, 39, 41 e 43**; o grupo teste II apresenta as moléculas **3, 5, 6, 14, 17, 18, 28, 44 e 47**, e o grupo teste III inclui as moléculas **9, 12, 15, 21, 32, 37, 48, 50 e 51** (Tabela 2).

4.1.2 Obtenção da estrutura cristalizada de Rho Quinase

As coordenadas da estrutura cristalográfica da enzima ROCK (Figura 3) (PIREDDU et al., 2012) foram obtidas no banco de dados de proteínas *Protein Data Bank* (código PDB: 3TV7) (BERMAN et al., 2000). Esta enzima é composta de quatro cadeias e apresenta resolução de 2,75 Å, estequiometria monomérica e simetria cíclica c_2 .

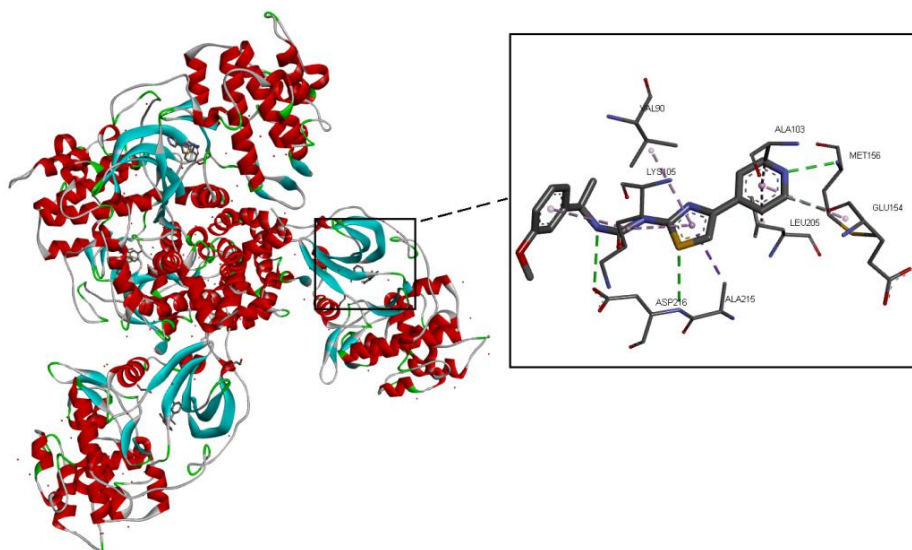


Figura 3 Enzima ROCK disponível no PDB (Código 3TV7). Em destaque, o ligante 1-[(1R)-1-(3-metoxifenil)etil]-3-(4-piridin-4-il-1,3-tiazol-2-il)urea da cadeia A e suas interações com os resíduos próximos

4.1.3 Construção e minimização dos modelos tridimensionais dos inibidores

A enzima foi previamente cortada em um raio de 10 Å ao redor do ligante. Esta distância é suficiente para analisar interações a curto (<4,5 Å) e longo alcance (>4,5 Å). Em seguida, adicionaram-se átomos de hidrogênio para manter a integridade geométrica e estrutural do receptor. Esta dimensão foi suficiente para selecionar os resíduos de aminoácidos que fazem parte do sítio de interação da ROCK1. Utilizou-se como base o ligante disponível no sítio catalítico para a construção das estruturas tridimensionais dos 51 análogos (Tabela 1). Eles foram otimizados via mecânica molecular através do campo de força CHARMM (BROOKS et al., 1983) e as cargas atômicas parciais calculadas por meio do *software Discovery Studio* (Discovery Studio 2.1, Accelrys Inc., San Diego, CA, USA), utilizado para todos os passos anteriormente descritos.

4.2 QSAR

Os compostos em 3D foram transferidos para o programa *4D-QSAR* (THE CHEM21 GROUP INC, 1997), a fim de investigar a correlação estrutura-atividade.

4.2.1 Simulação de dinâmica molecular

A quarta dimensão na metodologia de QSAR-4D corresponde à amostragem conformacional em função do tempo de simulação de dinâmica molecular (SDM). As estruturas dos complexos otimizados no sítio ativo da Rho Quinase foram submetidas ao processo de SDM, com o objetivo de gerar um perfil de amostragem conformacional (CEP, do inglês *conformational ensemble profile*) para cada composto, permitindo, assim, investigar a flexibilidade conformacional do ligante (THE CHEM21 GROUP INC, 1997; HOPFINGER et al., 1997).

Este método calcula o comportamento do sistema por meio das equações de movimento de Newton para um intervalo de tempo discreto. O conjunto de configurações de simulação irá formar a trajetória de átomos de sistema que fornecem conformações termodinamicamente acessíveis para os compostos em análise (SPOEL et al., 2005).

Essas conformações serão utilizadas no formalismo QSAR-4D. Então, é necessária uma estrutura contendo a posição inicial para todos os átomos do sistema; o próximo passo é calcular a força que age em cada átomo por meio da expressão da derivada da energia potencial V em relação às posições atômicas, como pode ser visto na equação 1.

$$F = -\frac{\partial V}{\partial q} \quad (1)$$

Com as informações de força e velocidade são calculadas as energias potenciais e cinéticas. Os potenciais de interação utilizados para calcular V_{total} são derivados de um campo de força molecular escolhido para realizar as simulações. Nesse caso, a energia é assumida como sendo a soma dos termos potenciais que descrevem o estiramento de ligações covalentes, as deformações angulares, as torções de ligação e as interações de van der Waals e eletrostáticas, entre outras, como pode ser visto na Equação 2.

$$V_{tot} = V_{estiramento} + V_{def\ angular} + V_{torção} + V_{vdW} + V_{coulomb} + \dots \quad (2)$$

Os cálculos de dinâmica foram executados à temperatura de 300 K, a fim de permanecer próximo à temperatura utilizada nos ensaios biológicos. O tempo de amostragem da simulação foi de 50 ps, com intervalos de 0,001 ps, gerando, assim, 500 conformações de cada composto. Aplicou-se também uma função dielétrica dependente da distância, $\epsilon_r = 3/r_{ij}$, no intuito de simular o efeito do solvente. Adicionalmente, os átomos correspondentes ao carbono alfa da proteína e o átomo de enxofre comum a todos os compostos foram fixados para prevenir uma grande mudança conformacional dos ligantes, o que poderia resultar em conformações estericamente proibidas.

4.2.2 Definição do alinhamento

O alinhamento é uma das etapas mais importantes na metodologia de QSAR-4D, pois ele é o caminho no qual os compostos do conjunto de

treinamento e teste são comparados. Neste trabalho, pressupõe-se que todos os compostos da série em estudo apresentam o mesmo modo de ligação no sítio ativo Rho quinase, atribuído ao fato de que os compostos são análogos estruturais e que, portanto, devem apresentar os mesmos tipos de interação, visto que são capazes de inibir a atividade da enzima. Foram realizados dez alinhamentos, utilizando-se as ordenadas pertencentes aos átomos de C α dos resíduos de aminoácidos presentes no sítio catalítico da proteína e a um átomo do ligante. São eles: (1) Ala103, Lys105, S25; (2) S25, N21, Ala103; (3) Ala103, Met153, S25; (4) Ala103, S25, Lys105; (5) S25, Ala103, Lys105; (6) Lys105, Ala103, S25; (7) S25, Lys105, Ala103 (8) Lys105, S25, Ala103 (9) S25, Ala103, N21 e (10) Ala103, S25, Met153 (Figura 4).

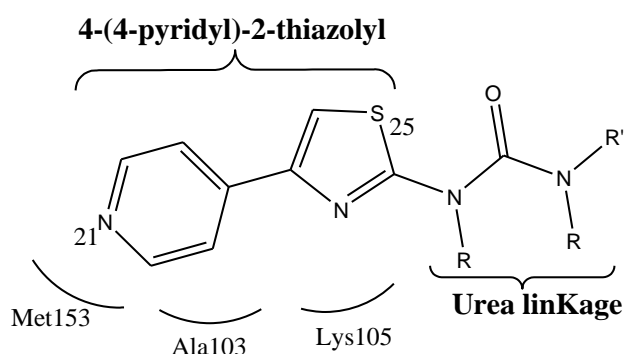


Figura 4 Representação dos átomos fixados nos alinhamentos destacados

4.2.3 Definição do tamanho das células

O CEP para cada composto obtido por SDM foi inserido dentro de uma caixa composta por células cúbicas que podem ter dimensões variadas. Quanto menor o tamanho da célula, mais refinados serão os modelos, entretanto, o tempo de análise aumenta devido ao maior número de variáveis geradas. O tamanho da célula de grade foi definido como 2 Å, o que corresponde ao número

inteiro mais próximo de duas vezes ao raio de van der Waals do átomo de hidrogênio ($r_{vdw} = 1,2 \text{ \AA}$).

4.2.4 Definição dos elementos de interação farmacofórica

Os átomos de cada composto foram classificados em seis tipos de elementos de interação farmacofórica (IPE, do inglês *interaction pharmacophore element*), correspondendo aos tipos de átomos que podem ocupar cada célula da caixa, de acordo com a metodologia de QSAR-4D (HOPFINGER et al., 1997). Os IPEs são classificados em não polares (np), polares com carga positiva (p+), polares com carga negativa (p-), aceptores de ligação hidrogênio (ha), doadores de ligação hidrogênio (hd) e sistemas aromáticos (ar).

A frequência de ocupação de cada célula da caixa por cada tipo de IPE forma o conjunto de descritores de ocupação da célula (GCOD, *Grid Cell Occupancy Descriptor*) que será utilizado na construção dos modelos de QSAR-4D.

4.2.5 Obtenção dos modelos de QSAR-4D

A fim de excluir ruído ou dados inúteis, bancos de dados nomeados DB1, DB2 e DB3 foram gerados. A base de dados DB1 foi construída excluindo-se as variáveis em que GCODs são iguais a zero para todas as moléculas; os bancos de dados DB2 e DB3 foram construídos excluindo-se as variáveis em que GCODs têm valores de corte de variância de até 0,1 e 1,0, respectivamente.

Os GCODs com alto peso em cada banco de dados provenientes da redução de dados foram processados em conjunto com os valores de atividade

biológica (variável dependente), utilizando-se uma metodologia que combina algoritmos genéticos (GA, do inglês *genetic algorithm*) (JIN; HOPFINGER, 1994) e mínimos quadrados parciais (PLS, do inglês *partial least squares*), denominada de aproximação da função genética (GFA, do inglês *genetic function approximation*). As grandes vantagens do GFA são a construção e a otimização de múltiplos modelos, em vez da otimização de apenas um único modelo (ROGERS; HOPFINGER, 1994).

Os cálculos de GFA foram realizados no programa Wolf 6.2 (ROGERS; HOPFINGER, 1994) e iniciados com 100 modelos (equações) gerados aleatoriamente, com taxa de probabilidade de mutação de 50%. O *crossover* foi variado entre 50.000 e 90.000 e o *smoothing factor*, que controla o número de descritores (GCODs) nos modelos de QSAR, foi variado de 0,1 a 0,5, no intuito de se obter equações com, no máximo, oito termos. Equações com mais de oito termos possivelmente provocará um superajuste (*overfitting*) nos dados, pois o número máximo de termos pode ser definido pelo número de compostos utilizado no conjunto de treinamento, dividido por 5 (KUBINYI, 2008).

Os modelos gerados para cada alinhamento foram ordenados pelo valor do fator de desajuste (LOF, do inglês *lack-of-fit*) (Equação 3) de Friedman, que é a medida dos mínimos quadrados penalizados; quando dois métodos têm o mesmo erro de mínimos quadrados, aquele que tiver o menor número de termos será escolhido (ROGERS; HOPFINGER, 1994).

$$LOF = \frac{LSE}{1 - c + \frac{dp}{M}} \quad (3)$$

em que **c** é o número de termos no modelo, **d** é *smoothing factor*, **p** é o número total de variáveis contido em cada termo e **M** é o número de compostos, **LSE** erro.

A validação é um passo crucial para avaliar a robustez dos modelos de QSAR. O primeiro índice que se utiliza é o coeficiente de determinação (r^2), que é uma medida de ajuste de um modelo estatístico linear generalizado em relação aos valores observados. Um valor de r^2 maior que 0,7 indica que o modelo está correlacionado e pode ser utilizado para calcular a atividade do conjunto de teste.

4.2.6 Validação interna dos modelos

Os melhores modelos obtidos por GFA foram submetidos à validação interna (validação cruzada). A técnica de validação interna é denominada de *leave-one-out cross-validation* (LOO-cv, q^2), na qual cada composto do conjunto de treinamento é eliminado por vez e os coeficientes da equação são recalculados, até que todos os componentes tenham sido eliminados uma vez. O índice estatístico gerado nesse processo é o coeficiente de correlação quadrático da validação cruzada, q^2 ou r^2_{cv} (SENESE et al., 2004).

Os valores de q^2 devem ser transformados em $q^2_{ajustado}$, demonstrado pela Equação 4, o que permite a comparação entre modelos com número diferente de variáveis e a punição quando o modelo inclui variáveis irrelevantes. Modelos com valores de $q^2_{ajustado}$ maiores que 0,5 são considerados bons modelos preditivos (CALDAS; RAMALHO; CUNHA, 2014).

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

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

4.2.7 Validação externa dos modelos

Os melhores modelos obtidos de cada alinhamento foram submetidos à validação externa, utilizando-se o conjunto de teste contendo nove compostos (grupo teste I: **1, 22, 30, 31, 35, 38, 39, 41 e 43**; grupo teste II: **3, 5, 6, 14, 17, 18, 28, 44 e 47** e grupo teste III: **9, 12, 15, 21, 32, 37, 48, 50 e 51**, disponíveis na Tabela 1) que não foram incluídos no desenvolvimento dos modelos de QSAR-4D. As equações são geradas com base nos compostos do conjunto de treinamento e a capacidade preditiva dos modelos é julgada com base na previsão R^2_{pred} . Este processo verificará a capacidade do modelo de prever a atividade biológica de moléculas congêneres.

O valor de R^2_{pred} é altamente dependente da média do grupo de treinamento. Então, somente o seu uso não garante uma verdadeira reflexão da predição externa. Portanto, Roy et al. (2009) definiram uma equação modificada para R^2_{pred} que dá a importância para a diferença entre r^2 e r_0^2 . Esta Equação está apresentada abaixo (Equação 5).

$$r_m^2 = r^2 \left(1 - \frac{r_0^2}{r^2} \right) \quad (5)$$

em que r^2 é o quadrado do coeficiente de correlação entre os valores observados e preditos e r_0^2 é o quadrado do coeficiente de correlação entre os valores observados e preditos com interseção no eixo zero. Valores maiores que 0,5 podem ser um bom indicativo de uma boa previsibilidade externa.

Validação *Y-randomization* ($R^2_{Y\text{-rand}}$) e R^2_P

A significância estatística da relação entre a atividade biológica e os descritores pode ser verificada pelo teste de aleatoriedade, que consiste na troca

dos valores das variáveis independentes de forma aleatória. Para um modelo de QSAR aceitável, o valor de $R^2_{Y\text{-rand}}$ deve ser menor que o coeficiente de correlação do modelo não randomizado. Utiliza-se um parâmetro que penaliza o modelo pela diferença entre $R^2_{Y\text{-rand}}$ e R^2 . Este parâmetro pode ser calculado pela Equação 6.

$$R_p^2 = r^2 \frac{r^2 - R_{Y\text{-rand}}^2}{r^2} \quad (6)$$

Na equação, r^2 é o coeficiente de determinação e $R^2_{Y\text{-rand}}$ é o valor encontrado após a aleatoriedade. O valor encontrado para R_p^2 deve ser maior que 0,5, para um modelo aceitável (ROY et al., 2009).

4.2.8 Definição do domínio de aplicabilidade

Um modelo de QSAR deve idealmente prever a propriedade alvo de qualquer composto, desde que cada descritor químico possa ser calculado. Porém, os modelos preveem a atividade de novas moléculas a partir da interpolação dos compostos do conjunto treinamento, o que pode implicar num erro, caso as estruturas químicas sejam muito diferentes das estruturas usadas na construção do modelo (TROPSHA, 2010).

Para impedir a extrapolação extrema por parte dos modelos, deve ser calculado o domínio de aplicação (AD, do inglês *applicability domain*), ou limite de similaridade, por meio da Equação 7.

$$AD = \bar{y} + Z\sigma \quad (7)$$

em que \bar{y} é a média da distância euclidiana dos k vizinhos mais próximos para cada composto do conjunto treinamento; Z é um parâmetro para controle do nível de significância e flexibilidade da extrapolação e σ é o desvio padrão das distâncias euclidianas.

O valor de Z foi definido como 0,5, valor este que introduz o limite para o qual cada composto será predito, representando, assim, metade do desvio padrão mencionado acima. Se a distância de um composto externo até pelo menos um dos seus vizinhos no conjunto treinamento exceder esse limite, a predição é considerada inviável (TROPSHA; GOLBRAIKH, 2007).

4.3 Otimização estrutural e *drug-likeness*

Com base na interpretação dos descritores obtidos pelo modelo de QSAR, foi possível propor alterações estruturais nos compostos estudados, empregando técnicas como bioisosterismo e hibridização molecular, a fim de otimizar a sua afinidade com a ROCK1.

O bioisosterismo é uma estratégia de modificação molecular de um protótipo baseada na troca de fragmentos moleculares que apresentem propriedades físico-químicas similares, como, por exemplo, volumes moleculares, formas, distribuições eletrônicas e propriedades físico-químicas semelhantes (FRIEDMAN, 1951). Já a hibridização molecular visa à reunião de características estruturais parciais de duas substâncias biotivas distintas em uma nova molécula. Esta reunião origina uma nova substância que poderá apresentar a atividade de uma das substâncias originais ou conjugar ambas as atividades em uma única molécula (BARREIRO; FRAGA, 2015).

Drug-likeness é um equilíbrio complexo de várias propriedades moleculares e características de estrutura particular que determinam se a molécula é semelhante às drogas conhecidas. Estas propriedades, principalmente

hidrofobicidade, distribuição eletrônica, as características de ligação de hidrogênio, tamanho da molécula e flexibilidade de ligação, influenciam o comportamento da molécula em um organismo vivo, incluindo a biodisponibilidade, as propriedades de transporte, a afinidade para proteínas, a reatividade, a toxicidade, a estabilidade metabólica e muitos outros (VISTOLI; PEDRETTI; TESTA, 2008). As moléculas propostas foram desenhadas no *software Molinspiration Online Property Calculation Toolkit* (<http://www.molinspiration.com/>), a fim de avaliar os critérios discutidos acima.

4.4 Docking

No intuito de avaliar se as moléculas propostas interagem com o sítio ativo da ROCK1 de maneira semelhante ao ligante cristalizado junto à enzima de código 3TV7, aplicou-se o estudo de *docking*.

Esta metodologia de ancoramento molecular visa prever os modos de interação do ligante dentro do alvo molecular, com base na hipótese de que um valor mais baixo de energia representa uma melhor interação entre a proteína e seu ligante, contribuindo, assim, para elucidar os principais elementos que favorecem essas interações (GUEDES; MAGALHÃES; DARDENNE, 2014).

O *docking* foi realizado utilizando-se o software Molegro Virtual Docker (MVD) (THOMSEN; CHRISTENSEN, 2006). Previamente, as estruturas propostas foram otimizadas por mecânica molecular e submetidas ao cálculo de energia parcial utilizando-se o método semiempírico RM1. Apenas os ligantes foram considerados flexíveis durante a simulação. Foram realizados 50 ancoramentos para cada ligante, utilizando-se o algoritmo de evolução diferencial guiada, em que cada ancoramento gera uma representação (pose). Os parâmetros utilizados foram o do *default*, exceto para a população, que foi

aumentada de 50 para 100. A função de pontuação selecionada foi a MolDock *score*, com resolução de grade de 0,30 Å.

MolDock é uma função de pontuação empírica derivada da função de pontuação potencial linear por partes, em inglês *piecewise linear potential* (PLP), proposta por Gehlhaar et al. (1995) e depois estendida por Yang e Chen (2004). O MolDock incrementa ainda mais essas funções considerando um termo para ligações de hidrogênio e um para cargas, denominado de E_{PLP} . A equação geral do MolDock score pode ser vista na Equação 8.

$$E_{score} = E_{inter} + E_{intra} \quad (8)$$

em que E_{inter} representa a energia intermolecular expressa por

$$E_{inter} = \sum_{ligante} \sum_{enzima} E_{plp} r_{ij} + 332,0 \frac{q_i q_j}{4r_{ij}^2} \quad (9)$$

O termo E_{intra} representa a energia interna do ligante.

$$E_{intra} = \sum_{ligante} \sum_{enzima} E_{plp} r_{ij} + \sum_{ligflexiveis} A (1 - \cos m \cdot \theta) - E_{clash} \quad (10)$$

em que θ é o ângulo torsional da ligação; o termo E_{clash} representa uma penalização de 1.000, se a distância entre dois átomos pesados (separados por mais de duas ligações) for menor que 2,0 Å. Se um átomo pesado do ligante é posicionado fora da região do sítio ativo, uma penalidade de 10.000 é aplicada à energia total.

As melhores poses foram selecionadas a partir da melhor pontuação *docking score* que busca as conformações que mais se assemelham à conformação do ligante cristalizado no sítio ativo.

5 RESULTADOS E DISCUSSÃO

5.1 Alinhamentos para grupos teste I, II e III

A partir do conjunto de equações geradas pelo programa 4D-QSAR, foram selecionadas as melhores equações de QSAR para cada alinhamento. Estas equações foram avaliadas pelos seus parâmetros estatísticos para selecionar a melhor equação para cada banco de dados. Os alinhamentos realizados com o grupo teste I (moléculas **1, 22, 30, 31, 35, 38, 39, 41 e 43**) apresentaram bons valores de validação (Tabela 3).

Tabela 8 Resumo dos resultados dos melhores modelos de QSAR-4D, obtidos para os dez alinhamentos com o grupo teste I e seus respectivos cortes

Alinhamento	GCOD	r^2	RMSE _C	q^2_{cv}	RMSE _{Cv}	q^2_{adj}	r^2_{pred}	$R^2_{m\ test}$	$R^2_{Y\ rand}$	R^2_p
Ali1_0.01_1	8	0,848	0,172	0,785	0,378	0,735	0,535	0,514	0,69	0,337
Ali1_0.1_1	8	0,832	0,173	0,727	0,411	0,664	0,482	0,518	0,56	0,433
Ali2_0.01_1	8	0,87	0,141	0,77	0,380	0,717	-0,160	0,013	0,83	0,174
Ali2_0.1_1	8	0,858	0,146	0,754	0,391	0,697	-0,023	0,082	0,81	0,187
Ali3_0.01_1	8	0,799	0,207	0,527	0,536	0,418	-0,978	-0,00	0,64	0,318
Ali3_0.1_1	8	0,794	0,233	0,593	0,496	0,499	-0,002	0,096	0,55	0,392
Ali4_0.01_1	8	0,856	0,155	0,721	0,415	0,657	0,182	0,248	0,26	0,660
Ali4_0.1_1	8	0,785	0,221	0,61	0,487	0,520	-0,117	0,017	0,63	0,309
Ali5_0.01_1	8	0,812	0,224	0,735	0,407	0,674	0,452	0,402	0,77	0,166
Ali5_0.1_1	8	0,751	0,234	0,64	0,475	0,557	-0,147	0,073	0,67	0,213
Ali6_0.01_1	8	0,852	0,185	0,63	0,480	0,545	0,273	0,251	0,66	0,373
Ali6_0.1_1	8	0,816	0,199	0,714	0,422	0,648	0,078	0,153	0,57	0,404
Ali7_0.01_1	8	0,807	0,229	0,7	0,434	0,631	-0,824	0,002	0,65	0,319
Ali7_0.1_1	8	0,786	0,243	0,7	0,430	0,631	0,241	0,224	0,67	0,267
Ali8_0.01_1	8	0,831	0,191	0,755	0,392	0,699	0,280	0,282	0,34	0,582
Ali8_0.1_1	7	0,806	0,186	0,698	0,435	0,639	-0,577	0,000	0,17	0,642
Ali9_0.01_1	8	0,866	0,138	0,786	0,365	0,737	-0,653	0,014	0,80	0,222
Ali9_0.1_1	7	0,856	0,138	0,785	0,367	0,743	-0,230	0,087	0,87	-
Ali10_0.01_1	8	0,845	0,184	0,729	0,411	0,667	-0,615	0,008	0,47	0,517
Ali10_0.1_1	8	0,802	0,204	0,687	0,443	0,615	-0,443	0,011	0,57	0,386

Pode-se observar que todos os alinhamentos apresentam um coeficiente de determinação (R^2) maior que 0,7 e um valor de Q^2_{cv} (parâmetro calculado na validação *leave-one-out*) maior que 0,5, que são valores aceitáveis para um bom modelo de QSAR (SANTOS-GARCIA et al., 2015). Analisando-se os valores de $Q^2_{ajustado}$ observa-se que o alinhamento **ali3_0.01_1** não apresenta valor superior a 0,5, que é o valor mínimo aceitável para esse parâmetro, levando, assim, à sua exclusão. Outro parâmetro utilizado na validação dos modelos foi o R^2_{pred} e, de acordo com os resultados obtidos nesta validação, somente o alinhamento **ali1_0.01_1** apresenta valor $\geq 0,5$ e, portanto, os outros alinhamentos não foram utilizados.

Os mesmos parâmetros foram avaliados (Tabela 4) para os alinhamentos realizados com o grupo teste II, que abrange as moléculas **3, 5, 6, 14 17, 18, 28, 44 e 47**.

Tabela 9 Resumo dos resultados dos melhores modelos de QSAR-4D obtidos para os dez alinhamentos com o grupo teste II e seus respectivos cortes

Alinhamento	GCOD	r^2	RMSE _C	q^2_{CV}	RMSE _{CV}	q^2_{adj}	r^2_{pred}	$R^2_{m\ test}$	$R^2_{Y\text{-rand}}$	R^2_p
Ali1_0.01_2	8	0,831	0,169	0,765	0,370	0,711	0,666	0,438	0,76	0,221
Ali1_0.1_2	8	0,778	0,202	0,692	0,424	0,621	0,754	0,525	0,66	0,267
Ali2_0.01_2	8	0,839	0,147	0,764	0,371	0,710	0,455	0,354	0,47	0,509
Ali2_0.1_2	8	0,836	0,181	0,753	0,378	0,696	0,145	0,124	0,46	0,512
Ali3_0.01_2	7	0,761	0,213	0,674	0,435	0,610	-0,356	-0,257	0,89	-
Ali3_0.1_2	8	0,717	0,258	0,578	0,496	0,481	0,583	0,444	0,17	0,530
Ali4_0.01_2	8	0,79	0,21	0,599	0,474	0,507	-0,816	-0,775	0,74	0,176
Ali4_0.1_2	8	0,721	0,243	0,558	0,503	0,456	0,421	0,342	0,66	0,178
Ali5_0.01_2	8	0,789	0,193	0,58	0,491	0,484	-2,738	1,578	0,73	0,191
Ali5_0.1_2	7	0,685	0,272	0,55	0,511	0,462	-6,405	15,820	0,14	0,505
Ali6_0.01_2	8	0,821	0,188	0,75	0,382	0,692	0,038	0,035	0,80	0,118
Ali6_0.1_2	8	0,743	0,27	0,557	0,499	0,455	0,576	0,330	0,64	0,238
Ali7_0.01_2	8	0,816	0,193	0,677	0,434	0,603	-0,920	-0,001	0,75	0,209
Ali7_0.1_2	8	0,747	0,279	0,606	0,476	0,515	0,233	0,184	0,64	0,244
Ali8_0.01_2	7	0,753	0,212	0,628	0,459	0,555	0,310	0,296	0,08	0,617
Ali8_0.1_2	8	0,711	0,264	0,56	0,505	0,459	0,566	0,603	0,10	0,555
Ali9_0.01_2	8	0,824	0,168	0,683	0,428	0,610	0,055	0,211	0,48	0,483
Ali9_0.1_2	8	0,854	0,133	0,779	0,358	0,728	0,741	0,662	0,44	0,549
Ali10_0.01_2	8	0,821	0,197	0,732	0,394	0,670	-1,464	-0,000	0,96	-
Ali10_0.1_2	7	0,758	0,244	0,644	0,453	0,574	-0,157	0,157	0,95	-

Como todos os alinhamentos apresentam um coeficiente R^2 maior que 0,7 e Q^2_{cv} maior que 0,5, passou-se a analisar o parâmetro $Q^2_{ajustado}$, o que permitiu excluir os alinhamentos **ali3_0.1_2**; **ali4_0.1_2**; **ali5_0.01_2**; **ali5_0.1_2**; **ali6_0.1_2** e **ali8_0.1_2**, por apresentarem valores inferiores a 0,5. Observando-se os valores de R^2_{pred} foi possível excluir os alinhamentos **ali2_0.01_2**; **ali2_0.1_2**; **ali3_0.01_2**; **ali4_0.01_2**; **ali6_0.01_2**; **ali7_0.01_2**; **ali7_0.1_2**; **ali8_0.01_2**; **ali9_0.01_2**; **ali10_0.01_2**; **ali10_0.1_2**. O valor de R^2_{pred} é dependente da média do grupo de treinamento. Para evitar superestimação de valores, fizeram-se os cálculos de $r_m^2_{teste}$, que é uma validação adicional para o grupo teste. Nesta validação, somente os alinhamentos **ali1_0.1_2** e **ali9_0.1_2** apresentaram resultados satisfatórios ($\geq 0,5$) e, portanto, somente estes dois alinhamentos foram considerados.

Analisando-se os alinhamentos com o grupo teste III (moléculas **9**, **12**, **15**, **21**, **32**, **37**, **48**, **50** e **51**), na Tabela 5, verificou-se que o mesmo apresenta valores superiores a 0,7 e 0,5, para os parâmetros R^2 e Q^2_{cv} , porém, o $Q^2_{ajustado}$ não apresenta valores $\geq 0,5$ para os alinhamentos **ali3_0.01_3**; **ali4_0.1_3**; **ali5_0.1_3** e **ali8_0.1_3**, o que causou a exclusão deles.

Tabela 10 Resumo dos resultados dos melhores modelos de QSAR-4D obtidos para os dez alinhamentos com o grupo teste III e seus respectivos cortes

Alinhamento	GCOD	r^2	RMSE _C	q^2_{CV}	RMSE _{CV}	q^2_{adj}	r^2_{pred}	$R^2_{m\ test}$	$R^2_{Y\text{-rand}}$	R^2_p
Ali1_0.01_3	8	0,856	0,17	0,779	0,368	0,728	0,386	0,356	0,69	0,348
Ali1_0.1_3	8	0,82	0,175	0,722	0,409	0,658	0,532	0,496	0,57	0,41
Ali2_0.01_3	8	0,842	0,161	0,758	0,387	0,702	-1,20	-0,149	0,75	0,255
Ali2_0.1_3	8	0,811	0,192	0,714	0,420	0,648	0,068	0,064	0,79	0,117
Ali3_0.01_3	8	0,749	0,281	0,576	0,509	0,479	0,387	0,335	0,69	0,181
Ali3_0.1_3	8	0,779	0,26	0,615	0,487	0,527	-0,95	0,001	0,52	0,396
Ali4_0.01_3	8	0,854	0,172	0,741	0,401	0,681	0,235	0,240	0,68	0,356
Ali4_0.1_3	8	0,703	0,276	0,567	0,517	0,468	-0,09	0,044	0,67	0,127
Ali5_0.01_3	7	0,773	0,234	0,672	0,447	0,608	0,503	0,520	0,19	0,590
Ali5_0.1_3	8	0,77	0,234	0,517	0,534	0,406	-1,84	-0,089	0,20	0,581
Ali6_0.01_3	7	0,824	0,167	0,743	0,397	0,693	0,019	0,051	0,66	0,333
Ali6_0.1_3	7	0,77	0,257	0,675	0,449	0,611	-1,63	0,504	0,67	0,243
Ali7_0.01_3	8	0,816	0,217	0,715	0,421	0,649	0,411	0,363	0,57	0,404
Ali7_0.1_3	8	0,747	0,283	0,601	0,495	0,509	0,051	0,171	0,67	0,207
Ali8_0.01_3	8	0,792	0,233	0,648	0,460	0,567	0,189	0,249	0,29	0,561
Ali8_0.1_3	8	0,729	0,252	0,548	0,522	0,444	0,505	0,491	0,16	0,549
Ali9_0.01_3	8	0,864	0,139	0,783	0,364	0,733	-1,92	-1,366	0,91	-
Ali9_0.1_3	8	0,85	0,168	0,8	0,353	0,754	0,366	0,326	0,89	-
Ali10_0.01_3	8	0,824	0,207	0,726	0,411	0,663	-3,69	0,047	0,85	-
Ali10_0.1_3	8	0,802	0,204	0,687	0,443	0,615	0,832	0,742	0,54	0,410

Os valores do parâmetro R^2_{pred} mostram que apenas os alinhamentos **ali1_0.1_3**; **ali5_0.01_3**; **ali10_0.1_3** mantiveram-se $\geq 0,5$, que é o mínimo aceitável para esta validação. Quando se observam os valores de $r_m^2_{\text{teste}}$, vê-se que apenas os alinhamentos **ali5_0.01_3** e **ali10_0.1_3** permaneceram no limite desejável.

Após as análises preliminares das validações para todos os grupos, apenas os alinhamentos **ali1_0.01_1**; **ali1_0.1_2**; **ali5_0.01_3**; **ali9_0.1_2** e **ali10_0.1_3** (Tabela 6) permaneceram para a construção do modelo.

Tabela 11 Resumo dos resultados dos melhores alinhamentos de RD-4D-QSAR obtidos após validação

Alinhamento	GCOD	R²	RMSE_C	Q²_{CV}	RMSE_{CV}	Q²_{ajustado}	R²_{pred}	r_m²_{teste}	R²_{Y-rand}	R_p²
Ali1_0.01_1	8	0,848	0,172	0,785	0,378	0,7358	0,5350	0,5142	0,69	0,3370
Ali1_0.1_2	8	0,778	0,202	0,692	0,424	0,6216	0,7541	0,5254	0,66	0,2672
Ali5_0.01_3	7	0,773	0,234	0,672	0,447	0,6082	0,5032	0,5201	0,19	0,5902
Ali9_0.1_2	8	0,854	0,133	0,779	0,358	0,7284	0,7411	0,6172	0,44	0,5494
Ali10_0.1_3	8	0,802	0,204	0,687	0,443	0,6154	0,8326	0,7590	0,54	0,4105

Analisando-se os parâmetros de $R^2_{Y\text{-rand}}$, observa-se que todos os alinhamentos foram satisfatórios, pois todos apresentaram valores inferiores aos respectivos R^2 . Já para o parâmetro R^2_p , somente os valores encontrados para os alinhamentos **ali5_0.01_3** e **ali9_0.1_2** estão dentro da faixa ideal $\geq 0,5$. Comparando-se o número de compostos *outliers* entre os alinhamentos, verificou-se que **ali9_0.1_2** apresentou três compostos e **ali5_0.01_3** não apresentou nenhum composto *outlier*. Por este motivo, o último foi escolhido para prosseguir com a análise dos GCODs.

Na Figura 5 observa-se o domínio de aplicabilidade do modelo e nota-se que tanto os compostos do conjunto treinamento quanto os de treino encontram-se dentro do domínio (entre -2,5 e 2,5). O composto 22, apesar de apresentar maior influência, o que poderia indicar uma dependência do modelo em relação a esse composto, apresenta uma variação de distância menor que 0,1 em relação à distribuição espacial do composto 27, o que contribui para demonstrar a fiabilidade do modelo selecionado.

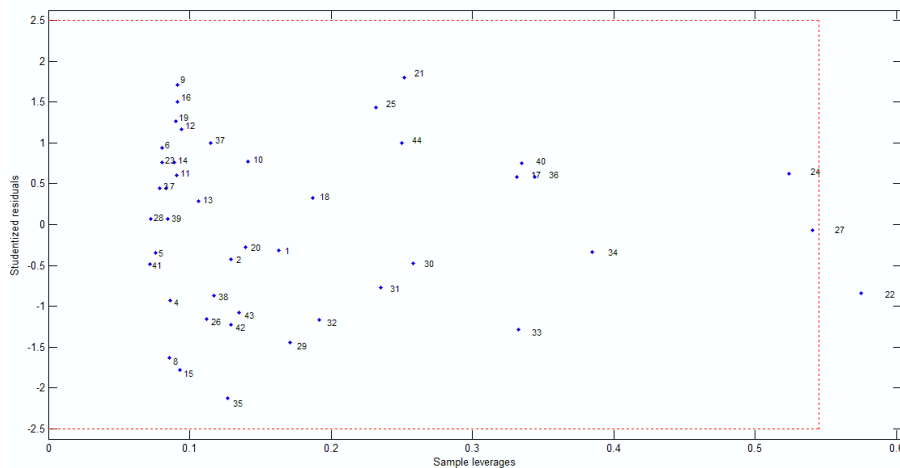


Figura 5 Gráfico de Williams para o modelo Ali5_001_3

5.2 Análise dos descritores

O melhor modelo do alinhamento **ali5_0.01_3** (Modelo 1) gerou 7 descritores (GCODs), dos quais cinco apresentaram coeficientes positivo, **(1)** (3,3,-3,np), **(2)** (1,4,-1,ar), **(3)** (2,3,1,np), **(4)** (0,1,-1,ha), **(5)** (0,-2,1,np), e correspondem a interações favoráveis entre os substituintes dos compostos e os resíduos de aminoácidos do sítio catalítico da enzima. Assim, substituintes com as características (IPEs) selecionadas nessas posições aumentam a potência dos compostos. Os GCODs **(6)** (4,2,-2,ar) e **(7)** (0,4,-3,np) apresentaram coeficientes negativos, correspondendo às interações desfavoráveis entre os substituintes dos compostos e os resíduos de aminoácidos da ROCK (a tabela de frequência de ocupação dos descritores encontra-se no Apêndice 1). Uma representação gráfica desse modelo e seus descritores pode ser vista na Figura 6.

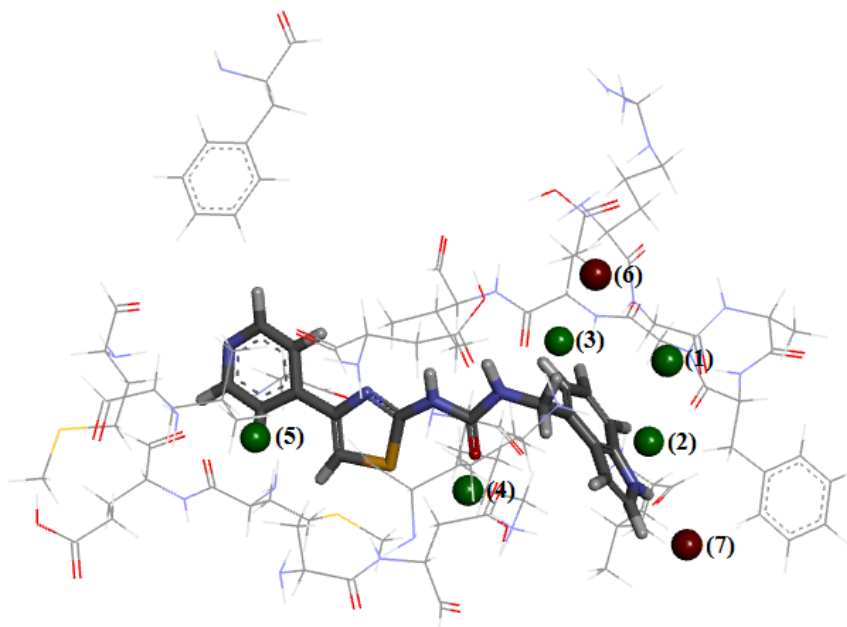


Figura 6 Representação dos resíduos de aminoácidos com a molécula (**29**) menos ativa. As esferas em verde representam os GCODs positivos e as em vermelho, os negativos

A Equação 11 representa o Modelo 1 do alinhamento **ali5_0.01_3**. Os números entre parênteses indicam as coordenadas cartesianas x, y e z da célula da caixa e as letras entre parênteses indicam o tipo de IPE dos átomos que ocupam a célula.

$$pIC_{50} = 4,43 - 0,64(4,2,-2,ar) + 1,24(2,3,1,np) + 0,90(0,1,-1,ha) + \quad (11)$$

$$1,38(1,4,-1,ar) + 4,31(3,3,-3,np) - 0,82(0,4,-3,np) + 0,59(0,-2,1,np)$$

A fim de compreender melhor o comportamento dos dados ajustados para o modelo, a matriz de correlação cruzada entre os diferentes GCODs no modelo 1 foi calculada. Não existe correlação entre pares de GCODs, pois apresentam $r < 0,5$, indicando pouca tendência de um descritor estar influenciando o outro, ou seja, cada descritor contribui de maneira única na construção do modelo de RD-4D-QSAR (KUBINYI, 2008).

O GCOD(1) (3,3,-3,np) está relacionado a substituintes apolares e é o descritor que mais contribui para o aumento na potência dos compostos. Está localizado próximo ao resíduo de aminoácido apolar Gly85 e o hidrogênio do anel benzênico, o que pode favorecer interações de van der Waals entre este aminoácido e o inibidor. Este GCOD apresentou maior frequência de ocupação para o composto **30** (Figura 7).

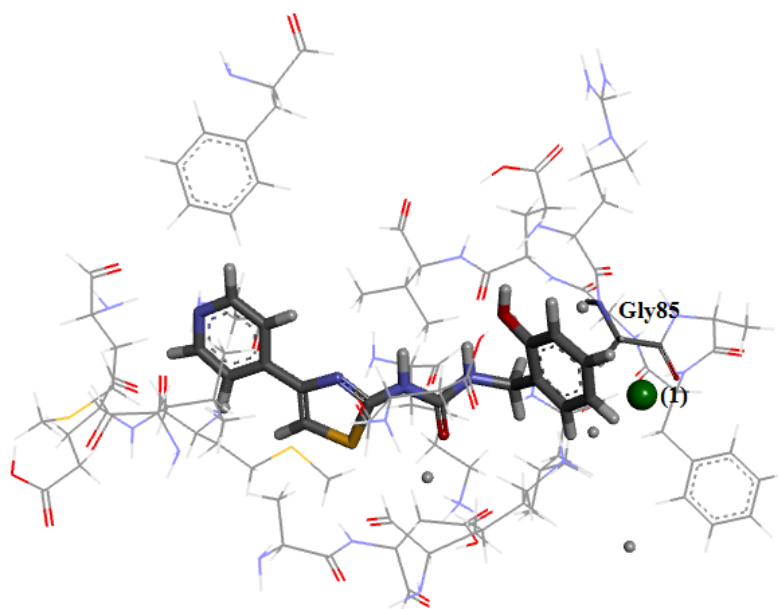


Figura 7 Representação dos resíduos de aminoácidos. Em destaque, Gly85 com a molécula 30. A esfera em destaque representa o GCOD(1)

O GCOD(2) (1,4,-1,ar) tem o segundo coeficiente mais positivo e mostra que o IPE do tipo aromático influencia bastante o aumento da atividade dos compostos. Este descritor apresentou alta frequência de ocupação para o composto **33** (Figura 8) e localiza-se próximo ao anel tolueno da molécula que se encontra próximo ao resíduo de aminoácido aromático Phe87, sugerindo uma interação do tipo π - π .

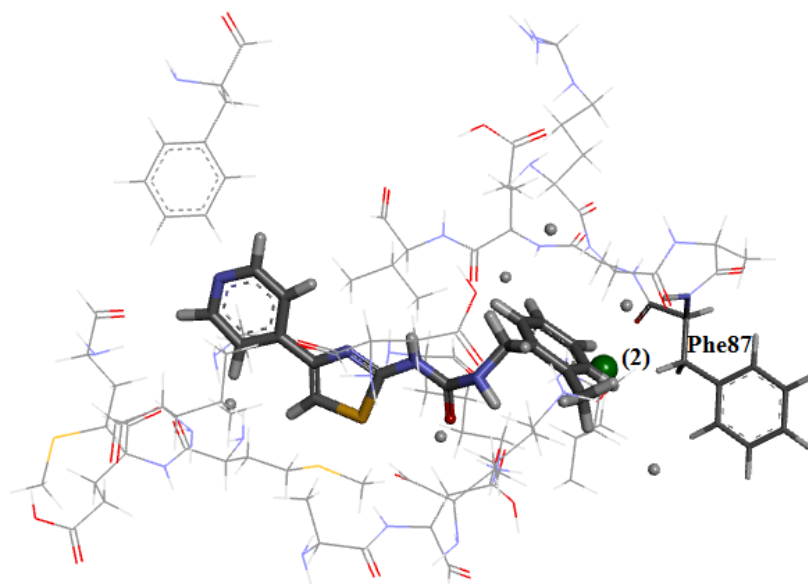


Figura 8 Representação dos resíduos de aminoácidos. Em destaque, Phe87 com a molécula **33**. A esfera em destaque representa o GCOD(2)

Para o GCOD(3) (2,3,1,np) (Figura 9), que é do tipo apolar com coeficiente positivo, os compostos **19** e **44** apresentaram maior frequência de ocupação e está situado entre o átomo de hidrogênio do anel benzênico e o resíduo de aminoácido Val90, favorecendo interações do tipo van der Waals.

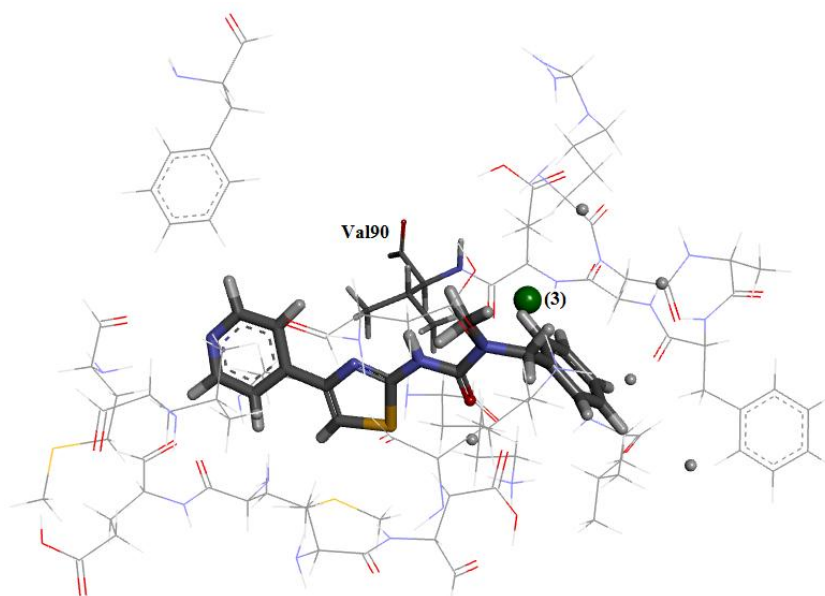


Figura 9 Representação dos resíduos de aminoácidos em destaque Val90, com a molécula 19. A esfera em destaque representa o GCOD(3)

O GCOD (4) (0,1,-1,ha) apresenta um IPE do tipo acceptor de ligação de hidrogênio com maior frequência de ocupação para o composto **2** (Figura 10), o qual se localiza próximo ao grupo carbonil (*urea linkage*) que atua como acceptor. Já o hidrogênio ligado ao nitrogênio do resíduo de aminoácido Lys105 atua como doador, favorecendo, assim, a ligação de hidrogênio, conforme pode ser visto no estudo publicado por Pireddu et al. (2012) (Figura 11), sugerindo a importância dessa ligação.

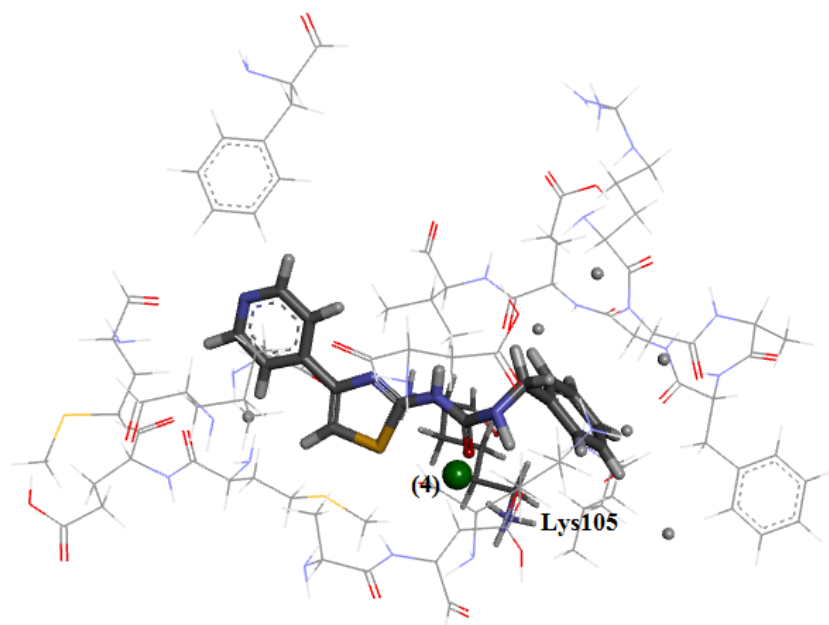


Figura 10 Representação dos resíduos de aminoácidos. Em destaque, Lys105 com a molécula 2. A esfera em destaque representa o GCOD(4)

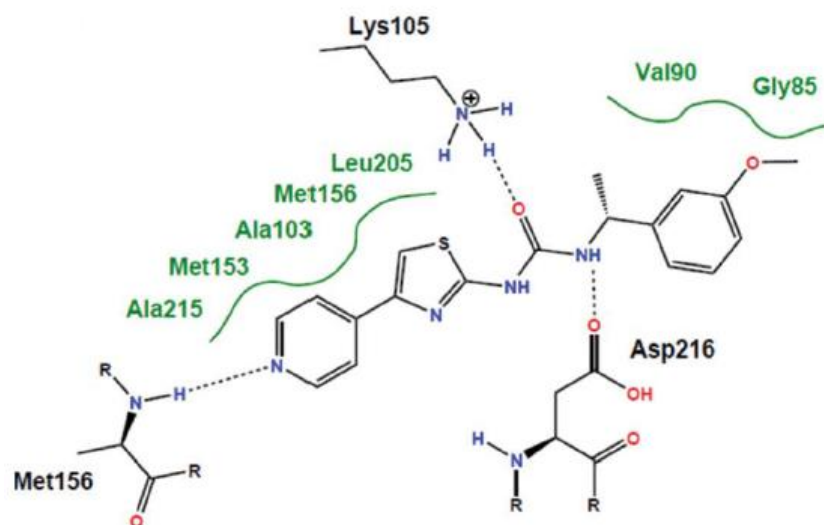


Figura 11 Interações entre os resíduos de aminoácidos Met 156, Lys105 e Asp 216 com um ligante da série congênere. As interações pontilhadas representam as ligações de hidrogênio. Adaptada de Pireddu et al. (2012)

O GCOD (5) (0,-2,1,np) é do tipo apolar com coeficiente positivo e apresenta frequência de ocupação para todos os compostos (Figura 12). Está localizado entre o hidrogênio do anel piridina (grupo 4-(4-piridil)-2-tiazol) e o resíduo de aminoácido Leu205. É importante ressaltar que os outros resíduos de aminoácidos da região (Met156, Ala215) também são apolares, o que favorece interações do tipo van der Waals. Além disso, nessa região é possível observar interações de ligação de hidrogênio entre o NH do resíduo de aminoácido Met156 e o átomo de nitrogênio do anel piridina, conforme descrito por Pireddu et al. (2012).

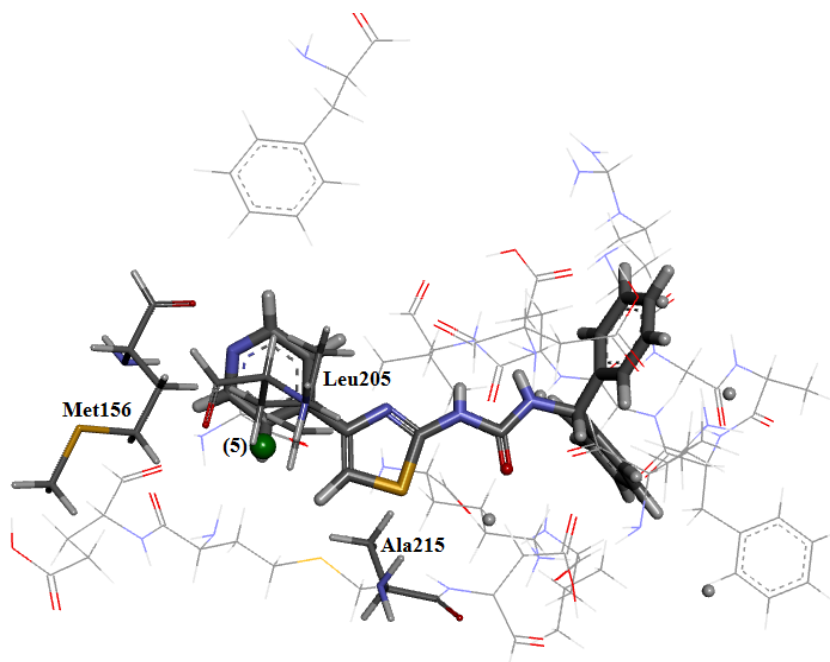


Figura 12 Representação dos resíduos de aminoácidos. Em destaque, Met156, Leu205, Ala215 com a molécula 25. A esfera em destaque representa o GCOD(5)

O GCOD(6) (4,2,-2,ar) é aromático e tem um coeficiente negativo (Figura 13), indicando que a ocupação desta célula da caixa por átomos aromáticos resulta em um decréscimo na potência dos compostos. Este descritor está situado entre o anel benzênico e o resíduo de aminoácido polar básico Arg84, o resíduo Lys200 e Asp 202. O composto, por apresentar um sistema π rico em elétrons, age como um “aceptor” da interação cátion- π do resíduo de Arg84 (GUIMARÃES, 2012). Os compostos com maior frequência de ocupação para esse descritor foram **17**, **25** e **27**. Este descritor pode explicar a diferença de atividade dos compostos **17-18** e **26-27** dependente do centro quiral. Quando o centro é *S*, a posição dos substituintes com ocupação do GCOD (4,2,-2,ar) gera um impedimento espacial. Essa limitação também ocorre quando o anel fenila do composto **25** ocupa esta posição.

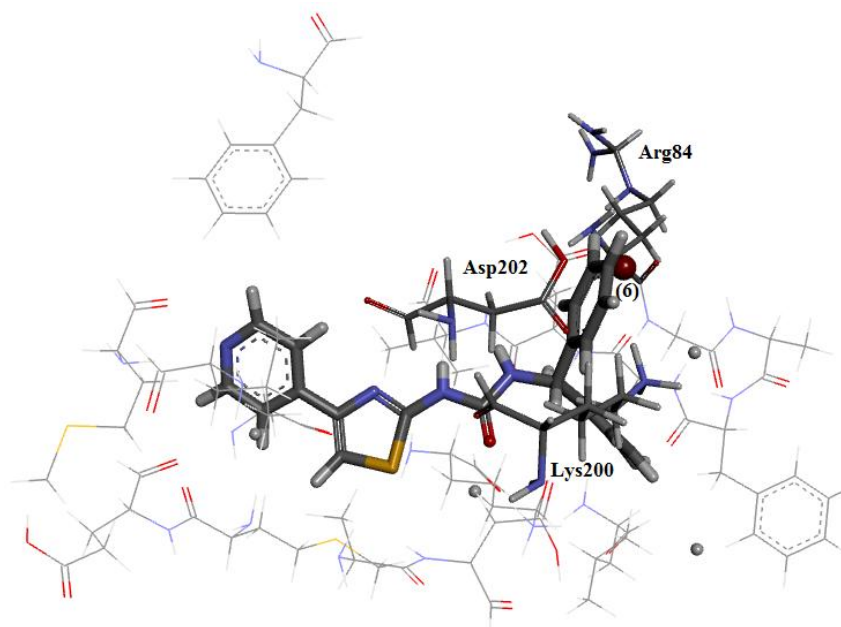


Figura 13 Representação dos resíduos de aminoácidos. Em destaque, Arg84 com a molécula 25. A esfera em destaque representa o GCOD(6)

O GCOD (7) (0,4,-3,np) apresenta o coeficiente mais negativo, provocando grande influência na atividade dos compostos e é do tipo apolar. Este descritor apresentou alta frequência de ocupação para os compostos **28**, **34**, **35** e **36** (Figura 14) e localiza-se próximo ao substituinte apolar ligado ao anel benzênico, indicando que substituintes apolares ligados nessa região não são favoráveis para a atividade, pois dificultam a interação descrita para o GCOD (2). Esta região situa-se na abertura de entrada do local do sítio ativo da ROCK1, perto da parte onde haverá solvente, indicando que substituintes polares devem ser explorados.

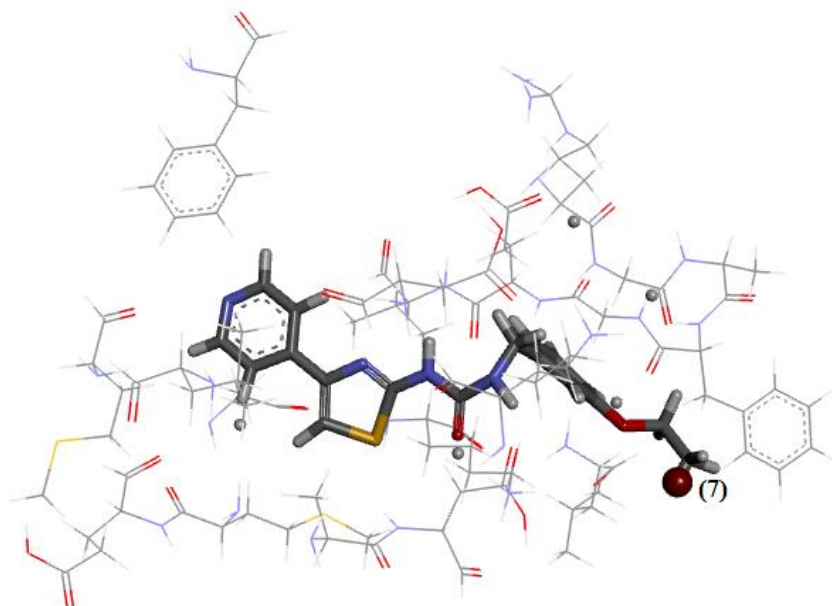


Figura 14 Representação dos resíduos de aminoácidos com a molécula **35**. A esfera em destaque representa o GCOD(7)

Analisando-se os resultados obtidos a partir do estudo de QSAR-4D foi possível propor modificações estruturais visando ao aumento da atividade desses compostos. As bioatividades das estruturas propostas foram realizadas por meio de um ensaio virtual de atividade na qual se aplicou a Equação 11 do Modelo 1 referente ao alinhamento ali5_001_3 e obtiveram-se os valores de atividade predita (Tabela 7). As estruturas foram construídas tomando como base a molécula **18**, que é a mais ativa da série.

Tabela 12 Moléculas propostas e seus respectivos valores de pIC_{50} predito

N°	Estrutura	pIC_{50}	N°	Estrutura	pIC_{50}
18		7,523	P13		5,985
P1		7,127	P14		7,173
P2		7,234	P15		7,121

Tabela 13, continuação

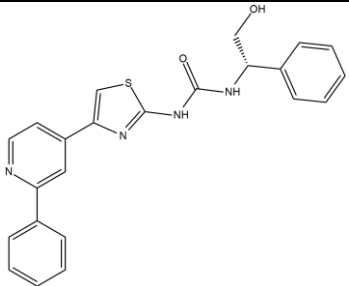
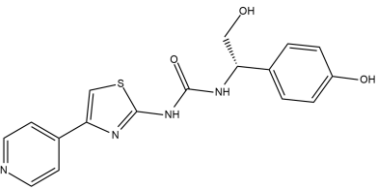
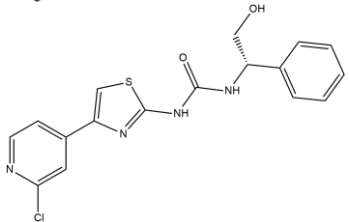
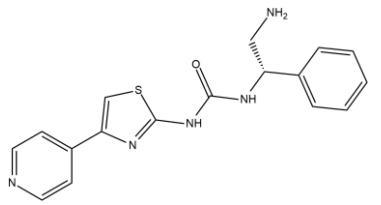
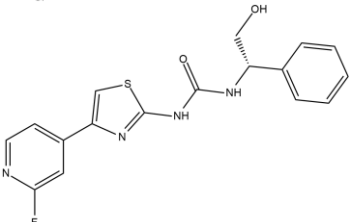
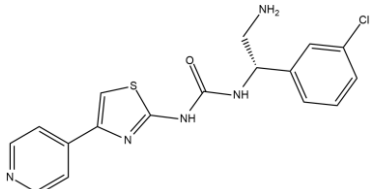
N°	Estrutura	pIC ₅₀	N°	Estrutura	pIC ₅₀
P3		7,518	P16		5,557
P4		6,967	P17		7,123
P5		6,474	P18		7,069

Tabela 14, continuação

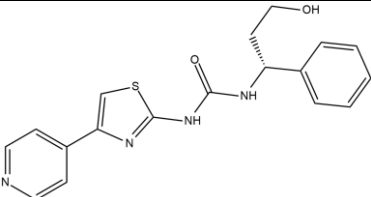
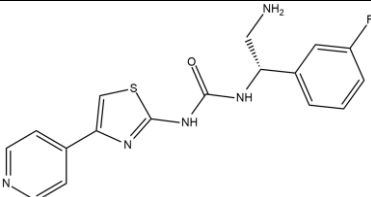
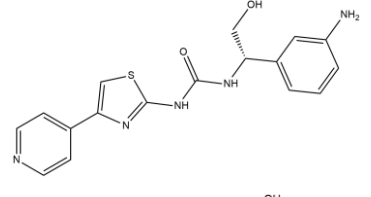
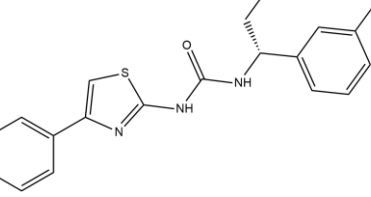
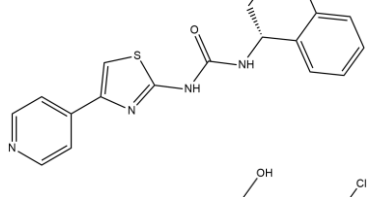
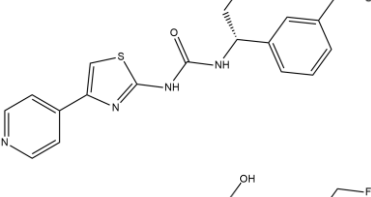
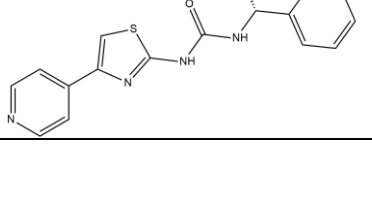
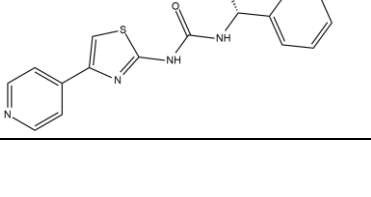
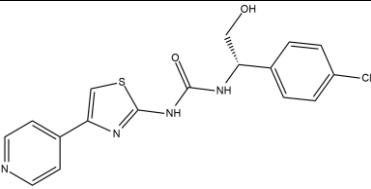
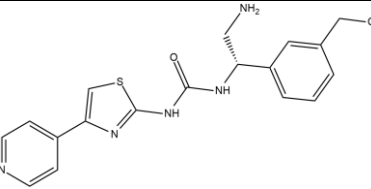
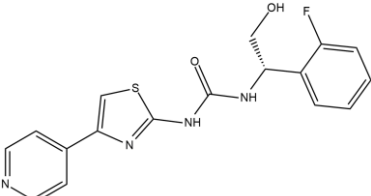
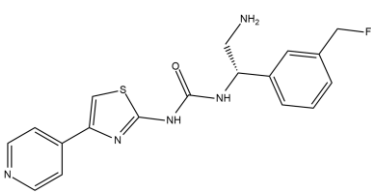
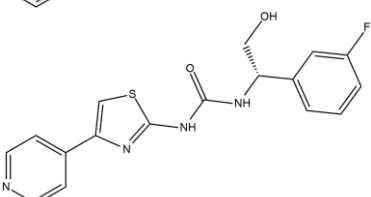
N°	Estrutura	pIC ₅₀	N°	Estrutura	pIC ₅₀
P6		6,813	P19		6,526
P7		7,403	P20		7,783
P8		7,001	P21		8,402
P9		7,091	P22		9,509

Tabela 15, conclusão

N°	Estrutura	pIC ₅₀	N°	Estrutura	pIC ₅₀
P10		5,555	P23		6,521
P11		6,977	P24		8,379
P12		7,123			

O GCOD (6) mostra a preferência pelo centro quiral R. Por esse motivo, manteve-se essa posição para todas as moléculas propostas.

Nas moléculas **P1**, **P2**, **P3**, **P4** e **P5** foram testadas diferentes substituintes no anel piridina onde se localiza o GCOD (5), que é do tipo apolar positivo. Comparando-se as moléculas **P1** e **P2** com as moléculas **P4** e **P5** é possível observar um decaimento do valor de atividade predita quando os substituintes são polares.

Observando-se o efeito da substituição em diferentes posições do anel benzênico, fica evidente a preferência pela posição meta (**P7**, **P9**, **P12** e **P15**), quando se compara a atividade com aquelas moléculas em orto (**P8**, **P11**) e para (**P10**, **P13** e **P16**).

As informações do GCOD (1), GCOD (3) e GCOD (7) mostram a preferência de átomos apolares ligados a átomos polares. Com essa informação, as moléculas **P21**, **P22**, **P23** e **P24** foram propostas.

As estruturas propostas **P20**, **P21**, **P22** e **P24** apresentaram uma atividade biológica predita maior que a atividade experimental do composto **18**, o qual é o composto mais ativo da série estudada. As moléculas **P3** e **P7**, apesar de apresentarem alto valor de atividade predita, não superaram o valor do composto **18**. Todavia, por apresentarem valores próximos aos do composto mais ativo, foram mantidas para o estudo da Regra de Lipinski (LIPINSKI, 2004).

5.3 Regra de Lipinski

O objetivo da Regra de Lipinski é estimar a solubilidade e a permeabilidade de fármacos administrados pela via oral, predizendo a influência da estrutura química na absorção de um composto. Segundo Lipinski, os critérios a serem analisados são: o log P, cujo valor limite é 5; área de superfície

polar (TPSA) menor ou igual a 140 Å; a massa molar (MW), a qual não deve exceder 500 g/mol e os grupos aceptores (n_{ON}) e doadores (n_{OHNH}) de ligação hidrogênio não devem ultrapassar a 10 e 5, respectivamente. Os resultados podem ser vistos na Tabela 8.

Tabela 16 Propriedades moleculares

Molécula	miLogP	TPSA	MW	n_{ON}	n_{OHNH}
18	1,83	87,14	340,41	6	3
P1	2,77	87,14	354,44	6	3
P2	2,85	87,14	368,46	6	3
P3	3,72	87,14	416,51	6	3
P4	2,85	87,14	374,85	6	3
P5	2,34	87,14	358,40	6	3
P6	1,20	87,14	354,44	6	3
P7	0,89	113,16	355,42	7	5
P8	2,46	87,14	374,85	6	3
P9	2,49	87,14	374,85	6	3
P10	2,51	87,14	374,85	6	3
P11	1,95	87,14	358,40	6	3
P12	1,97	87,14	358,40	6	3
P13	2,00	87,14	358,40	6	3
P14	1,77	107,36	356,41	7	4
P15	1,33	107,36	356,41	7	4
P16	1,35	107,36	356,41	7	4
P17	1,27	92,93	339,42	6	4
P18	1,92	92,93	373,87	6	4
P19	1,41	92,93	357,41	6	4
P20	0,77	113,16	355,42	7	5
P21	2,38	87,14	388,88	6	3
P22	2,08	87,14	372,43	6	3
P23	1,82	92,93	387,90	6	4
P24	1,51	92,93	387,90	6	4

Todas as moléculas propostas estão dentro das condições estabelecidas por Lipinski, não cometeram nenhuma violação, indicando não terem problemas

com a absorção, o que as torna possíveis candidatas à síntese e à avaliação farmacológica.

5.4 Docking molecular

O *docking* é uma técnica importante para a concepção e a otimização de moléculas. Esta metodologia foi empregada no intuito de avaliar as interações das moléculas propostas com a enzima ROCK. A estrutura **18** do conjunto de treinamento foi ancorada juntamente com as propostas, a fim de se realizar a comparação das interações. Esta molécula foi selecionada por ser a com maior resposta biológica. Na Tabela 9 apresentam-se as energias obtidas no *docking*.

Tabela 17 Energias de *docking* dos compostos analisados

Molécula	pIC ₅₀	MolDock Score	E _{interação}	H _{Bond}
18	7,523	-112,79	-129,02	-6,54
P1	7,127	-118,82	-136,69	-3,28
P2	7,234	-100,47	-115,60	-2,07
P3	7,518	-108,79	-129,90	-0,72
P4	6,967	-119,11	-137,75	-5,86
P5	6,474	-127,32	-132,74	-3,37
P6	6,813	-113,65	-133,10	-6,42
P7	7,403	-111,49	-129,10	-3,80
P8	7,001	-118,68	-133,96	-3,39
P9	7,091	-113,94	-133,77	-1,43
P10	5,555	-115,68	-129,62	-3,00
P11	6,977	-116,80	-136,55	-6,30
P12	7,123	-116,87	-134,11	-2,94
P13	5,985	-117,11	-133,22	-2,32
P14	7,173	-116,05	-132,80	-3,84
P15	7,121	-113,22	-134,01	-5,43
P16	5,557	-114,31	-131,81	-2,91
P17	7,123	-109,11	-128,15	-0,89
P18	7,069	-117,53	-133,84	-2,79
P19	6,526	-119,18	-136,02	-2,78

Tabela 18, conclusão

Molécula	pIC₅₀	MolDock Score	E_{interação}	H_{Bond}
P20	7,783	-113,96	-131,29	-3,19
P21	8,402	-121,89	-139,18	-2,90
P22	9,509	-117,33	-134,12	-6,70
P23	6,521	-123,49	-139,77	-2,65
P24	8,379	-123,93	-141,27	-3,19

A energia de interação das propostas **P24, P23, P21, P4, P1, P11, P19, P22, P12, P15, P8, P18, P9, P13, P6, P14, P5, P16, P20, P3** e **P7**, em ordem de menor energia, é melhor que a encontrada para a molécula **18** que, por sua vez, apresenta um valor próximo a **P17**. Apenas a proposta **P2** apresenta valor de interação superior.

Estes valores de energia corroboram o modelo obtido do estudo de QSAR-4D, visto que as estruturas propostas mais ativas, **P20, P21, P22** e **P24**, apresentaram melhor valor de energia, e aquelas de atividade similar, **P3** e **P7**, apresentaram energia de interação similar.

6 CONCLUSÃO

Neste estudo, uma série de inibidores piridiltiazólicos para ROCK1 foi avaliada por RD-4D-QSAR. Os melhores modelos foram obtidos a partir da avaliação do conjunto de treinamento (42 compostos) e de três grupos testes (9 compostos cada), com dez alinhamentos para cada combinação. O melhor modelo foi obtido do ali5_0.01_3, em que se empregou o tamanho de célula de grade de 2Å e se obtiveram valores de validações $r^2 = 0,773$; $Q^2_{CV} = 0,672$; $Q^2_{ajustado} = 0,608$, os quais confirmam uma robustez do modelo. Ademais, com base na avaliação dos GCODs, foram propostas 24 estruturas, tomando como base o composto mais ativo da série, tendo quatro delas (**P20**, **P21**, **P22** e **P24**) apresentado valores de atividade biológica predita maiores que o composto referência. Estes compostos foram avaliados pela Regra dos Cinco de Lipinski e não cometeram nenhuma violação; as energias de interação apresentadas no docking confirmam uma boa interação das estruturas propostas com a enzima, indicando que são promissoras candidatas à síntese e à avaliação farmacológica.

7 DIFUSÃO DE TECNOLOGIA

Os resultados deste trabalho foram enviados para um periódico de circulação internacional e está disponível no link <http://dx.doi.org/10.1155/2016/9198582>

REFERÊNCIAS

- ANDRADE, C. H. et al. 4D-QSAR: perspectives in drug design. **Molecules**, Basel, v. 15, n. 5, p. 3281-3294, May 2010.
- ASSIS, T. M. de. QSAR models guided by molecular dynamics applied to human glucokinase activators. **Chemical Biology & Drug Design**, Oxford, v. 87, n. 3, p. 455-466, Mar. 2016.
- AUSTIN, C. P. Molecular biology: NIH molecular libraries initiative. **Science**, New York, v. 306, n. 5699, p. 1138-1139, Nov. 2004.
- BARREIRO, E. J.; FRAGA, C. A. M. **Química medicinal as bases moleculares da ação dos fármacos**. 3. ed. Porto Alegre: Artmed, 2015. 489 p.
- BERMAN, H. M. et al. The protein data bank. **Nucleic Acids Research**, Oxford, v. 28, n. 1, p. 235-242, 2000.
- BOLAND, S. et al. Design, synthesis, and biological evaluation of novel, highly active soft ROCK inhibitors. **Journal of Medicinal Chemistry**, Washington, v. 58, n. 10, p. 4309-4324, 2015.
- 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.
- CASEY, P. J. Protein lipidation in cell signaling. **Science**, New York, v. 268, n. 5208, p. 221-225, 1995.
- CHEM21 GROUP INC, THE. **4D-QSAR user's manual**. Version 1.0. Lake Forest, 1997. Software.
- CHERKASOV, A. et al. QSAR modeling: where have you been? where are you going to? **Journal of Medicinal Chemistry**, Washington, v. 57, n. 12, p. 4977-5010, 2014.
- CRAMER, R. D.; PATTERSON, D. E.; BUNCE, J. D. Comparative molecular field analysis (CoMFA): 1., effect of shape on binding of steroids to carrier proteins. **Journal of the American Chemical Society**, Easton, v. 110, n. 18, p. 5959-5967, 1988.

DARENFED, H. et al. Molecular characterization of the effects of Y-27632. **Cell Motility and the Cytoskeleton**, New York, v. 64, n. 2, p. 97-109, 2007.

DOGGRELL, S. A. Rho-kinase inhibitors show promise in pulmonary hypertension. **Expert Opinion on Investigational Drugs**, London, v. 14, n. 9, p. 1157-1159, Sept. 2005.

ETIENNE-MANNEVILLE, S.; HALL, A. Rho GTPases in cell biology. **Nature**, London, v. 420, n. 6916, p. 629-635, 2002.

FERNANDES, L. B.; HENRY, P. J.; GOLDIE, R. G. Rho kinase as a therapeutic target in the treatment of asthma and chronic obstructive pulmonary disease. **Therapeutic Advances in Respiratory Disease**, Los Angeles, v. 1, n. 1, p. 25-33, 2007.

FERREIRA, M. M. C. Multivariate QSAR. **Journal of the Brazilian Chemical Society**, São Paulo, v. 13, n. 6, p. 742-753, 2002.

FRIEDMAN, H. L. Influence of isosteric replacements upon biological activity. In: SYMPOSIUM ON CHEMICAL-BIOLOGICAL CORRELATION, 1., 1951, Washington. **Proceedings...** Washington, 1951. p. 295-358.

GEHLHAAR, D. K. et al. **Docking conformationally flexible small molecules into a protein binding site through evolutionary programming.** [S.l.: s.n.], 1995. 627 p.

GUEDES, I. A.; MAGALHÃES, C. S. de; DARDENNE, L. E. Receptor-ligand molecular docking. **Biophysical Reviews**, New York, v. 6, n. 1, p. 75-87, 2014.

GUIMARÃES, C. R. W. The many contributions to protein-ligand binding: implications in drug design. **Revista Virtual de Química**, Niterói, v. 4, n. 4, p. 348-364, 2012.

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**, Washington, v. 86, n. 8, p. 1616-1626, 1964.

HARMSEN, S.; KOK, R. J. Kinase inhibitor conjugates. **Current Pharmaceutical Design**, San Francisco, v. 18, n. 20, p. 2891-900, 2012.

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

HUENTELMAN, M. J. et al. NIH public access. **Behavioral Neuroscience**, Washington, v. 123, n. 1, p. 218-223, 2010.

IIZUKA, K. et al. Evaluation of Y-27632, a Rho-kinase inhibitor, as a bronchodilator in guinea pigs. **European Journal of Pharmacology**, New York, v. 406, n. 2, p. 273-279, 2000.

ISHIZAKI, T. et al. Pharmacological properties of Y-27632, a specific inhibitor of rho-associated kinases. **Molecular Pharmacology**, New York, v. 57, n. 5, p. 976-983, 2000.

ITAGAKI, M. et al. Possible mechanisms underlying statin-induced skeletal muscle toxicity in L6 fibroblasts and in rats. **Journal of Pharmacological Sciences**, Tokyo, v. 109, n. 1, p. 94-101, 2009.

JIN, B.; HOPFINGER, A. J. A proposed common spatial pharmacophore and the corresponding active conformations of some TxA₂ receptor antagonists. **Journal of Chemical information and Computer Science**, Washington, v. 34, n. 4, p. 1014-1021, 1994.

KAMAI, T. et al. Overexpression of RhoA, Rac1, and Cdc42 GTPases is associated with progression in testicular cancer. **Clinical Cancer Research**, Philadelphia, v. 10, n. 14, p. 4799-4805, 2004.

KLEBE, G.; ABRAHAM, U.; MIETZNER, T. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. **Journal of Medicinal Chemistry**, Washington, v. 37, n. 24, p. 4130-4146, Nov. 1994.

KUBINYI, H. Applications of hansch analysis. In: GMBH, W. V. V. (Org.). **QSAR: hansch analysis and related approaches**. Weinheim: Wiley, 2008. p. 115-137.

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.

- KUBO, T. et al. The therapeutic effects of Rho-ROCK inhibitors on CNS disorders. **Therapeutics and Clinical Risk Management**, Bethesda, v. 4, n. 3, p. 605-615, 2008.
- KUZ'MIN, V. E.; ARTEMENKO, A. G.; MURATOV, E. N. Hierarchical QSAR technology based on the Simplex representation of molecular structure. **Journal of Computer-Aided Molecular Design**, Dordrecht, v. 22, n. 6/7, p. 403-421, 2008.
- LAWRENCE, N. J.; PIREDDU, R.; SEBTI, S. M. **Pyridylthiazole-based ureas as inhibitors of rho associated protein kinase (rock) and methods of use**. US 2013/0059839 A1, 7 mar. 2013.
- LEUNG, T. et al. The p160 RhoA-binding kinase ROK alpha is a member of a kinase family and is involved in the reorganization of the cytoskeleton. **Molecular and Cellular Biology**, The Hague, v. 16, n. 10, p. 5313-5327, 1996.
- LIPINSKI, C. A. Lead- and drug-like compounds: the rule-of-five revolution. **Drug Discovery Today: Technologies**, Kidlington, v. 1, n. 4, p. 337-341, 2004.
- LOGRASSO, P. V.; FENG, Y. Rho kinase (ROCK) inhibitors and their application to inflammatory disorders. **Current Topics in Medicinal Chemistry**, Hilversum, v. 9, n. 8, p. 704-723, 2009.
- LOWIS, D. R. HQSAR: a new, highly predictive QSAR technique. **Tripes Technical Notes**, Saint Louis, v. 1, n. 5, p. 1-17, 1997.
- MADAULE, P.; AXEL, R. A novel ras-related gene family. **Cell**, Cambridge, v. 41, n. 1, p. 31-40, 1985.
- MARTINS, P. A. et al. LQTA-QSAR: a new 4D-QSAR methodology. **Journal of Chemical Information and Modeling**, Washington, v. 49, n. 6, p. 1428-1436, 2009.
- MATSUI, T. et al. Rho-associated kinase, a novel serine/threonine kinase, as a putative target for small GTP binding protein Rho. **The EMBO Journal**, Oxford, v. 15, n. 9, p. 2208-2216, 1996.
- MILANI, P. **Synthesis of L-cystine and L-cysteine derivatives for use in studies of 20S proteasome inhibition**. 2011. 179 p. Thesis (Ph.D. in Chemistry) - Instituto de Química, Universidade de São Paulo, São Paulo, 2011.

MONG, P. Y.; WANG, Q. Activation of Rho kinase isoforms in lung endothelial cells during inflammation. **Journal of Immunology**, Baltimore, v. 182, n. 4, p. 2385-2394, 2009.

NAKAGAWA, O. et al. ROCK-I and ROCK-II, two isoforms of Rho-associated coiled-coil forming protein serine/threonine kinase in mice. **FEBS Letters**, Amsterdam, v. 392, n. 2, p. 189-193, 1996.

NOSSAMAN, B. D. et al. Role of the RhoA/Rho-kinase pathway in the regulation of pulmonary vasoconstrictor function. **Canadian Journal of Physiology and Pharmacology**, Ottawa, v. 88, n. 1, p. 1-8, Jan. 2010.

NUNES, C. A. et al. Chemoface: a novel free user-friendly interface for chemometrics. **Journal of the Brazilian Chemical Society**, São Paulo, v. 23, n. 11, p. 2003-2010, 2012.

NUNES, K. P.; RIGSBY, C. S.; WEBB, R. C. RhoA/Rho-kinase and vascular diseases: what is the link? **Cellular and Molecular Life Sciences**, Basel, v. 67, n. 22, p. 3823-3836, Nov. 2010.

OGATA, S. et al. Fasudil inhibits lysophosphatidic acid-induced invasiveness of human ovarian cancer cells. **International Journal of Gynecological Cancer**, Trevoise, v. 19, n. 9, p. 1473-1480, 2009.

OOMS, F. Molecular modeling and computer aided drug design: examples of their applications in medicinal chemistry. **Current Medicinal Chemistry**, Wageningen, v. 7, n. 2, p. 141-158, 2000.

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

ORGANIZATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. **Guidance document on the validation of (quantitative) Structure-Activity Relationship [(Q)SAR] models**. Paris, 2010. 154 p.

PAN, P. et al. Advances in the development of Rho-associated protein kinase (ROCK) inhibitors. **Drug Discovery Today**, Kidlington, v. 18, n. 23/24, p. 1323-1333, 2013.

PIREDDU, R. et al. Pyridylthiazole-based ureas as inhibitors of Rho associated protein kinases (ROCK1 and 2). **MedChemComm**, Cambridge, v. 3, n. 6, p. 699-709, 2012.

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

RIDLEY, A. J. Rho family proteins: coordinating cell responses. **Trends in Cell Biology**, Cambridge, v. 11, n. 12, p. 471-477, 2001.

RIENTO, K.; RIDLEY, A. J. Rocks: multifunctional kinases in cell behaviour. **Nature Reviews Molecular Cell Biology**, London, v. 4, n. 6, p. 446-456, 2003.

RIKITAKE, Y.; LIAO, J. K. ROCKs as therapeutic targets in cardiovascular diseases. **Expert Review of Cardiovascular Therapy**, London, v. 3, n. 3, p. 441-451, May 2005.

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, p. 854-866, 1994.

ROOKS, B. R. et al. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. **Journal of Computational Chemistry**, Washington, v. 4, n. 2, p. 187-217, 1983.

ROY, P. P. et al. On two novel parameters for validation of predictive QSAR Models. **Molecules**, Basel, v. 15, n. 1, p. 604-605, 2010.

SAMUEL, M. S. et al. Actomyosin-mediated cellular tension drives increased tissue stiffness and B-catenin activation to induce interfollicular epidermal hyperplasia and tumor growth. **Cancer Cell**, Maryland Heights, v. 19, n. 6, p. 776-791, 2011.

SANTOS-GARCIA, L. et al. QSAR analysis of nicotinamidic compounds and design of potential Bruton's Tyrosine Kinase (Btk) inhibitors. **Journal of Biomolecular Structure & Dynamics**, New York, p. 1-50, Aug. 2015.

SAPET, C. et al. Thrombin-induced endothelial microparticle generation : identification of a novel pathway involving ROCK-II activation by caspase-2 Thrombin-induced endothelial microparticle generation: identification of a novel pathway involving ROCK-II activation by c. **Blood**, Washington, v. 108, n. 6, p. 1868-1876, 2014.

SCHAAFSMA, D. et al. Rho kinase inhibitors: a novel therapeutical intervention in asthma? **European Journal of Pharmacology**, New York, v. 585, n. 2/3, p. 398-406, 2008.

SENESE, C. L. et al. 4D-fingerprints, universal QSAR and QSPR descriptors. **Journal of Chemical Information and Computer Sciences**, Washington, v. 44, n. 5, p. 1526-1539, 2004.

SHIMOKAWA, H.; RASHID, M. Development of Rho-kinase inhibitors for cardiovascular medicine. **Trends in Pharmacological Sciences**, Amsterdam, v. 28, n. 6, p. 296-302, 2007.

SIGMA-ALDRICH. Disponível em:
<<http://www.sigmaaldrich.com/catalog/product/sigma/y0503?lang=pt®ion=BR>>. Acesso em: 10 fev. 2016.

SPOEL, D. van der et al. GROMACS: fast, flexible, and free. **Journal of Computational Chemistry**, New York, v. 26, n. 16, p. 1701-1718, Dec. 2005.

TAKI, F. et al. Effects of Rho-kinase inactivation on eosinophilia and hyper-reactivity in murine airways by allergen challenges. **Clinical and Experimental Allergy**, Malden, v. 37, n. 4, p. 599-607, 2007.

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

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

TROPSHA, A. Best practices for QSAR model development, validation, and exploitation. **Molecular Informatics**, Malden, v. 29, n. 6/7, p. 476-488, 2010.

TROPSHA, A.; GOLBRAIKH, A. Predictive QSAR modeling workflow, model applicability domains, and virtual screening. **Current Pharmaceutical Design**, San Francisco, v. 13, n. 34, p. 3494-3504, 2007.

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, 2005.

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

VISTOLI, G.; PEDRETTI, A.; TESTA, B. Assessing drug-likeness: what are we missing? **Drug Discovery Today**, Kidlington, v. 13, n. 7/8, p. 285-294, 2008.

WETTSCHURECK, N.; OFFERMANN, S. Rho/Rho-kinase mediated signaling in physiology and pathophysiology. **Journal of Molecular Medicine**, Amsterdam, v. 80, n. 10, p. 629-638, 2002.

YANG, J. M.; CHEN, C. C. GEMDOCK: a generic evolutionary method for molecular docking. **Proteins: Structure, Function and Bioinformatics**, New York, v. 55, n. 2, p. 288-304, 2004.

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

APÊNDICE

Apêndice I

Frequência de ocupação

Molécula	GCOD1	GCOD2	GCOD3	GCOD4	GCOD5	GCOD6	GCOD7
1	0,002	0	0	0,004	0,812	0	0
2	0	0,972	0	1	0,604	0	0
3	0	0	0,006	0,002	1,632	0	0
4	0	0,604	0	0,928	1,084	0	0
5	0,004	0,156	0,004	0,614	1,230	0	0
6	0	0,412	0,040	0,002	1,750	0	0
7	0	0,004	0,032	0,004	1,732	0	0
8	0	0,444	0,004	0,002	1,724	0	0
9	0	0,166	0	0,996	1,008	0	0
10	0,014	0	0	0,010	1,762	0	0
11	0	0,042	0	0,824	0,644	0	0
12	0	0,772	0	0,912	1,254	0	0,382
13	0	0,664	0	0,998	0,644	0	0
14	0	0,032	0	0	1,798	0	0
15	0,082	0,004	0	0,010	1,902	0,924	0
16	0,002	0	0	0,008	1,866	0,052	0
17	0,166	0	0	0,238	1,488	0,428	0
18	0	0,798	0	0,968	0,918	0	0
19	0	0	1	0	1,786	0	0
20	0	0,800	0,002	0,004	1,494	0	0,662
21	0	0,416	0,326	0,966	0,594	0	0
22	0	0,750	0	0,998	0,762	0	0
23	0	0,710	0	0,004	1,766	0	0,004
24	0,386	0	0	0,272	1,026	0,688	0
25	0,032	0,828	0,166	0,002	1,890	1,462	0
26	0	0,746	0,026	0,924	0,904	0	0
27	0,068	0	0	0,004	1,136	1,422	0
28	0	0,632	0	1	0,734	0	1,026
29	0	0	0,008	0,002	1,026	0	0,202
30	0,504	0,006	0,002	0,040	0,340	0,706	0,004
31	0	0,582	0,002	0,996	0,730	0	0,328
32	0	0,404	0,002	0,002	1,830	0	0,002
33	0	1,172	0,018	1	0,734	0	0
34	0	0,1	0	0,536	1,234	0	0,936
35	0	0,790	0	1	0,704	0	1,060
36	0	0,750	0	1	0,562	0	0,936
37	0,526	0	0	0,232	1,082	1,064	0
38	0,360	0	0	0,402	1,734	0,484	0
39	0	0,228	0	0,996	0,826	0	0
40	0	1,164	0	0,002	1,068	0	0,616
41	0	0,416	0,004	0,002	1,626	0	0,556
42	0	0,516	0	0,002	1,890	0	0
43	0	0,798	0,050	0,736	0,798	0	0
44	0	0	1	0,006	1,640	0	0
45	0,010	0	0	0	1,500	0	0
46	0,192	0	0	0,002	1,088	0	0
47	0	0	0	0,008	0,922	0	0
48	0	0,148	0	0,006	1,778	0	0
49	0	0	0	0,212	0,386	0	0
50	0,276	0	0	0,010	1,048	0	0
51	0	0	0	0	1,194	0	0