ESTELLA GASPAR DA MOTA

PLANEJAMENTO RACIONAL DE AGENTES ANTI-HIPERTENSIVOS MULTI-ALVO

LAVRAS - MG

2015

ESTELLA GASPAR DA MOTA

PLANEJAMENTO RACIONAL DE AGENTES ANTI-HIPERTENSIVOS MULTI-ALVO

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Agroquímica, área de concentração em Química/Bioquímica, para a obtenção do título de Doutor.

Orientador

Dr. Matheus Puggina de Freitas

LAVRAS - MG

2015

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

Mota, Estella Gaspar da.
Planejamento racional de agentes anti-hipertensivos multi-alvo / Estella Gaspar da Mota. – Lavras: UFLA, 2015.
108 p.

Tese (doutorado) – Universidade Federal de Lavras, 2015.
Orientador: Matheus Puggina de Freitas.
Bibliografia.

1. Hipertensão arterial. 2. Anti-hipertensivos. 3. Renina. I. Universidade Federal de Lavras. II. Título.

ESTELLA GASPAR DA MOTA

PLANEJAMENTO RACIONAL DE AGENTES ANTI-HIPERTENSIVOS MULTI-ALVO

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Agroquímica, área de concentração em Química/Bioquímica, para a obtenção do título de Doutor.

APROVADA em 23 de fevereiro de 2015.

Dr. Teodorico de Castro Ramalho	UFLA
Dra. Elaine Fontes Ferreira da Cunha	UFLA
Dr. Alexandre Carvalho Bertoli	UNIFAL
Dra. Melissa Soares Caetano	UFOP

Dr. Matheus Puggina de Freitas

Orientador

LAVRAS - MG

2015

“Mas é preciso ter força
É preciso ter raça
É preciso ter gana sempre
Quem traz no corpo a marca
Maria, Maria
Mistura a dor e a alegria.”

(Milton Nascimento)

AGRADECIMENTOS

O final de mais uma etapa inspira gratidão... Gratidão pelos amigos, pela família, pelas situações, pelas grandes e pequenas coisas sem as quais nada valeria realmente a pena.

Agradeço, primeiramente, a Deus, por estar ao meu lado, inclusive nos momentos em que não merecia e por desejar, de uma forma ou de outra, que eu aqui estivesse.

Agradeço à UFLA e ao Departamento de Química, pelo acolhimento de longos anos e à CAPES, pela concessão da bolsa de estudos durante todo este trabalho.

Aos meus pais agradeço por toda confiança, todo apoio, todos os tipos de sacrifícios superados para que eu alcançasse essa realização. Agradeço por cada telefonema, cada palavra providencial dispensada, em especial pela minha mãe que, como ninguém, tem o dom de me devolver à vida todos os dias.

Agradeço à minha família, como um todo, por toda a torcida realizada durante esses anos de caminhada. Em especial, deixo registrado o meu carinho pelos meus avós, Joaquim e Fiinha; pelo meu tio Luís e minha madrinha Alaíde; pelo meu irmão, Gonzaga e pelas minhas primas, Nete, Taís e Elisa, que viveram comigo essa experiência, de maneira muito próxima.

A todo o pessoal do laboratório, agradeço, sobretudo, pelo acolhimento e pela forma como fui recebida e inserida no grupo da “Era do Gelo”! Tal atitude facilitou a criação de vínculos e o bom andamento do trabalho. Muito obrigada!

Não posso deixar de citar a família MIA-QSAR (Carol, Daniel e Mari), que me amparou em cada dificuldade com disponibilidade, boa vontade e carinho... Obrigada pela parceria no trabalho que se estendeu de forma tão espontânea à própria vida!

À Karina agradeço pela prestimosa ajuda no desenvolver do trabalho, com análises que vieram a enriquecer nossos resultados.

Agradeço aos meus amigos que foram lançar voos mais distantes por não permitirem que os quilômetros diminuam nosso carinho e cumplicidade... Adoro vocês! Destaco minha amigaaaa querida Pri Destro e meu brother Felipe, especialmente presentes nessa fase de minha vida.

Agradeço à Natalie e a Míriam, pelas injeções de ânimo semanais e pelo respeito e carinho que sempre nutriram por mim.

À professora Elaine agradeço pela parceria na realização do projeto, pela disponibilidade e pela disposição em ajudar sempre atuante! Tenha a certeza de que você enriqueceu muito nosso trabalho com sua valiosa colaboração e conhecimento.

Agradeço também a todos os professores do DQI que, de uma maneira ou de outra, estiveram envolvidos na realização deste trabalho, principalmente ao professor Téo, pela presença e disponibilidade e aos professores Guerreiro e Iara, pelo carinho e amizade.

Aos membros da banca, professores Téo, Elaine, Mel, Kátia e doutores Alexandre e Stephen, pela prontidão com que se colocaram disponíveis ao receberem nosso convite e, principalmente, pelas valiosas sugestões, correções e colaborações que aumentaram a relevância de nosso trabalho.

A Malu, minha ex-orientadora, amiga, confidente, por vezes, quase minha mãe, agradeço pela sensibilidade de notar o momento certo de me “tirar de debaixo das suas saias”, oferecendo-me a possibilidade de seguir um caminho totalmente novo. Todos os frutos que colho hoje e que são motivos de tantos agradecimentos, divido com você, que soube ver o depois com os olhos de antes e que me permitiu voos tão ousados.

Ao meu orientador, Matheus, agradeço primeiramente por ter me aceitado no seu grupo de pesquisa, confiando no meu trabalho e no meu

potencial. Obrigada pela sua disponibilidade, pela sua paciência, por ser esse grande profissional que você é. Trabalhar com pessoas como você é prazeroso e faz com que cresça dentro de mim a fé em um futuro de grandes realizações. Muito obrigada!

RESUMO

A hipertensão arterial, ou pressão alta, é uma doença caracterizada pela elevação dos níveis de tensão no sangue e por ter custos médicos e socioeconômicos elevados, decorrentes, principalmente, das suas complicações, tais como doença cerebrovascular, doença arterial coronariana e insuficiência cardíaca. Os medicamentos empregados para o controle dessa doença exibem maior eficácia quando ministrados em conjunto com outros fármacos, fato que potencializa os efeitos colaterais. Dentre as classes de anti-hipertensivos utilizadas no tratamento, destacam-se as responsáveis por inibir a ação da renina, os inibidores da enzima conversora de angiotensina (ECA) e os bloqueadores de canais de cálcio, entre outros. A proposta, neste trabalho, consistiu em avaliar a capacidade de modelagem do método MIA-QSAR para duas classes de agentes anti-hipertensivos: inibidores da renina e bloqueadores dos canais de cálcio. Os modelos construídos mostraram-se estatisticamente robustos e preditivos, sendo escolhidos para a determinação da atividade biológica dos novos inibidores da renina e bloqueadores de canais de cálcio, concebidos a partir da estratégia de combinar subestruturas químicas das moléculas altamente bioativas de cada série. Dentre os compostos propostos, os potencialmente mais ativos apresentaram valores estimados de pIC_{50} de 9,26 para os inibidores da renina e 12,77 para os bloqueadores de canais de cálcio, superando aqueles presentes na literatura. Tais predições foram validadas por estudos de *docking*. Além disso, os compostos propostos exibiram parâmetros farmacocinéticos calculados comparáveis aos compostos de referência, sendo potencialmente de fácil absorção, distribuição, metabolismo e excreção. Neste trabalho também descrevem-se os estudos de *docking* de compostos concebidos a partir da combinação de subestruturas de três tipos de anti-hipertensivos: inibidores da renina, bloqueadores do canal de cálcio e inibidores da enzima conversora de angiotensina (ECA), objetivando, assim, a obtenção de um composto multialvo. Alguns compostos propostos apresentaram valores animadores de *docking score* e de interação intermolecular ligante-enzima, para todos os três alvos estudados, pois estão acima dos valores calculados para os compostos de referência alisquireno (inibidor da renina), amlodipina (bloqueador de canais de cálcio) e captopril (inibidor da ECA), indicando perspectivas promissoras quanto a estudos posteriores.

Palavras chave: Hipertensão arterial. Anti-hipertensivos. Renina. Bloqueadores de canais de cálcio, Enzima conversora de angiotensina. MIA-QSAR. *Docking.*

ABSTRACT

Hypertension, or high blood pressure, is a disease characterized by elevated blood pressure levels and has high medical and socio-economical costs, arising mainly from its complications, such as cerebrovascular disease, coronary artery disease and heart failure. The drugs used to control this disease exhibit greater efficacy when given in combination with other drugs, a fact that enhances the side effects. Among the antihypertensive classes used for treatment, the highlights are those responsible for inhibiting the renin action, the angiotensin-converting enzyme (ACE) inhibitors, calcium channels blockers, among others. The purpose of this study was to evaluate the modeling capacity of the MIA-QSAR method for two classes of antihypertensive agents: renin inhibitors and calcium channel blockers. The models developed were statistically robust and predictive, chosen for determining the biological activity of new renin inhibitors and calcium channel blockers designed from the strategy of combining chemical substructures of highly bioactive molecules in each series. Among the compounds proposed, the most promising presented pIC_{50} values of 9.26 (renin inhibitors) and 12.77 (calcium channels blockers), exceeding those present in literature. Such predictions were validated by docking studies. Furthermore, the proposed compounds exhibited calculated pharmacokinetic parameters comparable to those of reference compounds, being potentially well absorbed, distributed, metabolized and excreted. This work also describes the docking studies of compounds designed on the basis of the combination of substructures of three types of antihypertensive drugs: renin inhibitors, calcium channel blockers and angiotensin-converting enzyme (ACE) inhibitors, in order to obtain a multi-target compound. Some proposed compounds showed encouraging values of docking score and intermolecular ligand-enzyme interaction for all three targets studied, since these data were higher than the values calculated for the reference compounds Aliskiren (renin inhibitor), Amlodipine (calcium channel blocker) and Captopril (ACE inhibitor), indicating promising prospects for further studies.

Keywords: Hypertension. Antihypertensives. Renin. Calcium channel blockers. Angiotensin-converting enzyme. MIA-QSAR. Docking.

LISTA DE FIGURAS

PRIMEIRA PARTE

SEGUNDA PARTE - ARTIGOS

ARTIGO 1

ARTIGO 2

ARTIGO 3

LISTA DE TABELAS

SEGUNDA PARTE

ARTIGO 1

ARTIGO 2

ARTIGO 3

SUMÁRIO

PRIMEIRA PARTE

1 INTRODUÇÃO

De acordo com a Organização Mundial da Saúde, a pressão arterial elevada é um problema de saúde pública. Estimativas apontam que a doença cardiovascular atinge 2 bilhões de pessoas, com índice de óbitos de 9 milhões de pessoas por ano, sendo a principal causa de morte no mundo e promovendo a ocorrência de ataques cardíacos e derrames (ORGANIZAÇÃO MUNDIAL DA SAÚDE - OMS, 2013). Embora ainda não haja uma cura definitiva para hipertensão, esta doença pode ser controlada pela adoção de um estilo de vida saudável e a utilização de anti-hipertensivos.

Os anti-hipertensivos consistem de uma classe de fármacos, dentre os quais se têm os bloqueadores dos canais de cálcio (BCC), os inibidores da renina (IR) e os inibidores da enzima conversora de angiotensina (ECA), entre outros. Os BCC são vasodilatadores e baixam a pressão arterial por reduzir a resistência periférica vascular, enquanto os IR e os inibidores da ECA atuam no sistema renina-angiotensina-aldosterona (S-RAA), impedindo a formação da angiotensina II. A angiotensina II é um potente vasoconstritor e estimula a produção de aldosterona, que promove a retenção de sódio e água.

O Brasil tem, aproximadamente, 20 milhões de hipertensos (BRASIL, 2014) e cerca de 80% deles não conseguem controlar a pressão arterial com apenas uma terapia. Esses pacientes, que necessitam da combinação de medicamentos e enfrentam o desafio da adesão a diferentes tratamentos, poderiam ser beneficiados pela utilização de agentes hipertensivos multialvo, isto é, princípios ativos que atuem em mais de um alvo enzimático.

O planejamento racional de novos fármacos caracteriza-se por sua complexidade e elevado custo. Assim, a possibilidade de projetar compostos

com propriedades bem definidas, evitando os custos das sínteses experimentais exploratórias de grande número de substâncias, tem impulsionado pesquisas na área da Química Computacional. Os fundamentos necessários para um projeto efetivo nessa área estão na relação quantitativa estrutura-atividade (QSAR). Nas técnicas utilizadas em QSAR, considera-se que existe uma relação entre as propriedades de uma molécula e sua estrutura química 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 (HANSCH et al., 2002; MARTINS, 2010).

Uma aproximação aos modelos QSAR convencionais, igualmente preditiva, mas com algumas vantagens em termos de simplicidade e custo computacional, foi desenvolvida em 2005 e nomeada *Multivariate Image Analysis applied to QSAR*, ou MIA-QSAR (FREITAS; BROWN; MARTINS, 2005). Os descritores MIA têm sido aplicados com sucesso não só para correlacionar estruturas químicas com atividades biológicas, mas também com propriedades físicas (GOODARZI; FREITAS, 2008, 2009; GOODARZI; FREITAS; RAMALHO, 2009).

Este trabalho foi realizado com o objetivo de avaliar a capacidade de modelagem do método MIA-QSAR para duas classes de agentes anti-hipertensivos: inibidores da renina e bloqueadores dos canais de cálcio. A proposta e o estudo das interações de um ou mais compostos que apresentem atividade anti-hipertensiva multialvo também foram realizados, por meio de estudos de *docking*. Esses compostos são uma miscelânea de estruturas das duas classes de anti-hipertensivos aqui estudadas, além da classe dos inibidores da enzima conversora de angiotensina, presente na literatura (SILVA, 2013). A fim de alcançar os objetivos propostos, este trabalho está dividido em duas partes. Na primeira, verifica-se uma abordagem geral sobre o tema e, na segunda, são descritos os resultados, demonstrados em três artigos.

2 REFERENCIAL TEÓRICO

2.1 Sistema renina-angiotensina-aldosterona

O sistema renina-angiotensina-aldosterona (S-RAA, Figura 1) é um dos principais reguladores da fisiologia humana. O aumento da atividade desse sistema está associado às doenças cardiovasculares e à hipertensão, assim como danos em órgãos como o coração e os rins (RAIZADA; FERREIRA, 2007).

Trata-se de um sistema que promove a liberação de angiotensina (Ang) II, a qual exerce seus efeitos pela interação com receptores específicos (PAUL; MEHR; KREUTZ, 2006). A Ang II é gerada pela ação da renina, uma enzima produzida pelos rins, sobre o angiotensinogênio plasmático, produzido pelo fígado, formando o decapeptídeo Ang I (Asp^{1}-Arg^{2}-Val^{3}-Tyr^{4}-Ile^{5}-His^{6}-Pro^{7}-Phe^{8}-His^{9}-Leu^{10}), que é clivado na ligação Phe^{8}-His^{9} pela enzima conversora de Ang I, presente em abundância no endotélio pulmonar, liberando o octapeptídeo ativo Ang II (Asp^{1}-Arg^{2}-Val^{3}-Tyr^{4}-Ile^{5}-His^{6}-Pro^{7}-Phe^{8}) (MACIEL, 2013).

A etapa principal do S-RAA que contribui para a desregulação da pressão arterial ocorre com a formação da Ang II, responsável por ativar negativamente um ciclo que sobrecarrega os órgãos envolvidos na estabilidade hemodinâmica. Após realizar sua função, o processo finaliza-se quando a Ang II é inativada. Isso ocorre por meio da ação de enzimas angiotensinases e, como resultado, forma-se uma nova proteína: a Ang III (ERIKSSON et al., 2003).

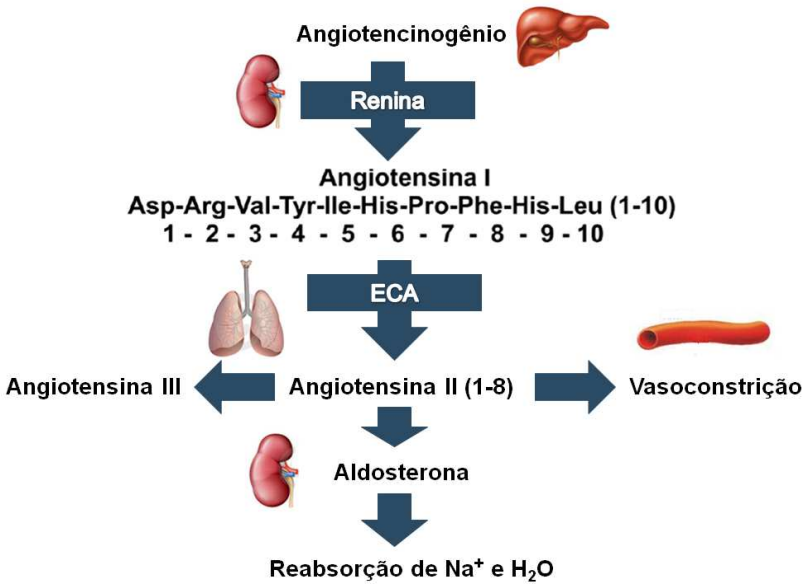


Figura 1 Sistema renina-angiotensina-aldosterona

2.2 Inibidores da renina

A renina é uma protease com dois lobos homólogos que, na sua fenda interna, tem dois resíduos de ácido aspártico com atividade catalítica. A renina cliva o angiotensinogênio, gerando Ang I, que é a precursora do produto ativo do sistema renina-angiotensina-aldosterona (S-RAA), a Ang II. A renina é formada no rim, a partir do seu precursor inativo, a prorenina que é inativa por apresentar um propeptídeo de 43 aminoácidos que recobre a fenda com atividade catalítica da renina. No rim, a remoção do propeptídeo forma a renina, também denominada de renina ativa. A renina participa na primeira etapa da ativação do S-RAA, clivando a ligação Leu10 – Val11 do angiotensinogênio. Sua localização privilegiada faz com que possa iniciar e determinar a velocidade de toda a cascata enzimática do S-RAA (RIBEIRO, 2006).

Os inibidores da renina (IR) têm grande afinidade pela renina, impedindo-a de clivar o angiotensinogênio. Agindo logo na primeira reação que gera a Ang II, os IR promoverão um bloqueio mais completo de S-RAA. Isso decorre do fato de a renina ser limitante da velocidade de reação e ter grande especificidade pelo seu substrato, o angiotensinogênio. O bloqueio direto da renina, logo no início da ativação do S-RAA, diminui a probabilidade de efeitos colaterais (GOESSLER; POLITO, 2012).

O alisquireno (Figura 2) é o representante atual dessa classe e promove inibição direta da ação da renina, com consequente diminuição da Ang II.

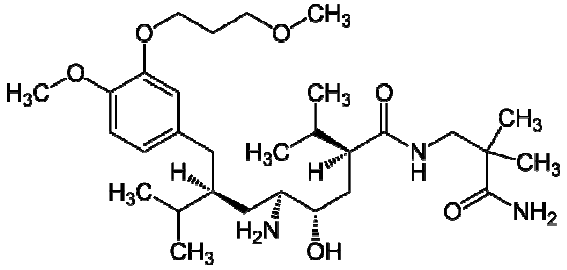


Figura 2 Estrutura química do alisquireno

2.3 Inibidores da enzima conversora de angiotensina

A enzima conversora de angiotensina (ECA) humana localiza-se na membrana plasmática e é uma metaloprotease que contém 1.278 resíduos de aminoácidos e tem dois domínios homólogos, cada um com um local catalítico e com uma região de ligação do zinco (SOUBRIE et al., 1988). Essa enzima glicoproteica zinco-dependente catalisa a remoção de aminoácidos de diferentes substratos peptídicos e é responsável pela conversão da Ang I em Ang II. Trata-se, portanto, de uma enzima chave no processo de regulação da pressão arterial, por apresentar especificidade pela Ang I.

Os inibidores da ECA precisam competir com o substrato natural pelo zinco (II) no sítio ativo, bem como por meio de ligações de hidrogênio e interações hidrofóbicas (PAULA et al., 2005). A Ang II é um potente vasoconstritor e estimula a produção de aldosterona, que promove a retenção de sódio e água. A enzima é estimulada pela secreção de renina pelos rins quando estes reagem à diminuição da sua perfusão sanguínea. Ao inibir essa enzima, os inibidores da ECA produzem vasodilatação periférica, diminuindo a pressão arterial (BAKRIS, 2009).

Os principais representantes dessa classe de medicamentos são o captopril e o enalapril (Figura 3), além do cilazapril, delapril, lisinopril e fosinopril, entre outros (SILVA, 2013).

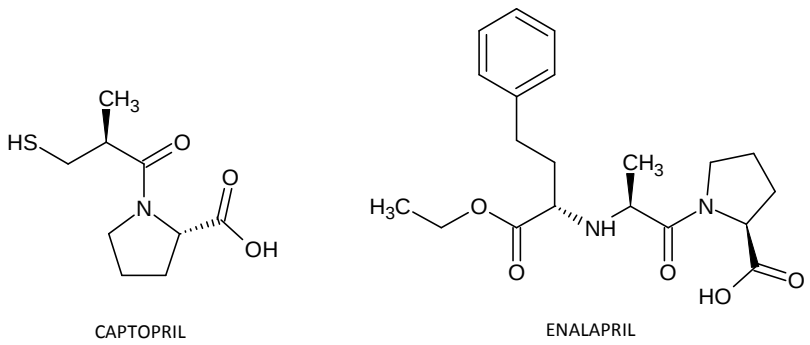


Figura 3 Estrutura química do captopril e do enalapril

2.4 Bloqueadores de canais de cálcio

O íon cálcio (Ca^{2+}) é um importante mensageiro intracelular, sendo fundamental nos mecanismos de excitação e contração da musculatura lisa do miocárdio e dos vasos.

A concentração extracelular de Ca^{2+} é de, aproximadamente, 5 M; sua concentração intracelular oscila entre 0,1 e 10 µM. Apesar desse gradiente

eletroquímico transmembranar favorável, o Ca^{2+} tem a sua entrada na célula restringida, pois é mediada por canais e transportadores específicos existentes na membrana plasmática (PIMENTEL, 2003).

Os bloqueadores dos canais de cálcio (BCC) impedem o fluxo de cálcio para dentro as células, incluindo células musculares cardíacas, células do sistema de condução do coração e da musculatura lisa do vaso, por bloqueio competitivo com o Ca^{2+} que entra pelos canais lentos voltagem-dependentes (dependente de um estímulo elétrico). Portanto, reduzem a excitabilidade do coração e a frequência cardíaca. O período de relaxamento é prolongado, o que leva à perda da velocidade de condução dos sinais do marca-passo fisiológico por todo o miocárdio (DOLLERY, 1991). Assim, a ação anti-hipertensiva dos BCC é decorrente da diminuição da resistência vascular periférica pela redução da concentração de Ca^{2+} nas células musculares lisas do vaso, relaxando a musculatura lisa e promovendo vasodilatação (BOMBIG; PÓVOA, 2009).

Há pelo menos seis tipos de canais voltagem-dependentes encontrados em vários tecidos: L, N, P, Q, R e T (OIGMAN; FRITSCH, 1998). Os BCC terapeuticamente importantes atuam sobre os canais do tipo L que compreendem três classes quimicamente distintas de fármacos: fenilalquilaminas, di-hidropiridinas e benzotiazepinas (BOMBIG; PÓVOA, 2009).

A estrutura proteica do canal L é bem conhecida e trata-se de uma macromolécula composta por várias subunidades, incluindo alfa1, alfa2, beta e delta (a gama existe apenas no músculo esquelético) (OIGMAN; FRITSCH, 1998).

Os fármacos de cada uma das três classes químicas mencionadas anteriormente ligam-se à subunidade α1 do canal de Ca^{2+} do tipo L, mas em locais distintos, que interagem alostericamente entre si e com o maquinário de controle de passagem do canal, impedindo sua abertura e, consequentemente, reduzindo a entrada de Ca^{2+}.

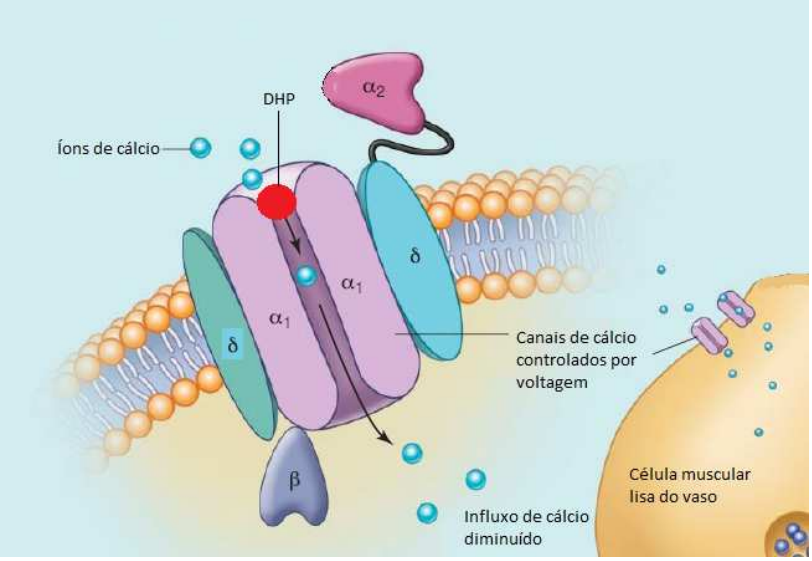


Figura 4 Canal de cálcio tipo L, bloqueado por di-hidropiridinas

Amlodipina e felodipina (Figura 4) são derivadas de di-hidropiridinas bastante conhecidas comercialmente como fármacos anti-hipertensivos (MOTA et al., 2013).

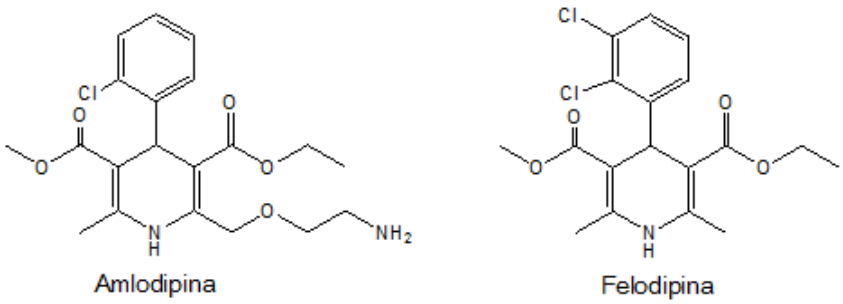


Figura 5 Estrutura química da amlodipina e felodipina

2.5 Biodisponibilidade e regra de Lipinski

A biodisponibilidade é um termo farmacocinético que descreve a velocidade e o grau com que uma substância ativa ou a sua forma molecular terapeuticamente ativa é absorvida a partir de um medicamento e se torna disponível no local de ação. A avaliação da biodisponibilidade é realizada com base em parâmetros farmacocinéticos calculados a partir dos perfis de concentração plasmática do fármaco ao longo do tempo (KELLER; PICHOTA; YIN, 2006).

A regra de Lipinski, também denominada de regra dos 5, tem como objetivo estimar a solubilidade e a permeabilidade de fármacos administrados pela via oral, predizendo a influência da estrutura química na absorção de um composto, uma vez que a previsão dos processos farmacocinéticos, logo nos estágios iniciais da pesquisa, é de extrema importância para o desenvolvimento de um candidato a fármaco (DUCHOWICZ et al., 2007). Segundo Lipinski, os critérios a serem analisados são: a massa molar, a qual não deve exceder a 500 g/mol; o log P, cujo valor limite é 5 e os grupos doadores (NH + OH) e aceptores (N + O) de ligação hidrogênio, cujas somatórias não devem ultrapassar a 5 e 10, respectivamente (LIPINSKI et al., 2001). Embora existam fármacos de sucesso que transgridam a regra de Lipinski, bem como outros que as obedeçam e que não tenham chegado ao mercado, a regra dos 5 de Lipinski tem servido de base para traçar tendências de perfis farmacocinéticos, com frequência, há muitos anos. Os parâmetros da regra de Lipinski, em conjunto com a área superficial polar da molécula (TPSA, relacionada com a biodisponibilidade oral), podem ser avaliados utilizando-se o programa Molinspiration (MOLINSPIRATION CHEMINFORMATICS, 1986). Essas propriedades, principalmente hidrofobicidade, ligações de hidrogênio, tamanho e flexibilidade da molécula e, claro, presença de características farmacofóricas, influenciam o

comportamento da molécula em um organismo vivo e devem ser levadas em consideração para propostas de novas drogas.

Introspecções similares acerca disso podem ser alcançadas utilizando-se o programa Hologram QSAR Technique, presente na literatura (MODA et al., 2008). Os modelos associados à base de dados desse programa conseguem avaliar a absorção intestinal humana, a biodisponibilidade oral, a penetração à barreira hemato-encefálica e a solubilidade de novos fármacos.

2.6 Desenvolvimento de novos fármacos

A descoberta de fármacos pela indústria farmacêutica é considerada, por especialistas, uma atividade complexa, multifatorial, cara, demorada, envolvendo a aplicação de técnicas e metodologias modernas, e cuja produtividade é questionada com base em dados que demonstram a relação inversamente proporcional entre os investimentos em pesquisa e desenvolvimento (P&D) e a descoberta de "New Chemical Entities" (NCEs) (MILNE, 2003).

As etapas envolvidas no processo de P&D baseiam-se, simplificadamente, no estudo de compostos para o tratamento de doenças. A primeira etapa consiste na descoberta de um composto com atividade terapêutica. Na segunda etapa, são feitos testes *in vitro* para avaliação das propriedades biológicas das moléculas obtidas, bem como por bioensaios *in vivo*, estudando o metabolismo e investigando a farmacocinética e a farmacodinâmica nos animais, o que é considerado o estudo pré-clínico. Na terceira e última etapa do processo são realizados estudos clínicos em humanos, em várias fases, parte denominada estudo clínico (FERREIRA et al., 2009; GUIDO; ANDRICOPULO; OLIVA, 2010; LOMBARDINO; LOWE, 2004).

Introduzir um novo medicamento na terapêutica é um processo longo e bastante oneroso, como citado anteriormente. Todo o processo de P&D dura cerca de doze anos, com probabilidade de sucesso muito pequena (LIMA et al., 2003). Assim, de cada 30.000 moléculas sintetizadas, cerca de 20.000 (66,7%) entram na fase de estudos pré-clínicos, 200 (0,67%) entram na fase I dos estudos clínicos, 40 (0,13%) passam para a fase II, 12 (0,04%) entram na fase III e somente 9 (0,027%) são aprovadas pelos órgãos regulatórios. É importante mencionar, ainda, que apenas um medicamento aprovado (0,003%) é incluído nos protocolos terapêuticos (CALIXTO; SIQUEIRA JÚNIOR, 2008).

O sucesso comercial de um fármaco depende de alguns fatores, tais como simplicidade estrutural (com possibilidade de modificações, visando à otimização de seu perfil farmacoterapêutico), ser membro de uma série congênere com relação estrutura/atividade (SAR) bem estabelecida e apresentar boas propriedades farmacocinéticas (OPREA et al., 2001).

2.7 *Quantitative Structure-Activity Relationship* (QSAR)

Com o desenvolvimento da tecnologia, computadores, softwares e hardwares, foram criados programas específicos para serem utilizados em problemas químicos. Criou-se, assim, um novo ramo da química, a Química Computacional.

Uma técnica que tem se destacado nesse meio é a relação quantitativa entre estrutura-atividade, ou QSAR, do inglês *Quantitative Structure-Activity Relationship*. O método baseia-se na possibilidade de a atividade e/ou propriedade ser uma função da estrutura molecular, ou seja, correlaciona a estrutura química com determinada propriedade química, física ou biológica, com o objetivo de fundamentar o planejamento de novas substâncias que tenham

perfil terapêutico mais adequado às necessidades (NEVES et al. 1998; TAVARES et al., 2004).

A área desperta grande interesse, visto que dispensa sínteses experimentais exploratórias de grande número de compostos, pois consegue projetar moléculas com propriedades bem definidas, o que diminui os custos e o tempo (HANSCH et al., 2002; MARTINS et al., 2009).

As técnicas utilizadas em QSAR surgiram em 1863, quando Cros, da Universidade de Estrasburgo, observou que a toxicidade de álcoois em mamíferos aumentava quando suas solubilidades em água diminuíam. Crum-Brown e Fraser postularam, em 1868, que existia uma relação entre as atividades fisiológicas e as estruturas químicas. Mais tarde, Richet propôs que a toxicidade de alguns álcoois e éteres era inversamente proporcional às suas solubilidades em água. Por volta de 1900, Meyer e Overton, trabalhando independentemente, estabeleceram relações lineares entre a ação narcótica de alguns compostos orgânicos e uma distribuição de coeficientes de solubilidade em água e em lipídios, descrevendo um parâmetro que pode ser considerado um precursor do atual log P, o coeficiente de partição octanol-água. Em 1939, Ferguson estudou o comportamento de propriedades diversas (solubilidade em água, partição, capilaridade e pressão de vapor) em relação à atividade tóxica de diferentes séries homólogas de compostos (CARBÓ-DORCA et al., 2000). Mesmo considerando esses procedimentos como as raízes do atual QSAR, no final da década de 1930, Hammett propôs o primeiro procedimento metodológico de propósito geral, tendo verificado que as constantes de equilíbrio de ionização dos ácidos benzóicos meta e para substituídos estavam relacionadas. Essa relação levou à definição da chamada constante de Hammett, parâmetro que se tornou um descritor capaz de caracterizar a atividade de muitos conjuntos de moléculas.

Em 1964, Free e Wilson postularam que, para uma série de compostos similares, diferindo entre si apenas pela presença de certos substituintes, a contribuição desses substituintes para a atividade biológica seria aditiva e dependeria apenas do tipo e da posição do substituinte (MARTINS, 2010).

A sistematização das análises em QSAR deve ser associada ao trabalho de Hansch e Fujita, surgido em 1964. As bases para o modelo de Hansch-Fujita é considerar que a atividade biológica observada é o resultado da contribuição de diferentes fatores que se comportam de maneira diferente. Cada contribuição para a atividade é representada por um descritor estrutural, enquanto a atividade biológica de um conjunto de compostos é ajustada em um modelo multilinear. Os descritores mais utilizados nas primeiras análises de QSAR foram o coeficiente de partição octanol/água (log P), a constante de Hammett e o parâmetro de lipofilicidade.

Nas últimas décadas, muitos avanços significativos foram observados na área de QSAR, destacando-se o desenvolvimento de métodos bidimensionais (QSAR 2D) (HONÓRIO; GARRATT; ANDRICOPOLOS, 2005) e tridimensionais (QSAR 3D) (HONÓRIO et al., 2007), além de métodos que envolvem informações moleculares mais avançadas (QSAR 4D e 5D). A principal diferença entre estes métodos está relacionada com o tipo de informação molecular utilizada para a construção do modelo quantitativo (DEBNATH, 2001).

A construção de um modelo QSAR segue cinco etapas principais: (i) seleção de um conjunto de dados com uma série de dados de resposta conhecidos, (ii) cálculo de descritores, (iii) divisão do conjunto de dados para treinamento e conjuntos de testes para o desenvolvimento do modelo e sua posterior validação, (iv) construção de modelos utilizando diferentes ferramentas quimiométricas e (v) validação do modelo desenvolvido com base interna e estatísticas de validação externa.

2.8 *Multivariate image analysis applied to quantitative structure-activity relationships* (MIA-QSAR)

O método de análise multivariada de imagem aplicada às relações quantitativas entre estrutura-atividade, ou MIA-QSAR (FREITAS; BROWN; MARTINS, 2005), consiste numa técnica simples, bastante utilizada para auxiliar nas propostas de novas drogas. Essa abordagem inicia-se com a criação de descritores binários que são obtidos a partir dos *pixels* de desenhos de estruturas químicas 2D. Essa técnica pode envolver o uso de programas de livre acesso, facilitando o uso da metodologia. Por ser uma técnica com utilização de imagens 2D para gerar descritores, eliminam-se muitas complicações envolvidas em QSAR 3D, tais como varredura conformacional e regras de alinhamento tridimensional dos ligantes. Assim, pode-se alcançar um avanço no custo computacional (FREITAS et al., 2006).

A imagem 2D de uma molécula não representa efeitos particulares, como, por exemplo, interações estéricas ou de qualquer parâmetro físico-químico, mas é uma forma geral para representar toda a estrutura química. Pequenas diferenças no desenho de uma estrutura podem causar mudanças significativas nas suas propriedades, ou seja, cada pixel atua como um código. Os descritores são binários, em que os pixels pretos correspondem ao dígito 0 e a cor branca corresponde ao número 765, de acordo com o sistema de composição de cores *red-green-blue*, ou RGB. Como exemplo, os enantiômeros *R* e *S* de uma molécula quiral podem ser distinguidos pelo desenho de cunha ou tracejado, para as ligações nesse centro (GOODARZI; FREITAS, 2009); isso fará com que, se for o caso, diferenças substanciais na propriedade estudada sejam notadas (por exemplo, para as atividades biológicas), ou seja, esse método é altamente sensível ao desenho. Tal fato é confirmado por Tropsha ao ressaltar que o resultado não depende do nível de representação da estrutura química, 2D

ou 3D, mas da capacidade do método de distinguir pares de enantiômeros (TROPSHA, 2010). Nesse sentido, o desenho das estruturas químicas compreende a etapa mais importante na construção de um modelo MIA-QSAR. As moléculas devem ser desenhadas cuidadosamente, mantendo a mesma disposição dos átomos, grupos substituintes e direção da cadeia carbônica para todas as estruturas (GOODARZI; FREITAS; RAMALHO, 2009), o que consiste no alinhamento bidimensional do conjunto de moléculas.

2.9 Parâmetros de validação

A validação de modelos de QSAR tem sido realizada por meio da predição, tanto interna (utilizando os compostos da série de treinamento) quanto externa (utilizando compostos que não participaram da série de treinamento) (LINDGREN et al., 1991).

Inicialmente, para validar um modelo QSAR, deve-se avaliar a qualidade do coeficiente de determinação r^2 (ou R^2), que analisa a qualidade da calibração e depende da variância total da variável dependente (*Syy*), de acordo com a equação

$$r^2 = 1 - \frac{\sum (y_{obs} - y_{calc})^2}{S_{yy}}$$

em que y_{obs} é o valor experimental da atividade biológica e y_{calc} é o valor calculado pelo modelo (KUBINYI, 1993). De forma geral, considera-se que um modelo QSAR pode ser aceito se o coeficiente de determinação for maior que 0,70, embora a literatura não seja taxativa a respeito desse número.

Na estratégia de validação cruzada, amostras do conjunto de treinamento são retiradas do modelo inicial, enquanto um novo modelo é construído para as restantes. A partir desse modelo, as atividades biológicas das amostras retiradas

são preditas, calculando-se os resíduos entre os valores obtidos e os valores reais. Esse procedimento é repetido até que todas as amostras tenham sido excluídas ao menos uma vez. O procedimento de validação cruzada mais comum é o *leave-one-out* (deixe uma fora), em que uma única amostra é retirada de cada vez para a construção do modelo. Como resultado deste procedimento, obtém-se o coeficiente de determinação da validação cruzada, q^2 (ou Q^2), calculado de acordo com a equação

$$q^2 = 1 - \frac{\sum (y_{obs} - y_{calc})^2}{\sum (y_{obs} - y_{médio})^2}$$

O valor de q^2 é utilizado como um critério para avaliar a robustez e a habilidade preditiva do modelo. Valores de $q^2 > 0{,}5$ são considerados aceitáveis para a validação do modelo. A diferença entre r^2 e q^2 não deve ser maior do que 0,30, pois uma diferença substancialmente maior pode indicar um modelo com ajuste forçado (*overfitting*) ou presença de variáveis independentes irrelevantes ou, ainda, de *outliers* nos dados (ERIKSSON et al., 2003).

Outra abordagem que deve ser utilizada paralelamente a essas técnicas de validação é a randomização das atividades biológicas (WOLD; ERIKSSON, 1995). Essa abordagem consiste em repetir o procedimento de construção do modelo com o bloco das variáveis dependentes (atividades biológicas, por exemplo) randomizadas e, em seguida, fazer uma avaliação estatística dos modelos assim produzidos. Se todos os modelos obtidos com as atividades randomizadas apresentarem coeficientes de correlação r^2 e q^2 elevados, significa que a correlação foi obtida por acaso e não porque a atividade biológica está definitivamente relacionada aos descritores empregados no modelo. Isso implica que não será possível obter um modelo QSAR confiável para esse conjunto de dados utilizando o método empregado. Em geral, modelos aceitáveis apresentam

valor de R^2_r (ou r^2_{rand}) menor que 0,50 (GRAMATICA, 2007; TROPSHA, 2005).

O processo de randomização analisa a robustez do processo de construção do modelo. Para avaliar se a diferença entre R^2 e R^2_r é estatisticamente relevante, isto é, para se certificar que R^2_r é, de fato, um valor pobre e que o modelo pode ser considerado confiável, utiliza-se o parâmetro R^2_P, do inglês *penalized* R^2 (MITRA; SAHA; ROY, 2009; ROY et al., 2009). O parâmetro R^2_P deve ser superior a 0,5 para que o modelo QSAR possa ser considerado confiável e não o resultado de um mero acaso. O valor de R^2_p é obtido utilizando-se a seguinte fórmula:

$$R^2_p = R^2\sqrt{(R^2 - R^2_r)}.$$

No entanto, no caso ideal, o valor médio de R^2 para os modelos randomizados deve ser zero, isto é, R^2_r deveria ser zero. Consequentemente, o valor de R^2_P deve ser igual ao valor de R^2 para o modelo QSAR desenvolvido. Assim, a fórmula corrigida de R^2_P ($^cR^2_p$), como proposto por Todeschini e Consonni (2010), é dada por

$$^cR^2_p = R\sqrt{(R^2 - R^2_r)}.$$

Para refinar ainda mais a análise da capacidade preditiva dos modelos QSAR, outro parâmetro foi desenvolvido para detectar a proximidade entre a atividade observada e a prevista: o parâmetro r_m^2, do inglês *modified* r^2 (ROY; ROY, 2008). O r_m^2 é calculado com base na correlação dos dados observados e preditos (com e sem intercepto) e também pela troca dos eixos (ROY et al., 2013).

$$r_m^2 = r^2 \times (1 - \sqrt{r^2 - r_0^2})$$

De acordo com esse critério, para um modelo ser considerado preditivo, o valor de r_m^2 deve ser superior a 0,5 (OJHA et al., 2011).

2.10 Análise multivariada de dados

Para análises quantitativas pode-se utilizar a regressão por mínimos quadrados parciais (PLS) e, para análises qualitativas, pode-se utilizar a análise de componentes principais (PCA).

2.10.1 Análise por componentes principais (PCA)

A análise de componentes principais (PCA, do inglês *principal component analysis*) é uma técnica multivariada que consiste em transformar um conjunto de variáveis originais em outro conjunto de variáveis de mesma dimensão, denominado componentes principais. É realizada sobre uma matriz de dados que relaciona um conjunto de variáveis e amostras e, geralmente, essa matriz de dados X pré-processada é, então, decomposta no produto de três matrizes (BROWN, 1995).

Utiliza-se a PCA no intuito de reduzir a dimensionalidade dos dados originais por meio da geração de novas variáveis, denominadas componentes principais, preservando o maior número possível de informação contida nas variáveis originais (MATOS et al., 2003).

Como resultados da análise de componentes principais, são gerados dois conjuntos de dados chamados de escores e pesos, os quais trazem informações

sobre as amostras e as variáveis, respectivamente (CORREIA; FERREIRA, 2007; MATOS et al., 2003).

Portanto, a PCA é um tratamento matemático que identifica, no hiperespaço das variáveis, a direção na qual está contida a maior parte das informações. Assim, por meio da projeção dos resultados analíticos de cada amostra no espaço formado pelas novas componentes principais, é possível demonstrar diferenças entre as várias amostras, ou grupos de amostras (gráfico de pontuação), determinando, ao mesmo tempo, quais variáveis principais estão envolvidas (BRUNS; FAIGLE, 1985; LAVINE, 2000; LAVINE; WORKMAN, 2006).

2.10.2 Análise de regressão por mínimos quadrados parciais (PLS)

Esta técnica foi desenvolvida na década de 1970, por Herman Wold. Foi usada, primeiramente, em associação à espectroscopia na região do infravermelho próximo, em que é difícil designar bandas para componentes particulares. No modelo PLS, as informações espectrais (matriz X) e as informações das concentrações (matriz Y) são correlacionadas, a fim de se obter uma relação linear na fase de calibração. A regressão por mínimos quadrados parciais, para a construção do modelo, retira informações do conjunto de dados da matriz espectral (matriz X) para correlacionar com as informações retiradas do conjunto de dados de referência (matriz Y) e obter o número de variáveis latentes necessários necessárias para fazer correlação entre os espectros e as concentrações (HOPKE, 2003; MILLER; MILLER, 2000).

A variável latente descreve a direção de máxima variância que também se correlaciona com a concentração. Portanto, as variáveis latentes são, na realidade, combinações lineares das componentes principais calculadas pelo método PCA (FERREIRA et al., 1999). Para a construção do modelo de

calibração, é utilizado um número de variáveis latentes que proporcione o menor erro possível de previsão.

REFERÊNCIAS

BAKRIS, G. L. Is blockade of the renin-angiotensin system appropriate for all patients with diabetes? **Journal of the American Society of Hypertension**, New York, v. 5, n. 7, p. 288-290, July 2009.

BOMBIG, M. T. N.; PÓVOA, R. Interações e associações de medicamentos no tratamento anti-hipertensivo: antagonistas dos canais de cálcio. **Revista Brasileira de Hipertensão**, Rio de Janeiro, v. 16, n. 4, p. 226-230, Oct. 2009.

BRASIL. Ministério da Saúde. **Hipertensão**. Disponível em: <http://portalsaude.saude.gov.br/>. Acesso em: 6 jan. 2014.

BROWN, S. D. Chemical systems under indirect observation: latent properties and chemometrics. **Applied Spectroscopy**, Baltimore, v. 49, n. 12, p. 14A-31A, Dec. 1995.

BRUNS, R. E.; FAIGLE, J. F. G. Quimiometria. **Química Nova**, São Paulo, v. 8, n. 2, p. 84-99, mar./abr. 1985.

CALIXTO, J. B.; SIQUEIRA JÚNIOR, J. M. Desenvolvimento de medicamentos no Brasil: desafios. **Gazeta Médica da Bahia**, Salvador, v. 78, n. 1, p. 98-106, Feb. 2008.

CARBÓ-DORCA, R. et al. **Molecular quantum similarity in QSAR and drug design**. Girona: University of Girona, 2000. 123 p.

CORREIA, P. R. M.; FERREIRA, M. M. C. Reconhecimento de padrões por métodos não supervisionados: explorando procedimentos quimiométricos para tratamento de dados analíticos. **Química Nova**, São Paulo, v. 30, n. 2, p. 481-487, mar./abr. 2007.

DEBNATH, A. K. Quantitative structure-activity relationship (QSAR) paradigm--Hansch era to new millennium. **Mini Reviews in Medicinal Chemistry**, Beijing, v. 1, n. 2, p. 187-195. July 2001.

DOLLERY C. Clinical pharmacology of calcium antagonists. **American Journal of Hypertension**, New York, v. 2, n. 4, p. 88-95, Feb. 1991.

DUCHOWICZ, P. R. et al. Applicationof descriptores based on Lipinski's rules in the QSPR study of aqueous solubilities. **Bioorganic & Medecinal Chemistry**, Oxford, v. 15, n. 4, p. 3711-3719, Mar. 2007.

ERIKSSON, L. et al. Methods for reliability and uncertainty assessment and for applicability evaluations of classification- and regression-based QSARs. **Environmental Health Perspectives**, New York, v. 111, p. 1361-1375, Aug. 2003.

FERREIRA, F. G. et al. Fármacos: do desenvolvimento à retirada do mercado. **Revista Eletrônica de Farmácia**, Goiânia, v. 6, n. 1, p. 14-24, jan. 2009.

FERREIRA, M. M. C. et al. Quimiometria I: calibração multivariada, um tutorial. **Química Nova**, São Paulo, v. 22, n. 5, p. 724-731, set./out. 1999.

FREITAS, M. P. MIA-QSAR modelling of anti-HIV-1 activities of some 2-amino-6-arylsulfonylbenzonitriles and their thio and sulfinyl congeners. **Organic & Biomolecular Chemistry**, Cambridge, v. 4, n. 6, p. 1154-1159, Mar. 2006.

FREITAS, M. P.; BROWN, S. D.; MARTINS, J. A. MIA-QSAR: a simple 2D image-based approach for quantitative structure-activity relationship analysis. **Journal of Molecular Structure**, Amsterdam, v. 738, n. 1/3, p. 149-154, Mar. 2005.

GOESSLER, K. F.; POLITO, M. D. Inibidor direto da renina e hipertensão arterial: uma revisão. **Revista Brasileira de Cardiologia**, Rio de Janeiro, v. 25, n. 3, p. 241-255, jun. 2012.

GOODARZI, M.; FREITAS, M. P. Ant colony optimization as a feature selection method in the QSAR modeling of anti-HIV-1 activities of 3-(3,5-dimethylbenzyl)uracil derivatives using MLR, PLS and SVM regressions. **Chemometrics and Intelligent Laboratory Systems**, Cambridge, v. 98, n. 2, p. 123-127, Oct. 2009.

GOODARZI, M.; FREITAS, M. P. Predicting boiling points of aliphatic alcohols through multivariate image analysis applied to quantitative structure#property relationships. **Journal of Physical Chemistry A**, Washington, v. 112, n. 44, p. 11263-11265, Oct. 2008.

GOODARZI, M.; FREITAS, M. P. Prediction of electrophoretic enantioseparation of aromatic amino acids/esters through MIA-QSPR. **Separation and Purification Technology**, London, v. 68, p. 363-366, Aug. 2009.

GOODARZI, M.; FREITAS, M. P.; RAMALHO, T. C. Prediction of ^{13}C chemical shifts in methoxyflavonol derivatives using MIA-QSPR. **Spectrochimica Acta Part A**, Kidlington, v. 74, p. 563-568, Oct. 2009.

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

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

HANSCH, C. et al. Chem-bioinformatics: comparative QSAR at the interface between chemistry and biology. **Chemical Reviews**, Washington, v. 102, n. 3, p. 783-812, Mar. 2002.

HONÓRIO, K. M. et al. 3D QSAR comparative molecular field analysis on nonsteroidal farnesoid X receptor activators. **Journal of Molecular Graphics and Modelling**, New York, v. 25, n. 6, p. 921-927, Mar. 2007.

HONÓRIO, K. M.; GARRATT, R. C.; ANDRICOPOLOS, A. D. Hologram quantitative structure-activity relationships for a series of farnesoid X receptor activators. **Bioorganic & Medicinal Chemistry Letters**, Oxford, v. 15, n. 12, p. 3119-3125, June 2005.

HOPKE, P. K. The evolution of chemometrics. **Analytica Chimica Acta**, Amsterdam, v. 500, n. 1/2, p. 365-377, 2003.

KELLER, T. H.; PICHOTA, A.; YIN, Z. A practical view of 'druggability'. **Current Opinion in Chemical Biology**, London, v. 10, p. 357-361, Aug. 2006.

KUBINYI, H. **QSAR:** Hansch analysis and related approaches. New York: VCH, 1993. 240 p.

LAVINE, B. K. Chemometrics. **Analytical Chemistry**, Washington, v. 72, n. 12, p. 91-97, Dec. 2000.

LAVINE, B. K.; WORKMAN, J. Chemometrics. **Analytical Chemistry**, Washington, v. 78, n. 12, p. 4137-4145, Dec. 2006.

LIMA, J. S. et al. Pesquisa clínica: fundamentos, aspectos éticos e perspectivas. **Revista da Sociedade de Cardiologia do Estado do Rio de Janeiro**, Rio de Janeiro, v. 16, n. 4, p. 225-233, mar. 2003.

LINDGREN, F. et al. A strategy for ranking environmentally occurring chemicals: part IV, development of chemical model systems for characterization of halogenated aliphatic hydrocarbons. **Molecular Informatics**, New York, v. 10, n. 1, p. 36-42, 1991.

LIPINSKI, C. A. et al. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. **Advanced Drug Delivery Reviews**, London, v. 46, n. 1, p. 3-26, Mar. 2001.

LOMBARDINO, J. G.; LOWE, J. A. The role of the medicinal chemist in drug Discovery: then and now. **Nature Reviews Drug Discovery**, London, v. 3, n. 10, p. 853-862, Oct. 2004.

MACIEL, R. P. **Efeitos da inibição da enzima conversora de angiotensina sobre a doença periodontal induzida experimentalmente em ratos**. 2013. 84 p. Dissertação (Mestrado em Ciências) - Universidade de São Paulo, Bauru, 2013.

MARTIN, J. 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, June 2009.

MARTINS, J. P. A. **Química computacional aplicada a QSAR**. São Paulo: CENAPAD, 2010. 123 p.

MATOS, G. D. et al. Análise exploratória em química analítica com emprego de quimiometria: PCA e PCA de imagens. **Analytica**, São Paulo, v. 6, n. 3, p. 38-40, ago./set. 2003.

MILLER, J. N.; MILLER, J. C. **Statistic and chemometrics for analytical chemistry**. New York: Prentice Hall, 2000. 238 p.

MILNE, G. M. Pharmaceutical productivity: the imperative for new paradigms. **Annual Reports in Medicinal Chemistry**, Palo Alto, v. 38, n. 2, p. 383-396, Oct. 2003.

MITRA, I.; SAHA, A.; ROY, K. Exploring quantitative structure–activity relationship studies of antioxidant phenolic compounds obtained from traditional

Chinese medicinal plants. **Molecular Simulation**, New York, v. 36, n. 36, p. 1067-1079, Nov. 2010.

MODA, T. L. et al. PK/DB: database for pharmacokinetic properties and predictive in silico ADME models. **Bioinformatics**, Oxford, v. 24, n. 19, p. 2270-2271, Oct. 2008.

MOLINSPIRATION CHEMINFORMATICS. **Molinspiration**. New Ulica, 1986. Disponível em: <http://www.molinspiration.com/>. Acesso em: 5 ago. 2013.

MOTA, E. G. et al. Computer-assisted design of novel 1,4-dihydropyridine calcium channel blockers. **Molecular Simulation**, New York, v. 40, n. 12, p. 959-965, Sept. 2013.

NEVES, P. J. et al. TOP: um programa de cálculo de descritores topológicos para uso em correlações entre estrutura e atividade. **Química Nova**, São Paulo, v. 21, n. 6, p. 709-713, abr. 1998.

OIGMAN, W.; FRITSCH, M. T. Antagonistas de canais de cálcio. **HiperAtivo**, São Paulo, v. 5, n. 2, p. 104-109, jun. 1998.

OJHA, P. K. et al. Further exploring r^2_m metrics for validation of QSPR models. **Chemometrics and Intelligent Laboratory Systems**, Berlin, v. 107, n. 1, p. 194-205, May 2011.

OPREA, T. I. et al. Is there a difference between leads and drugs?: a historical perspective. **Journal of Chemical Information and Computer Sciences**, Washington, v. 41, n. 5, p. 1308-1315, Oct. 2001.

ORGANIZAÇÃO MUNDIAL DA SAÚDE. **Hipertensão no mundo**. Disponível em: <http://www.who.int/en/>. Acesso em: 6 jan. 2013.

PAUL, M.; MEHR, A. P.; KREUTZ, R. Physiology of local renin-angiotensin systems. **Physiological Reviews**, Baltimore, v. 86, n. 3, p. 747-803, July 2006.

PAULA, W. X. A química inorgânica no planejamento de fármacos usados no controle da hipertensão. **Cadernos Temáticos de Química Nova na Escola**, São Paulo, v. 1, n. 6, p. 19-23, jul. 2005.

PIMENTEL, S. **Canais e transportadores de cálcio**. Porto: Faculdade de Medicina da Universidade do Porto, 2003. 11 p.

RAIZADA, M. K.; FERREIRA, A. J. ACE2: a new target for cardiovascular disease therapeutics. **Cardiovascular Pharmacology**, New York, v. 50, n. 2, p. 112-119, Aug. 2007.

RIBEIRO, A. B. Inibidores da renina: uma nova classe de hipertensivos. **Revista Brasileira de Hipertensão**, São Paulo, v. 13, n. 3, p. 219-220, jun. 2006.

ROY, K. et al. Some case studies on application of "$r(m)^2$" metrics for judging quality of quantitative structure-activity relationship predictions: emphasis on scaling of response data. **Journal of Computational Chemistry**, Washington, v. 34, n. 12, p. 1071-1082, Jan. 2013.

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

ROY, P. P.; ROY, K. On some aspects of variable selection for partial least squares regression models. **QSAR Combinatorial Science**, New York, v. 27, n. 3, p. 302-313, Mar. 2008.

SILVA, D. G. **Modelagem racional de novos dipeptídeos modificados como inibidores da enzima conversora de angiotensina**. 2013. 121 p. Dissertação (Mestrado em Agroquímica) - Universidade Federal de Lavras, Lavras, 2013.

SOUBRIER, F. et al. Two putative active centers in humamangiotensin I-converting enzyme revealed by molecular cloning. **PNAS**, New York, v. 85, n. 24, p. 9386-9390, Dec. 1988.

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

TODESCHINI, R.; CONSONNI, V. **Handbook of molecular descriptors**. Weinheim: Wiley-VCH, 2000. 524 p.

TROPSHA, A. Application of predictive QSAR models to database mining. In: OPREA, T. (Ed.). **Chemoinformatics in drug discovery**. Weinheim: Wiley-VCH, 2005. p. 440-458.

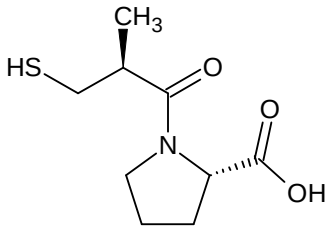
TROPSHA, A. Best practices for QSAR model development, validation, and exploration. **Molecular Informatics**, New York, v. 29, n. 6/7, p. 476-488, July 2010.

WOLD, S.; ERIKSSON, L. Statistical validation of QSAR results. In: WATERBEEMD, H. van der (Ed.). **Chemometric methods in molecular design**. Weinheim: Wiley-VCH, 1995. p. 309-318.

SEGUNDA PARTE - ARTIGOS

Esta parte do trabalho é dedicada à exposição dos artigos elaborados durante o desenvolvimento desta tese. Além de apresentar um estudo computacional para o desenvolvimento de modelos MIA-QSAR e a proposição de novos compostos inibidores da renina e bloqueados de canais de cálcio, em um terceiro artigo descrevem-se estudos de *docking* de compostos concebidos a partir da combinação de subestruturas de três tipos de anti-hipertensivos: inibidores da renina, bloqueadores do canal de cálcio e inibidores da enzima conversora de angiotensina (ECA), a fim de se obter um composto multialvo.

Os resultados aqui apresentados que façam referência aos inibidores da ECA se utilizaram de dados obtidos por SILVA (2013), em seu trabalho sobre modelagem de novos peptídeos modificados como inibidores dessa enzima, no qual o dipeptídeo modificado captopril, um dos fármacos mais utilizados no tratamento da hipertensão, foi utilizado como composto de referência (Figura 1).



CAPTOPRIL

Figura 1 Estrutura química do captopril

Um breve resumo do trabalho de Silva (2013) é apresentado a seguir, para fins de contextualização acerca dos inibidores da ECA empregados nos estudos de *docking* desta tese (Artigo 3).

Para a proposta de novos dipeptídeos, Silva (2013) fez uso de técnicas computacionais (MIA-QSAR, QSAR-tradicional e *docking*) associadas à

estratégia de combinar subestruturas químicas de moléculas altamente bioativas como inibidores da ECA. Assim, foram propostas oito novas moléculas, presentes na Figura 2.

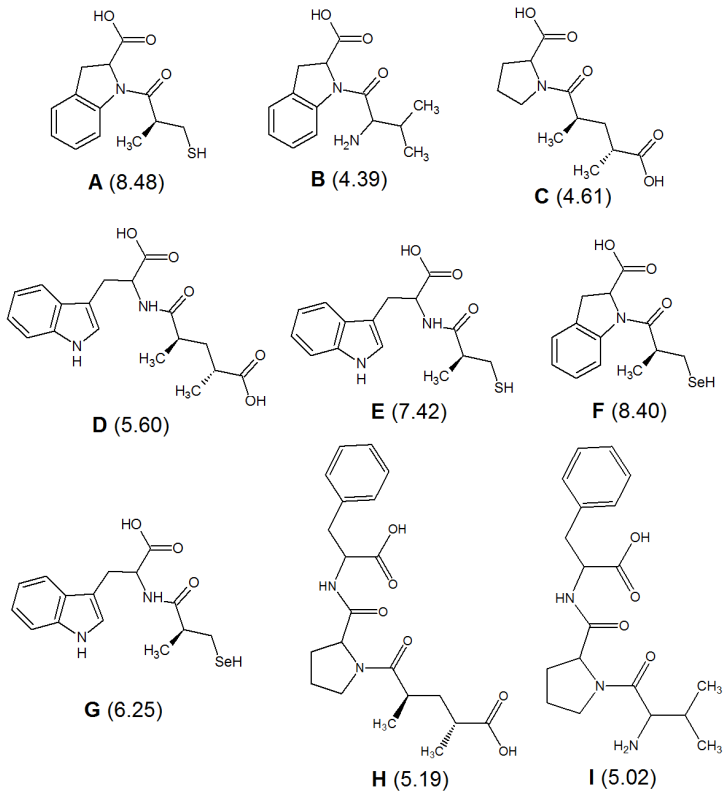
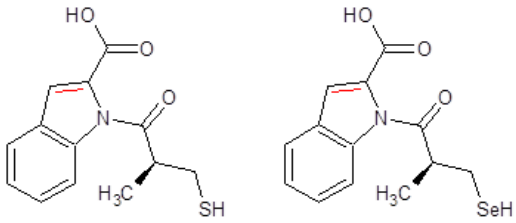


Figura 2 Estruturas propostas usando a estratégia de combinação de subestruturas dos compostos mais ativos da série. Os valores preditos de pIC_{50}, obtidos pelo modelo MIA-QSAR, estão apresentados entre parêntesis

Pode-se notar que as moléculas que contêm um grupo SH ou SeH foram mais promissoras do que as outras concebidas a partir da combinação de subestruturas químicas sem esses grupos. A análise de *docking* demonstrou a influência dos átomos de enxofre e selênio, que aumentam o poder de inibição,

pois realizam uma ligação de hidrogênio a mais (com a Glu385) do que as outras moléculas, validando as metodologias QSAR empregadas.

A última modificação nas estruturas consistiu na adição da dupla ligação no anel de cinco membros dos compostos **A** e **F**, que deu origem às moléculas **J** e **K** (Figura 3), que apresentaram melhor comportamento. De fato, a parte indol dessas moléculas pode exercer influência na atividade biológica em conjunto com os átomos de enxofre/selênio, validando os estudos computacionais utilizados no trabalho em questão.



J **K**

Figura 3 Novas estruturas propostas **J** e **K**; uma dupla ligação foi acrescentada nos compostos **A** e **F**, resultando em um indol modificado

As estruturas propostas exibiram propriedades farmacocinéticas similares ou melhores que as do captopril, demonstrando que essas moléculas apresentam potenciais propriedades farmacocinéticas para uma adequada absorção, distribuição, metabolismo e excreção.

ARTIGO 1

THEORETICAL DESIGN OF NEW INDOLE-3-CARBOXAMIDE DERIVATIVES AS RENIN INHIBITORS

Artigo redigido conforme normas da revista Medicinal Chemistry Research

Estella G. da Mota, Mariene H. Duarte, Elaine F. F. da Cunha, Matheus P. Freitas

Abstract

Renin inhibitors pertain to a new generation class of anti-hypertensive agents. There are only a few studies on the computational modeling of such class of compounds and only one available drug in the market used as renin inhibitor for the treatment of hypertension, Aliskiren. The present study reports the QSAR modeling of the activities of a series of indole-3-carboxamide derivatives using MIA-QSAR in order to propose new promising analogues as renin inhibitor candidates. The proposed structures were submitted to docking evaluation to search for the interaction modes responsible for the calculated bioactivities. In addition, the drug-likeness of the proposed compounds were investigated using theoretical data related to pharmacokinetic properties. Overall, at least two promising candidates are proposed as highly active and pharmacokinetically acceptable renin inhibitors.

Keywords: docking studies, hypertension, indole-3-carboxamides, MIA-QSAR, renin inhibitors

1. Introduction

According to the World Health Organization (www.who.int), high blood pressure is a major public health problem. Estimates indicate that cardiovascular diseases affect 2 billion people worldwide, while about 9 million people die each year as a consequence of cardiovascular diseases. Although there are not a permanent cure for hypertension, this disease can be controlled by adopting a healthy lifestyle and through the use of antihypertensive drugs (da Mota *et al.*, 2014).

The renin–angiotensin system plays a key role in the regulation of blood pressure (Jing *et al.*, 2012). The mechanism involves conversion of angiotensin I (inactive) into angiotension II through the angiotension-converting enzyme (ACE), which is active and responsible for most of pressor effects (Paul *et al.*, 2006). Some antihypertensives act by inhibiting the ACE, while recent studies are devoted to develop drugs capable of acting in an earlier stage, *i.e.* the inhibition of renin, which is responsible for hydrolyzing angiotensinogen from the liver into angiotensin I (Jing *et al.*, 2012; Wood *et al.*, 1987).

In the last decades, substantial efforts have been reported with the aim at discovering renin inhibitors and overcoming issues like low oral bioavailability. Aliskiren (SPP100), an orally active renin inhibitor with four chiral centers (Fig. 1), is currently the only compound which has reached the market (Scheiper *et al.*, 2010). Aliskiren-based drug design against HIV-I protease has also been performed to try eliminating side-effects related to hypertension and diabetes and also to facilitate AIDS therapy (Tzoupis *et al.*, 2012).

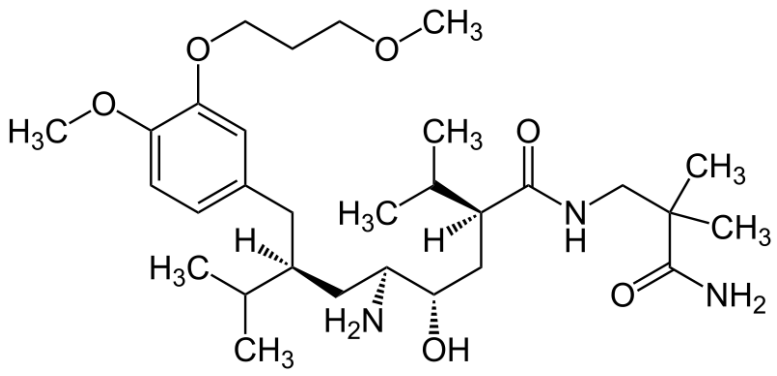


Figure 1 Aliskiren: Common renin inhibitor for the treatment of hypertension

In this study, we propose the building of a QSAR model capable of estimating accurately the bioactivities of new indole-3-carboxamide derivatives as renin inhibitors. In order to realize new antihypertensive drugs with low structural complexity and high biological activity and oral bioavailability, some new classes of direct renin inhibitors have been recently developed (Scheiper *et al.*, 2010; Corminboeuf *et al.*, 2010), among which nonpeptide indole-3-carboxamide derivatives appear to be good candidates. Indole-3-carboxamide derivatives have no chiral centers and possess low molecular weight and considerable rigidity. Thus, their synthesis can be driven by molecular structures conceived computationally.

Computational molecular modeling, like quantitative structure–activity relationship (QSAR) analysis, can be effective in drug design, since the bioactivities of new drug candidates can be estimated even before synthesis and biological tests. This can be achieved by correlating experimentally available bioactivities of a given class of drug like molecules with the corresponding molecular descriptors, which are numerical data encoding molecular structures (Guimarães *et al.*, 2014).

In this work, multivariate image analysis applied to QSAR (MIA-QSAR) was the technique used to generate MIA descriptors for further

regression with the corresponding biological data using partial least squares (PLS) regression. In MIA-QSAR (Freitas, 2007), descriptors are pixels (binaries) of images, corresponding to the 2D chemical structures as wireframes; the substructural diversity among the set of compounds (usually the substituents) explains the variance in the bioactivity values, while the congruent substructure of the series is used for 2D alignment.

To validate the predictions obtained by the QSAR model and to understand the interactions ruling the bioactivities of new proposed indole-3-carboxamide derivatives as direct renin inhibitors, docking studies were carried out to accompany the trends in calculated pIC_{50} in terms of docking scores (the energy of interaction between ligand and enzyme).

2. Methods and Materials

The QSAR model was developed using the MIA-QSAR strategy, in which descriptors correspond to pixels (binaries) of 2D chemical structures drawn with the aid of the ChemSketch program (ACD/ChemSketch, 2006). Then, each image was saved as bitmaps (.bmp files), a 2D array in which the variable pixels (substituents) correspond to the variable moiety in each molecule; these changes along with the series of compounds explain the variance in the activities block (the **y** column vector in the regression analysis). The pixels were converted to numerical values using the Matlab (Mathworks, 2007) platform, and the superposed images of Figure 2 illustrate the variance in the chemical drawings.

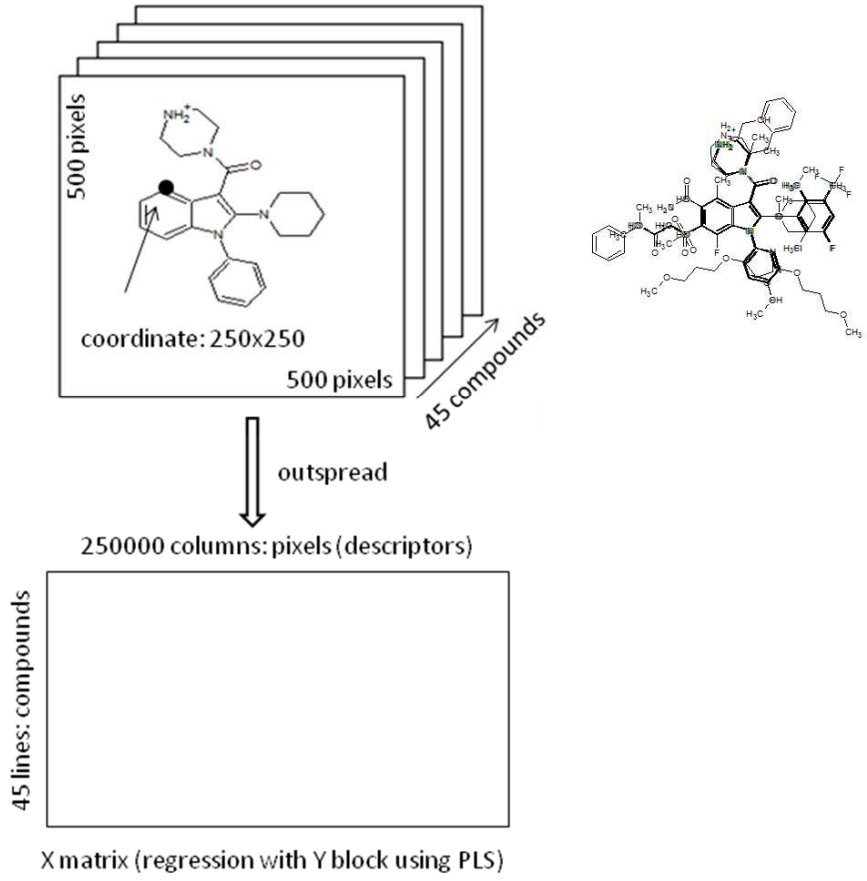
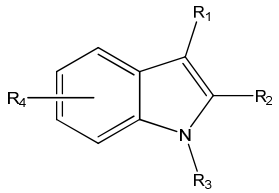


Figure 2 Construction of the 3D-array and subsequent unfolding to a 2D-array (**X** matrix). The arrow in the structure indicates the pixel used for the 2D alignment. The detail on the right represents the superposed images to illustrate the data variance

A series of 45 indole-3-carboxamide derivatives were taken from the literature (Jing *et al.*, 2012) (Table 1) and were converted into images of 500 × 500 pixels size each, as described elsewhere (Cormanich *et al.*, 2011). The images were grouped to give a three-way array of 45 × 500 × 500 dimension,

which was subsequently unfolded to a 45 × 250,000 matrix. Columns with zero variance (corresponding to blank space or congruent substructures in the workspace) were removed to minimize computing costs, giving a 45 × 6951 matrix, the **X** matrix.

Table 1 Series of 45 indole-3-carboxamide compounds and the corresponding experimental, fitted and predicted biological data (Jing *et al.*, 2012)



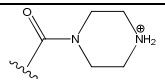
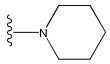
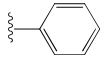
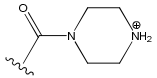
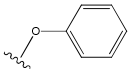
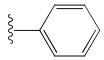
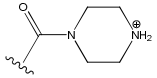
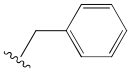
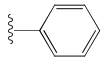
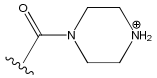
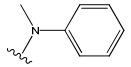
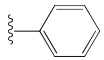
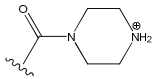
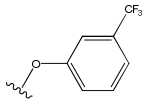
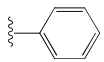
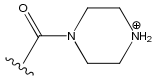
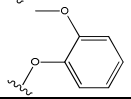
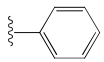
Cpd	R_1	R_2	R_3	R_4	pIC_{50}			
					Exp.	Calib.	LOOCV	Test
1				—	5.42	5.27	6.43	
2				—	6.38	6.37	6.37	
3				—	7.04	6.49	6.39	
4				—	5.50	5.98	6.38	
5 [a]				—	5.05			6.08
6				—	6.17	6.68	6.91	

Table 2. Continuation

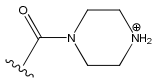
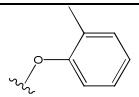
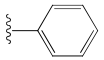
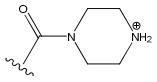
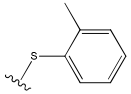
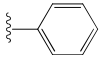
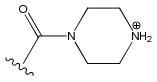
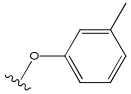
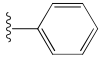
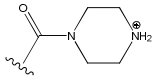
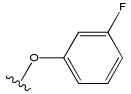
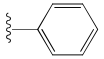
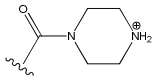
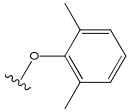
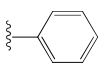
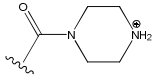
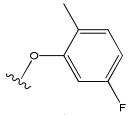
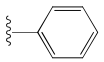
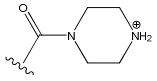
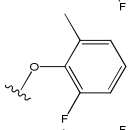
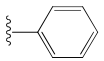
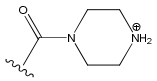
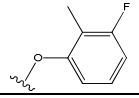
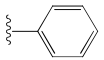
Cpd	R_1	R_2	R_3	R_4	pIC_{50}			
					Exp.	Calib.	LOOCV	Test
7				—	7.68	7.38	7.30	
8				—	7.15	7.40	7.41	
9 [a]				—	6.11			6.20
10				—	7.20	6.42	6.26	
11				—	7.19	7.57	7.55	
12				—	8.30	7.68	7.38	
13				—	7.80	7.58	7.40	
14 [a]				—	7.96			7.43

Table 3. Continuation

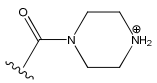
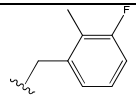
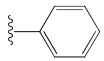
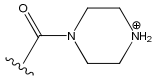
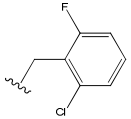
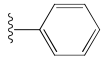
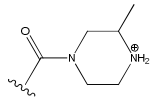
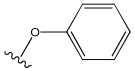
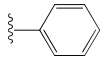
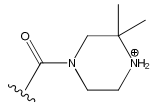
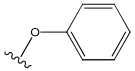
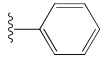
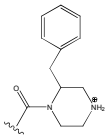
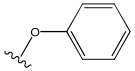
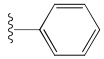
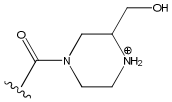
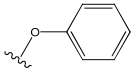
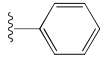
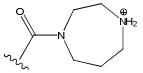
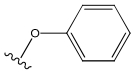
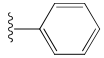
Cpd	R_1	R_2	R_3	R_4	pIC_{50}			
					Exp.	Calib.	LOOCV	Test
15				—	8.05	7.55	7.42	
16				—	7.74	7.16	6.81	
17				—	5.19	5.89	6.35	
18[a]				—	5.14			5.77
19				—	5.35	5.36	5.87	
20				—	5.80	6.01	6.29	
21				—	5.47	5.60	5.93	

Table 4. Continuation

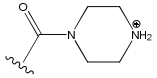
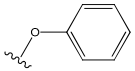
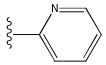
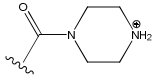
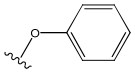
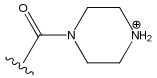
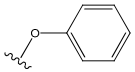
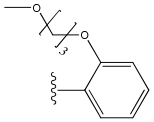
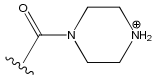
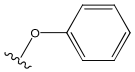
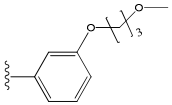
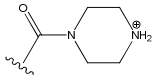
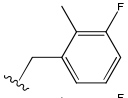
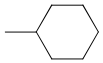
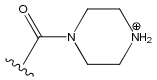
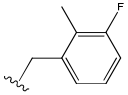
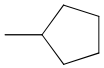
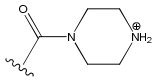
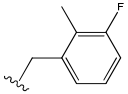
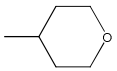
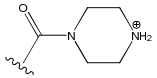
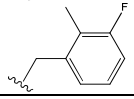
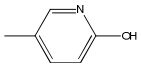
Cpd	R_1	R_2	R_3	R_4	pIC_{50}			
					Exp.	Calib.	LOOCV	Test
22				—	5.30	6.16	6.49	
23				—	6.15	5.50	5.46	
24[a]				—	6.76			6.55
25				—	6.12	5.99	5.65	
26				—	7.47	7.27	7.23	
27				—	7.70	7.28	6.98	
28[a]				—	6.12			6.94
29				—	5.03	5.56	7.12	

Table 5. Continuation

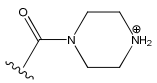
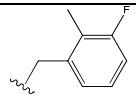
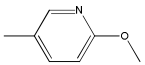
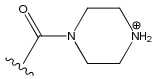
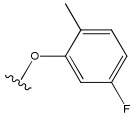
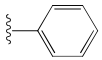
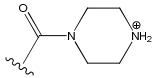
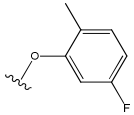
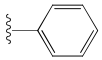
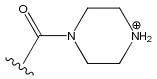
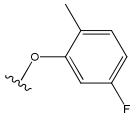
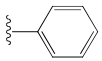
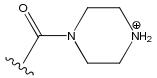
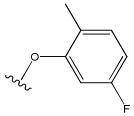
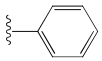
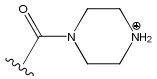
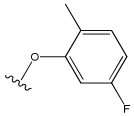
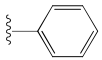
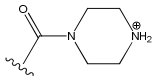
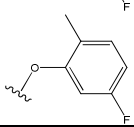
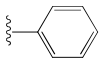
Cpd	R_1	R_2	R_3	R_4	pIC_{50}			
					Exp.	Calib.	LOOCV	Test
30				—	6.44	6.20	6.81	
31				6-OCH_3	7.62	7.78	7.79	
32 [a]				5-OCH_3	5.87			8.13
33				5-OH	8.70	9.00	8.64	
34				5-$CONH_2$	8.22	8.29	7.46	
35 [a]				6-CO_2H	6.91			7.62
36				6-$OCH_2CH_2N(CH_3)_2$	7.54	7.65	8.08	

Table 6. Continuation

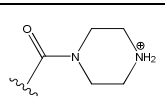
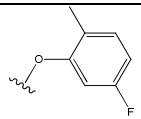
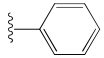
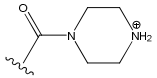
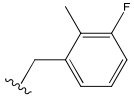
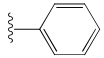
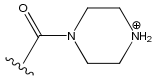
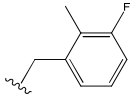
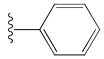
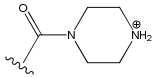
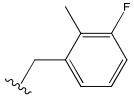
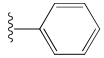
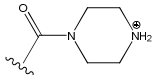
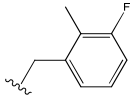
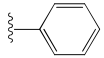
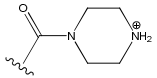
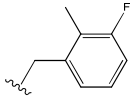
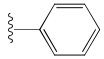
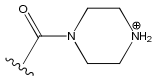
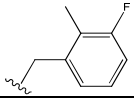
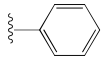
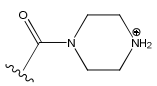
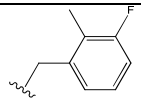
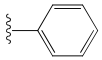
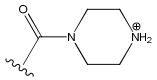
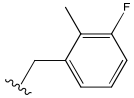
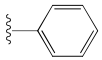
Cpd	R_1	R_2	R_3	R_4	pIC_{50}			
					Exp.	Calib.	LOOCV	Test
37				6-OCH_2CO_2H	6.89	7.09	7.69	
38				6-OCH_3	7.92	7.74	7.59	
39				6-OH	8.40	8.48	8.24	
40[a]				5-OH	8.40			8.87
41				4-CH_3, 5-OH	8.70	9.01	8.60	
42				7-F	7.08	5.57	7.57	
43				6-OCH_2CH_2OPh	7.62	7.39	7.04	

Table 7. Conclusion

Cpd	R_1	R_2	R_3	R_4	pIC_{50}			
					Exp.	Calib.	LOOCV	Test
44[a]				6-OCH_2CH_2OH	7.51			7.28
45				6-SO_2CH_3	7.36	7.47	7.65	

[a] Test set compounds

The **X** matrix was split into training (35 compounds, *ca.* 78%) and test (10 compounds, *ca.* 22%) set compounds (Table 1). These matrices were regressed against the **y** column vector (the pIC_{50} values, IC_{50} in mol L^{-1} as the 50% inhibitory concentration toward renin inhibitor in fluorogenic peptide substrate) using partial least squares (PLS) regression. The calibration performance was evaluated by means of RMSEC (root mean square error of calibration) and r^2, defined as $1 - ((\Sigma(y_i-\hat{y}_i)^2/\Sigma(y_i-\bar{y})^2)$, in which y_i corresponds to the experimental pIC_{50} values, $\hat{y}_i$ are the predicted pIC_{50} values, and $\bar{y}$ corresponds to the mean pIC_{50} values. Leave-one-out cross-validation (statistically evaluated using RMSECV and q^2, defined similarly as above) and external validation (statistically evaluated using RMSEP and r^2_{test}) were used to validate the QSAR model; also, Y-randomization tests (mean of 10 repetitions) were carried out to attest the model robustness (statistically evaluated using $RMSEC_{rand}$ and r^2_{rand}). Additional statistical parameters proposed by Roy et al. (2006) and Mitra et al. (2010), namely modified r^2 (r^2_m) and corrected penalized r^2 ($^cr^2_P$), were used for validation purposes. While $r^2_m \geq 0.5$ guarantees that not only a good correlation between the experimental and predicted pIC_{50} values was obtained for the test set, but also that the absolute experimental and predicted pIC_{50} values are congruent, the $^cr^2_P$ parameter gives insight about the statistical difference between r^2 and r^2_{rand} (values ≥ 0.5 are acceptable). The modified r^2 is described as $r_m^2 = r^2 (1 - (r^2 - r_0^2)^{1/2})$, wherein r^2 and r_0^2 correspond to the squared correlation coefficient values between measured and predicted values for the test set with and without intercept, respectively, while the corrected penalized r^2 is described as $^cr^2_P = r(r^2 - r^2_{rand})^{1/2}$.

The activities of the proposed compounds were estimated using the calibration model built. These procedures were all performed using the Chemoface program (Nunes *et al.*, 2012) and the Matlab platform (Mathworks, 2007).

New indole-3-carboxamide derivatives were proposed based on the combination of substructures of the most active compounds within the series. Their bioactivities were estimated using the regression parameters obtained in the MIA-QSAR model and the results were validated comparing the predicted pIC_{50} with the docking scores obtained from receptor-based analyses. The crystal structure of 2-(3-fluoro-2-methylbenzyl)-4-methyl-1-phenyl-3-(piperazin-1-ylcarbonyl)-1H-indol-5-ol inside the renin active site was obtained from the Protein Data Bank (PDB code: 3OOT) (Scheiper *et al.* 2010) and used for docking and alignment of the proposed structures **A**–**I**. The calculation of the docking energies of the ligands inside the renin active site was performed using the software Molegro Virtual Docker® (MVD®) (Thomsen and Christensen, 2006). The MolDock scoring function was used to superimpose the flexible molecule onto a template molecule. The docking search algorithm used in MVD was based on optimization using guided differential evolution (da Cunha *et al.*, 2013), which combines the differential evolution optimization technique with a cavity prediction algorithm during the search process. The active site exploited in docking studies was defined as a subset region of 10.0 Å from the center of the ligand. The interaction modes between the ligand and renin active site were determined as the highest energy score in the protein-ligand complex used during docking, and the conformers of each compound were mostly associated with bioactive conformations of 2-(3-fluoro-2-methylbenzyl)-4-methyl-1-phenyl-3-(piperazin-1-ylcarbonyl)-1H-indol-5-ol (Ramalho *et al.*, 2010).

The pharmacokinetic profiles of the proposed drug candidates were evaluated using calculated data for the Lipinski´s rule of five and other direct/indirect parameters related to the molecular polar surface area, human intestinal absorption, solubility (in buffer with pH 6.5) and plasma protein binding. These data were achieved using the Molinspiration program (www.molinspiration.com) and the Percepta Module of the ACD/Labs program.

3. Results and Discussion

The optimum number of latent variables (PLS components) used for further analysis was four, given the decay in the root mean square error of cross-validation (RMSECV) as a function of the number of latent variables. The calibration gives insight about the modeling ability of the MIA descriptors and it was performed over 35 out of 45 compounds; the r^2 of 0.863 (RMSEC = 0.40) attested the high performance of the model (illustrated in Figure 3), which is suggested to be achieved when $r^2 > 0.8$. To guarantee that this value was not due to chance correlation, a robustness test based on randomization of the **y** block and subsequent regression with the intact **X** matrix using 4 PLS components gave an r^2_{rand} of 0.590 (mean of 10 repetitions). In order to show that r^2 and r^2_{rand} are statistically distinct, the ${}^{c}r^2_{p}$ parameter (Roy *et al.*, 2009) was calculated to be 0.485 (only marginally inferior to the recommended value ≥ 0.50); then, the acceptable result of calibration is neither a fortuitous correlation nor over fitted.

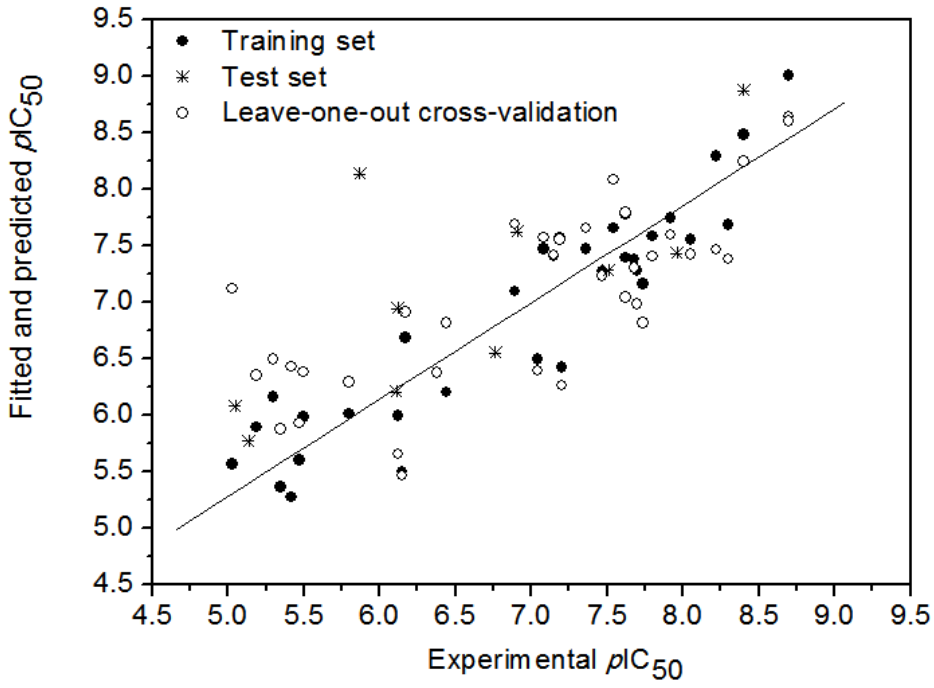


Figure 3 Plot of experimental vs. fitted/predicted pIC_{50} using the MIA-QSAR model for the series of indole-3-carboxamide derivatives as renin inhibitors

Model validation is required to attest the reliability and prediction performance of the QSAR model. The leave-one-out procedure was firstly performed for cross-validation, giving q^2 of 0.567 (RMSECV = 0.72), which is satisfactory, giving the recommended value of > 0.5. However, external validation is assumed to be the only way to establish a reliable QSAR model (Golbraikh and Tropsha, 2002); therefore, 10 samples randomly chosen from the overall 45 compounds (attempting to avoid border pIC_{50} values and having broad range of bioactivities) were used and the prediction of their pIC_{50} values fitted very well with the experimental ones, *i.e.* r^2_{test} = 0.519 (RMSEP = 0.91). This result is strongly affected by compound 32, whose residual in the predicted

pIC_{50} compared to the experimental value was significantly high (2.26). Thus, this compound was considered as an outlier and, after its exclusion, r^2_{test} enhanced to 0.788 (RMSEP = 0.60). To prove that not only a good correlation between the experimental and predicted pIC_{50} values was obtained, but also that the absolute experimental and predicted pIC_{50} values are consistent, an r^2_m calculation was performed according to described in the literature (Roy *et al.*, 2006; Ojha *et al.*, 2011), resulting in the 0.482 value, which is only marginally inferior to the recommended ≥ 0.5. The use of the recently developed aug-MIA-QSAR method (Nunes and freitas, 2013) to model the activities of this series of compounds did not give improved statistical results when compared to the traditional MIA-QSAR. Because ${}^{C}r^2_P$ and r^2_m were slightly worse than the recommended values, further validation using comparison with docking studies will be useful to attest the reliability of the MIA-QSAR model. This will be shown further, while new drug candidates can be proposed and their bioactivities predicted using the MIA-QSAR model.

Novel indole-3-carboxamide derivatives with expected high biological activities can be designed by mixing substructures of two or more highly active compounds of different series, keeping the common scaffold unaltered. This combination has shown to be synergistic in a variety of studies (da Mota *et al.*, 2014; Guimarães *et al.*, 2014; Silva *et al.*, 2012; Deeb *et al.*, 2012). Accordingly, the most active compounds of each series are **33**, **39**, **40** and **41**, whose substructures were combined to give the proposed compounds **A-I** of Figure 4. It is worth mentioning that the pharmacophoric group should contain a piperazin-1-yl-methanone moiety as R_1 and a phenyl ring as R_3, while the aromatic rings of the indolyl and R_2 (either as -OAr or -CH_2Ar groups) moieties should be substituted by 5- or 6-OH and/or 4-CH_3 groups in the former and 3- or 5-F and 2-CH_3 groups in the later. Thus, these groups capable of performing dipolar, hydrophobic and hydrogen bond interactions with the active site of renin (further

analyzed using docking studies) should play the determinant role for the bioactivities of the proposed molecules.

The pIC_{50} values of these indole-3-carboxamide derivatives were predicted using the MIA-QSAR model, and compounds **A**, **B**, **D**, **E** and **F** exhibited pIC_{50} values higher than the largest value found within the experimental data, indicating the synergistic effect mentioned above. The chemical structure of the most active, predicted compound **B** compared to the other purposes suggests that the effect of R_2 is independent on the group bridging the indolyl moiety with the phenyl ring, the pattern substitution in the later makes difference, with a 2-CH_3,5-F di-substitution preferred over both monosubstitution and 2-CH_3,3-F di-substitution. In addition, di-substitution in the indolyl benzene ring at the 4-CH_3 and 5-OH positions is preferred over monosubstitution and 4-CH_3 and 6-OH di-substitution.

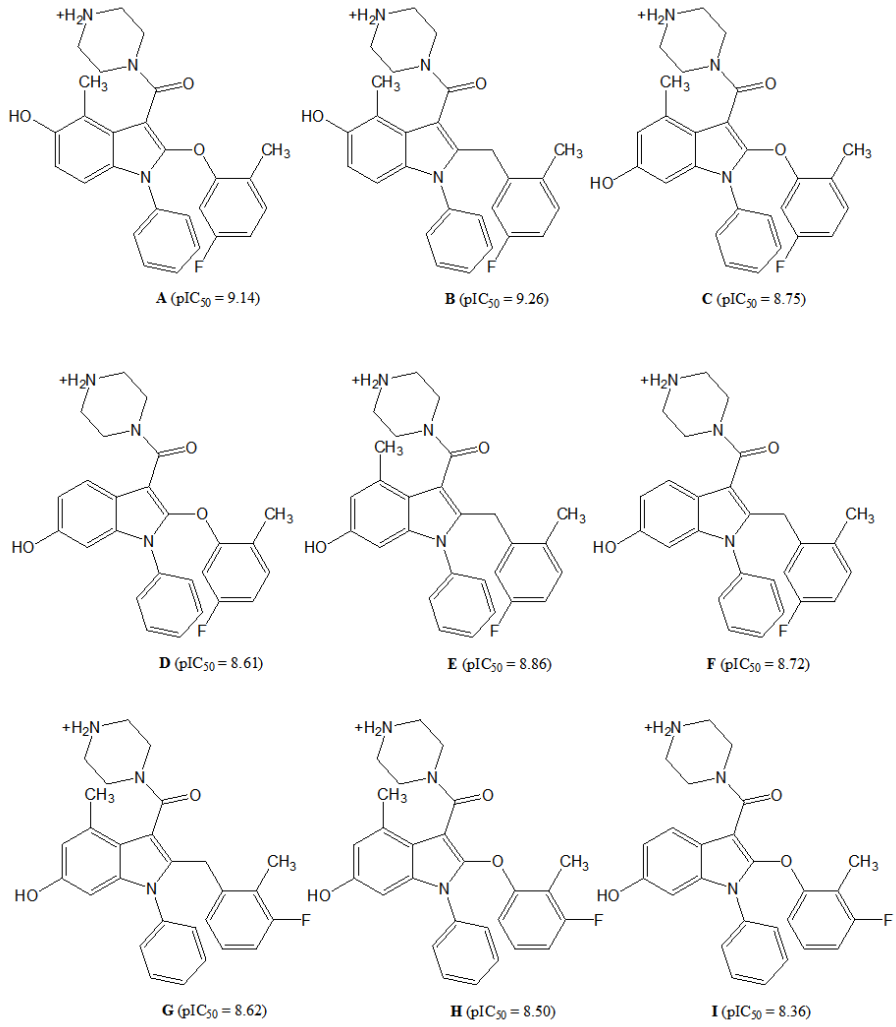


Figure 4 The proposed compounds A-I (the predicted pIC_{50} are given in parenthesis)

In order to attest if these QSAR predictions are really reliable, docking studies were performed to compare the docking scores with the predicted pIC_{50} for **A-I**. Initially, we have performed a re-docking calculation. The ligand binding process was validated by "re-docking" of the co-crystallized 2-(3-

fluoro-2-methylbenzyl)-4-methyl-1-phenyl-3-(piperazin-1-ylcarbonyl)-1H-indol-5-ol inside the renin active site. In docking simulations, we expect that the best results generate RMSD values below 2.0 Å, when compared to crystallographic structures. We verified that the ligand became bound in the renin crystal structure, similarly to preexisting co-crystallized ligand (RMSD = 0.54). After that, the docking scores was compared with the experimental pIC_{50} of selected compounds with low, moderate and high activities (Figure 5a, containing data for compounds **12**, **15**, **20**, **26**, **29**, **33**, **39**, **40** and **41**). The determination coefficient of 0.706 guarantees that pIC_{50} values is well reproduced by the docking scores and, therefore, this receptor-based approach can be used to validate the MIA-QSAR predictions for the proposed compounds, as well as to understand the interactions governing their biological activities. The correlation between the predicted pIC_{50} and docking scores for **A**-**I** indicate that either the first data were underestimated or the docking scores are overestimated for **C**, **G** and **H**, since they shift from the linear relationship (R^2 = 0.884) observed in Figure 5b for the remaining proposed compounds. Nevertheless, the predicted pIC_{50} for **A**, **B**, **D**, **E**, **F** and **I** were validated by the docking methodology, confirming that **A** and especially **B** are promising direct renin inhibitors. Figure 6 shows **A** and **B** interacting with the amino acid residues through hydrogen bond, which is an important descriptor for the binding mode and activity of compounds in an active site.

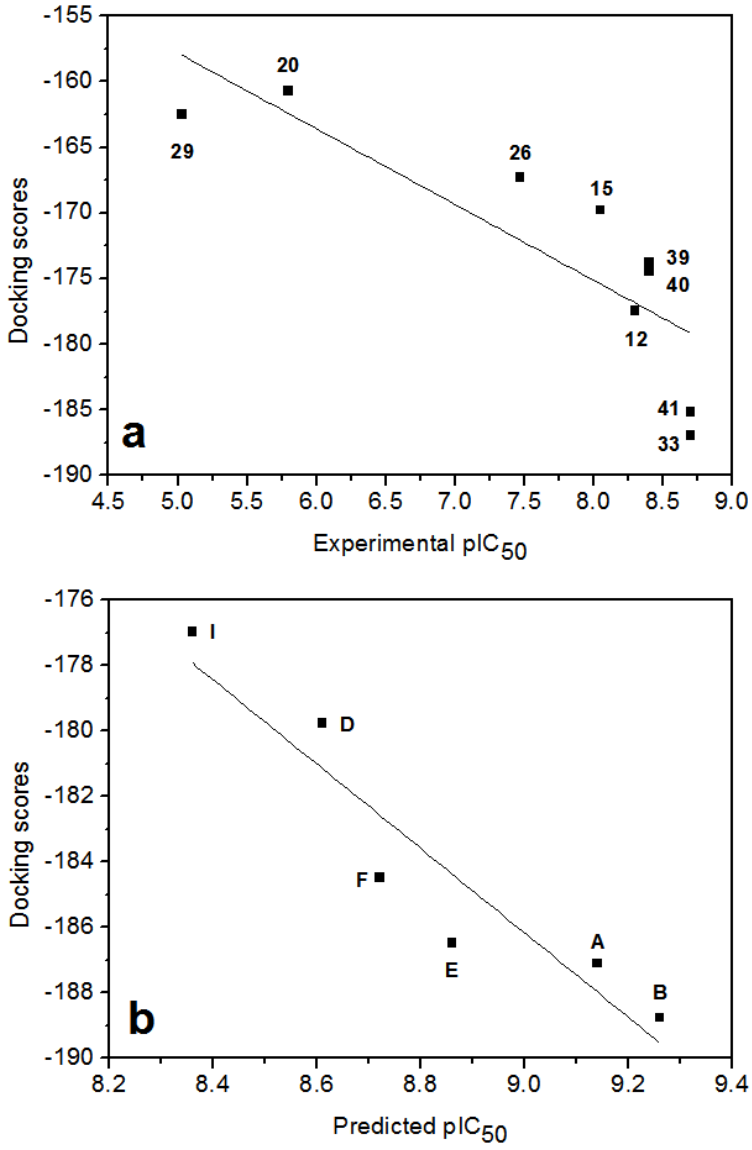


Figure 5 Relationship between docking scores and a) experimental and b) predicted pIC_{50}. The docking score obtained for Aliskiren is -130.5 kcal mol^{-1}

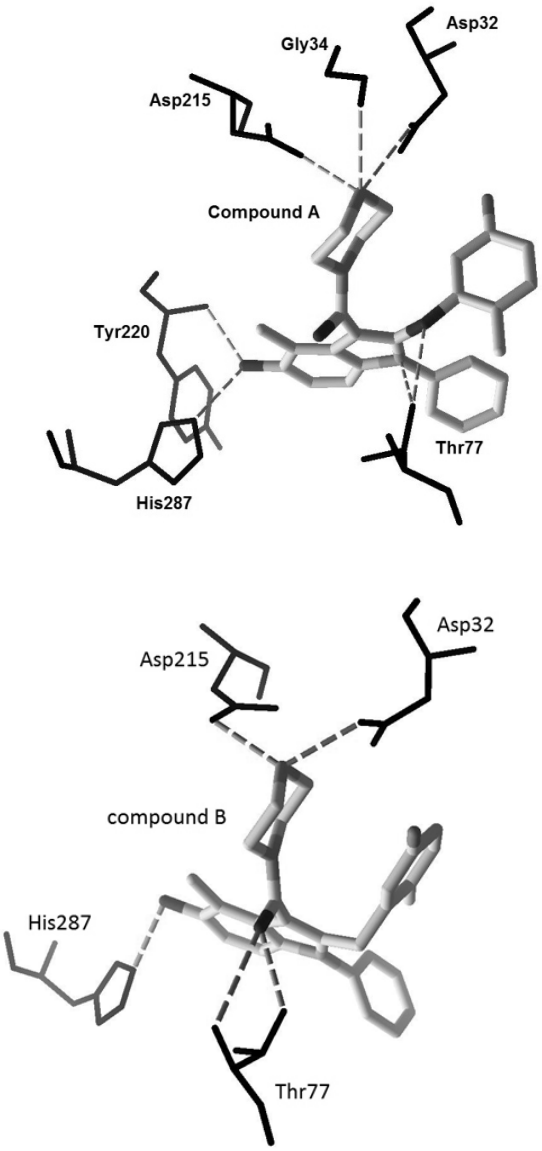


Figure 6 Compounds **A** and **B** docked inside the renin active site

A drug must have good ADMET (absorption, distribution, metabolism, excretion and toxicity) profile in addition to high bioactivity. Some of these

pharmacokinetic parameters can be calculated to give insight about the drug-likeness of the proposed compounds in comparison with the commercially available Aliskiren and some experimental drug candidates of Table 1. The Lipinski´s rule of five (Lipinski *et al.*, 1997) states that most "drug-like" molecules have log*P* ≤ 5, molecular weight ≤ 500, number of hydrogen bond acceptors (n_{ON}) ≤ 10, and number of hydrogen bond donors (n_{OHNH}) ≤ 5; molecules that violate more than one of these rules may have problems with bioavailability. These parameters, together with the molecular polar surface area (TPSA, related to oral bioavailability), were assessed using the Molinspiration program (www.molinspiration.com) (Table 2). The results are encouraging for the proposed compounds, since no rule of five was violated for any of them, while Aliskiren has an exceedingly high molecular volume and number of hydrogen bond acceptors. In addition, according to the criteria established by the Percepta program (≤ 0.01 mg mL^{-1} highly insoluble, 0.01 to 0.10 insoluble, ≥ 0.10 soluble), compound **B** would be soluble at pH 6.5 even in the neutral form (the protonated form is not possible to be calculated), while the human intestinal absorption (HIA) is 100% and the plasma protein binding (PPB) is similar to Aliskiren.

Table 8 Calculated parameters of the Lipinski´s rule of five, total polar surface area (TPSA) and computed pharmacokinetic parameters for the proposed compounds (A-I), for reference compounds and Alinkiren.[a]

Cpd	log*P*	MW	n_{ON}	n_{OHNH}	TPSA	Violations	HIA (%)	PPB (%)	Solub.
Aliskiren	4.15	553.7	10	7	166.4	2	95	86	10.88
33	0.94	446.5	6	3	71.3	0	100	97	0.06
39	1.24	444.5	5	3	62.1	0	100	98	0.01
40	1.24	444.5	5	3	62.1	0	100	98	0.07
41	2.09	458.5	5	3	62.1	0	100	97	0.04
A	1.79	460.5	6	3	71.3	0	100	97	0.08
B	2.11	458.5	5	3	62.1	0	100	97	0.18
C	1.32	460.5	6	3	71.3	0	100	97	0
D	0.94	446.5	6	3	71.3	0	100	97	0.01
E	1.64	458.5	5	3	62.1	0	100	97	0
F	1.27	444.5	5	3	62.1	0	100	98	0.01
G	1.62	458.5	5	3	62.1	0	100	97	0
H	1.30	460.5	6	3	71.3	0	100	97	0
I	0.92	446.5	6	3	71.3	0	100	97	0.01

[a] HIA, PPB and solubility (mol mL^{-1}) were calculated for the neutral, deprotonated indole-3-carboxamide derivatives

4. Conclusions

Our QSAR model was reliable in predicting the activities of new indole-3-carboxamide derivatives as direct renin inhibitors, with the support of docking studies. Compound **B** was found the be the most promising drug candidate, because of its high predicted activity and good calculated pharmacokinetic profile, related to ADME. Di-substitution at the benzene ring of the indolyl moiety with 4-CH_3 and 5-OH groups, together with di-substitution at the benzene ring in R_2 with 2-CH_3 and 5-F appear to be relevant when designing corresponding analogues for further synthesis. Hydrogen bonds involving some of these groups with amino acid residues of the active site of renin should play a

significant role for the intermolecular enzyme-substrate interaction governing the biological activities of this class of compounds.

Acknowledgements

Authors are thankful to FAPEMIG for the financial support of this research, as well as to CAPES for the studentship (E.G.M.) and to CNPq for fellowships (to E.F.F.C. and M.P.F.).

References

ACD/ChemSketch Version 10.02, Advanced Chemistry Development, Inc., Toronto, 2006.

Cormanich RA, Freitas MP, Rittner R (2011) 2D Chemical drawings correlate to bioactivities: MIA-QSAR modelling of antimalarial activities of 2,5-diaminobenzophenone derivatives. J Braz Chem Soc 22:637-642.

Corminboeuf O, Bezencon O, Grisostomi C, Remen L, Richard-Bildstein S, Bur D, Prade L, Hess P, Strickner P, Fischli W, Steiner B, Treiber A (2010) Design and optimization of new piperidines as renin inhibitors. Bioorg Med Chem Lett 20:6286–6290.

da Cunha EFF, Resende JE, França TCC, Gonçalves MA, Souza FR, Santos-Garcia L, Ramalho TC (2013) Molecular modeling studies of piperidine derivatives as new acetylcholinesterase inhibitors against neurodegenerative diseases. J Chem 27872.

da Mota EG, Silva DG, Guimarães MC, da Cunha EFF, Freitas MP (2014) Computer-assisted design of novel 1,4-dihydropyridine calcium channel blockers. Mol Simul. 40:959-965.

Deeb O, da Cunha EFF, Cormanich RA, Ramalho TC, Freitas MP (2012) Computer-assisted assessment of potentially useful non-peptide HIV-1 protease inhibitors. Chemom Intell Lab Sys 116:123-127.

Freitas MP (2007) Multivariate QSAR: From classical descriptors to new perspectives. Curr Comp-Aid Drug Des 3:235-239.

Golbraikh A, Tropsha A (2002) Beware of q^2! J Mol Graph Model 20:269-276.

Guimarães MC, Silva DG, da Mota EG, da Cunha EFF, Freitas MP (2014) Computer-assisted design of dual-target anti-HIV-1 compounds. Med Chem Res 23:1548-1558.

Jing T, Feng J, Zuo Y, Ran B, Liu J, He G (2012) Exploring the substructural space of indole-3-carboxamide derivatives binding to renin: a novel active-site spatial partitioning approach. J Mol Model 18:4417–4426.

Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 23:3-25.

Matlab 7.5, Mathworks Inc., Natick, 2007.

Mitra I, Saha A, Roy K (2010) Exploring quantitative structure-activity relationship studies of antioxidant phenolic compounds obtained from traditional Chinese medicinal plants. Mol Simul 36:1067-1079.

Nunes CA, Freitas MP (2013) Introducing new dimensions in MIA-QSAR: a case for chemokine receptor inhibitors. Eur J Med Chem 62:297-300.

Nunes CA, Freitas MP, Pinheiro ACM, Bastos SC (2012) Chemoface: A novel free user-friendly interface for chemometrics. J Braz Chem Soc 23:2003-2010.

Ojha PK, Mitra I, Das RN, Roy K (2011) Further exploring r^2_m metrics for validation of QSPR models. Chemom Intell Lab Sys 107:194-205.

Paul M, Poyan MA, Kreutz R (2006) Physiology of local renin–angiotensin systems. Physiol Rev 86:747–803.

Ramalho TC, França TCC, Rennó MN, Guimarães AP, da Cunha EFF, Kuca K (2010) Development of new acetylcholinesterase reactivators: Molecular modeling *versus in vitro* data. Chem-Biol Inter 187:436–440.

Roy K, Mitra I, Kar S, Ojha PK, Das RN, Kabir H (2006) Comparative studies on some metrics for external validation of QSPR models. J Chem Inf Model 52-396-408.

Roy PP, Paul S, Mitra I, Roy K (2009) On two novel parameters for validation of predictive QSAR models. Molecules 14:1660-1701.

Scheiper B, Matter H, Steinhagen H, Stilz U, Böcskei Z, Fleury V, McCort G (2010) Discovery and optimization of a new class of potent and non-chiral indole-3-carboxamide-based renin inhibitors. Bioorg Med Chem Lett 20:6268-6272.

Silva DG, Freitas MP, da Cunha EFF, Ramalho TC, Nunes CA (2012) Rational design of small modified peptides as ACE inhibitors. Med Chem Commun 3:1290-1293.

Scheiper B, Matter H, Steinhagen H, Stilz U, Bocskei UZ, Fleury V, McCort G (2010) Discovery and optimization of a new class of potent and non-chiral indole-3-carboxamide-based renin inhibitors. Bioorg Med Chem Lett 20:6268–6272.

Thomsen R, Christensen M H (2006) MolDock: A new technique for high-accuracy molecular docking. J Med Chem 49:3315–3321.

Tzoupis H, Leonis G, Megariotis G, Supuran CT, Mavromoustakos T, Papadopoulos MG (2012) Dual inhibitors for aspartic proteases HIV-1 PR and renin: Advancements in AIDS-hypertension-diabetes linkage via molecular docking, inhibition assays, and binding free energy calculations. J Med Chem 55:5784-5796.

Wood JM, Stanton JL, Hofbauer KG (1987) Inhibitors of renin as potential therapeutic agents. J Enz Inhib Med Chem 1:169–185.

ARTIGO 2

COMPUTER-ASSISTED DESIGN OF NOVEL 1,4-DIHYDROPYRIDINE CALCIUM CHANNEL BLOCKERS

Artigo redigido conforme normas da revista Molecular Simulation

Estella G. da Mota, Daniel G. Silva, Maria C. Guimarães, Elaine F. F. da Cunha, Matheus P. Freitas

Abstract

The rational design of novel 1,4-dihydropyridine calcium channel blockers (CCB) is reported in this work. First, a QSAR modeling was carried out to predict the pIC_{50} of new drug-like compounds, which are miscellany of the substructures of the most active molecules within the data set. Subsequently, the descriptors based on multivariate image analysis were analyzed using principal component analysis (PCA) for pattern recognition, *i.e.* to classify drug-like compounds into three levels of efficacy according to the substituents. The QSAR estimations were validated by docking studies, which have not been performed extensively for CCB in current studies in the field, and insights about ADME properties are given.

Keywords: calcium channel blockers; hypertension; 1,4-dihydropyridine derivatives; QSAR; docking studies.

1. . Introduction

Cardiovascular diseases are the main cause of death worldwide according to the World Health Organization (www.who.int, accessed in July 20th, 2013). There is not cure for hypertension, but it can be controlled by healthy lifestyle changes and by the use of antihypertensives. Calcium channel blockers (CCBs) have widespread use as antihypertensives, like the well-known amlodipine and felodipine (Figure 1), which are some 1,4-dihydropyridine (DHP) derivatives. They act by blocking the movement of calcium (Ca^{2+}) through calcium channels, *i.e.* CCBs work by blocking voltage-gated calcium channels in cardiac muscle and blood vessels. This decreases intracellular calcium, leading to a reduction in cardiac and vascular contraction (vasodilation) and, consequently, to an increase in arterial diameter. Vasodilation decreases total peripheral resistance, while a decrease in cardiac contractility decreases cardiac output. Since blood pressure is determined by cardiac output and peripheral resistance, blood pressure drops [1]. The L-type CCBs are assumed to be responsible for the regulation of smooth and cardiac muscle contraction [2-4]. The access of DHPs to L-type CCBs has been studied by several groups [5-7] and the consensus is that DHPs reach their binding site within the pore-forming α_1-subunit from the extracellular side.

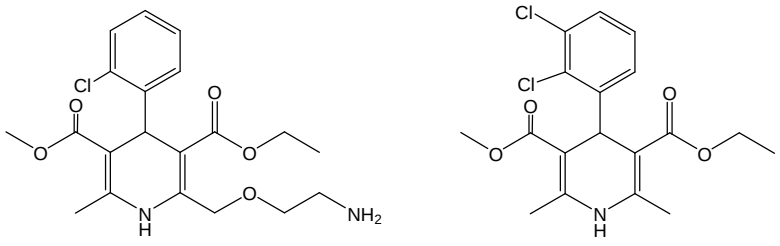


Figure 1 Common DHPs used as CCB for the treatment of hypertension

A series of DHPs has been synthesized [8-11] and some of them have shown to be more active than amlodipine and felodipine (pIC_{50} of 6.6 and 8.3, respectively [12]; IC_{50} in M). Moreover, a QSAR analysis for this series of compounds provided predictive models based on a pool of physicochemical descriptors [13], while many properties were found to be important for calcium channel antagonist activity, namely electronic features of some groups, dipole moment, molar refractivity, surface area, electrophilicity and electronegativity. In this work, we aimed at systematizing substituent profiles of this series of compounds using principal component analysis and building a QSAR model capable of estimating accurately the bioactivities of new DHPs.

In order to achieve this, a multivariate image analysis applied to QSAR (MIA-QSAR) was performed, and the descriptors were regressed against the corresponding biological data using partial least squares (PLS) regression. In MIA-QSAR [14], descriptors are pixels (binaries) of images, corresponding to the two-dimensional chemical structures as wireframes; the variable moiety of the whole set of chemical structures (usually the substituents) explains the variance in the bioactivities block, while the congruent substructure of the series is used for 2D alignment. The QSAR model was applied to predict the antagonist activity toward L-type calcium channel of new, proposed DHPs. To validate these predictions, docking studies were carried out to accompany the trends in calculated pIC_{50} in terms of docking scores (the energy of interaction between ligand and enzyme).

2. Materials and Methods

The compound classification using principal component analysis (PCA) and the QSAR model were accomplished using multivariate image analysis (MIA), in which descriptors correspond to pixels (binaries) of 2D chemical structures drawn with the aid of the ChemSketch program [15]. Then, each

image was saved as bitmaps (.bmp files), a 2D array in which the variable pixels (substituents) correspond to the variable portion of the molecules; these changes explain the variance in the activities block (the **y** column vector in the regression analysis). The pixels were converted to numerical values using the Matlab [16] platform and the superposed images of Figure 2 illustrate the variance in the chemical drawings.

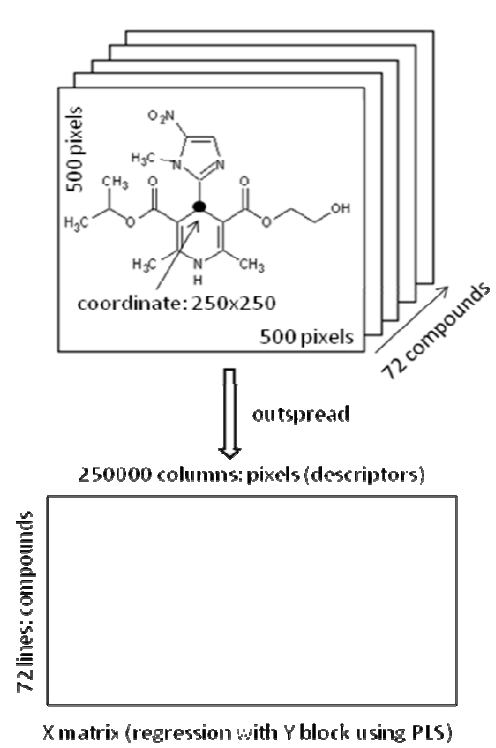
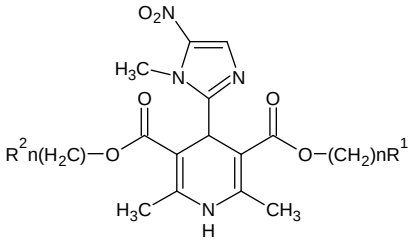


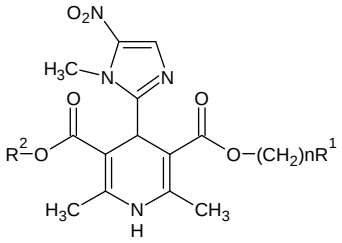
Figure 2 Construction of the three-dimensional arrangement and unfolding of the 3-way array to give the X matrix. The arrow in the structure indicates the pixel used for the 2D alignment

A series of 72 DHP derivatives was taken from the literature [13] (see data set in the (Table 1) and converted into images of 500×500 pixels size each, according to described above and elsewhere [14,17,18].

Table 1 Series of calcium channel blockers used in the MIA-QSAR modeling and the experimental and fitted/predicted values of pIC_{50}: (A) symmetrical and (B) unsymmetrical compounds



(A)



(B)

Compound	R^1	R^2	N	Exp.	Calib.	LOOCV	Test
1	n-Butyl	Ethyl	0	11.49			10.37
2	n-Butyl	n-Butyl	0	11.29	11.28	10.09	
3	OCH_3	i-Propyl	2	12.21	12.29	11.64	
4	OH	Methyl	2	11.89	11.53	9.94	
5	$COCH_3$	Methyl	3	11.83	12.18	11.57	
6	CN	i-Propyl	2	11.83	11.63	11.07	
7	OH	i-Propyl	2	11.53	11.47	10.85	
8[a]	OH	Ethyl	2	11.45			10.80
9	OCH_3	Methyl	2	11.38	11.36	10.95	
10[a]	OCH_3	Ethyl	2	11.13			11.48
11	c-hexyl	Methyl	0	9.01	8.99	9.18	
12	c-hexyl	Ethyl	0	8.40	8.29	8.38	
13	c-hexyl	Methyl	1	8.57	8.55	8.63	
14[a]	c-hexyl	Ethyl	1	8.26			8.94
15	c-hexyl	Methyl	2	8.40	8.32	9.07	
16[a]	c-hexyl	Ethyl	2	7.82			8.12
17	c-hexyl	Methyl	3	8.33	8.36	8.42	
18	c-hexyl	Ethyl	3	8.28	8.40	8.82	
19	c-hexyl	Methyl	4	8.30	8.41	9.83	
20[a]	c-hexyl	Ethyl	4	8.22			8.20
21	c-Pentyl	Methyl	3	8.73	8.58	8.50	
22	c-Pentyl	Ethyl	3	8.40	8.64	8.94	
23	c-Propyl	Methyl	1	7.48	7.69	9.75	
24	c-Propyl	Ethyl	1	7.13	7.40	8.30	
25	Phenyl	Methyl	1	7.70	7.73	8.48	
26[a]	Phenyl	Ethyl	1	7.70			7.85

Table 2, Continuation

Compound	R^1	R^2	N	Exp.	Calib.	LOOCV	Test
27	Phenyl	Methyl	2	7.70	7.59	7.79	
28	Phenyl	Ethyl	2	7.56	7.66	7.89	
29	p-tolyl	Methyl	2	7.90	7.68	7.35	
30	p-tolyl	Ethyl	2	7.32	7.32	7.78	
31	Phenyl	Methyl	3	7.61	7.55	7.48	
32[a]	Phenyl	Ethyl	3	7.42			7.55
33	Phenyl	Methyl	4	7.57	7.45	7.30	
34	Phenyl	Ethyl	4	7.21	7.29	7.31	
35	Phenyl	Methyl	5	7.14	7.13	7.10	
36	Phenyl	Ethyl	5	7.08	7.02	7.49	
37	c-hexyl	c-hexyl	0	7.85	7.87	9.18	
38[a]	c-hexyl	c-hexyl	1	9.31			8.53
39	c-hexyl	c-hexyl	2	8.88	8.84	8.45	
40	c-hexyl	c-hexyl	3	7.55	7.50	7.35	
41	n-Propyl	n-Propyl	0	10.05	10.61	11.45	
42	c-Propyl	c-Propyl	1	7.57	7.41	7.14	
43	c-Pentyl	c-Pentyl	3	7.89	7.88	8.48	
44[a]	Phenyl	Phenyl	1	8.58			7.76
45	Phenyl	Phenyl	2	8.42	8.24	7.69	
46	p-tolyl	p-tolyl	2	7.28	7.17	7.70	
47	Phenyl	Phenyl	3	7.35	7.45	7.22	
48	Phenyl	Phenyl	4	6.89	6.97	7.49	
49	Phenyl	Phenyl	5	5.95	5.95	7.03	
50	i-Butyl	i-Butyl	0	11.36	11.45	10.65	
51	ONO_2	Methyl	2	11.59	11.61	11.29	
52[a]	ONO_2	Ethyl	2	10.93			11.31
53	ONO_2	i-Propyl	2	12.02	11.99	11.41	
54	ONO_2	Methyl	3	11.33	10.84	10.21	
55	ONO_2	Ethyl	3	11.11	11.30	11.59	
56	ONO_2	i-Propyl	3	11.75	11.98	12.16	
57	ONO_2	Methyl	4	11.60	11.82	9.95	
58[a]	ONO_2	Ethyl	4	10.87			11.11
59	$CH(CH_2ONO_2)_2$	Methyl	0	11.41	11.52	11.56	
60	$CH(CH_2ONO_2)_2$	Ethyl	0	11.66	11.16	10.72	
61	$N(CH_3)_2$	i-Propyl	0	11.61	11.74	12.08	
62	$N(CH_3)_2$	Methyl	3	8.14	8.52	9.73	
63[a]	$N(CH_3)_2$	Ethyl	3	8.96			8.77
64	$N(CH_3)_2$	i-Propyl	3	9.33	8.80	9.01	
65	$N(CH_3)_2$	Methyl	2	8.50	9.03	9.72	

Table 3, Conclusion

Compound	R^1	R^2	N	Exp.	Calib.	LOOCV	Test
66	$N(CH_3)_2$	Ethyl	2	9.31	9.34	9.66	
67	$N(CH_3)_2$	i-Propyl	2	10.13	9.75	8.84	
68[a]	t-Butyl	Methyl	0	9.74			10.05
69	t-Butyl	Ethyl	0	10.31	10.39	10.36	
70	n-Pentyl	Ethyl	0	12.13	12.13	11.45	
71	Methyl	Ethyl	0	11.13	11.07	9.90	
72	n-Pentyl	Methyl	0	12.33	12.07	11.47	

[a] Test set compounds

The images were grouped to give a tree-way array of 72×500×500 dimension, which was subsequently unfolded to a 72×250000 matrix. Columns with zero variance (corresponding to blank parts or congruent substructures in the workspace) were removed to minimize computing costs, giving a 72×4634 matrix, the **X** matrix. These data were mean-centered and PCA was carried out using the Pirouette program [19]. Samples were classified as highly active (pIC_{50} > 11), moderately active (8 < pIC_{50} < 11), and poorly active (pIC_{50} < 8). New compounds obtained from the combination of substructures of the most active compounds within the series were proposed and then introduced in PCA to evaluate in which class they would fall within. Since the chemical structures have a C_2 symmetry, the classification capability of R_1 and R_2 groups were analyzed individually, in order to find out whether bioactivity has any relationship with a given half of the molecule; to achieve this, the images were taken only partially (an image corresponding only to the R_1 part, and another corresponding only to the R_2 moiety), and PCA was carried out for both data, named $\mathbf{X_{R1}}$ and $\mathbf{X_{R2}}$ matrices.

The **X** matrix, as well as the partial $\mathbf{X_{R1}}$ and $\mathbf{X_{R2}}$ matrices, were randomly split into training (58 compounds, *ca.* 80%) and test (14 compounds, *ca.* 20%) set compounds (Table 1). These matrices were regressed against the **y**

column vector (the pIC_{50} values, IC_{50} in mol L^{-1} as the 50% inhibitory concentration toward calcium channel in Guinea-pig Ileum) using partial least squares (PLS) regression. The calibration performance was evaluated by means of RMSEC (root mean square error of calibration) and r^2, defined as $1 - [(\Sigma(y_i - \hat{y}_i)^2/\Sigma(y_i - \bar{y})^2]$, in which y_i corresponds to the experimental pIC_{50} values, $\hat{y}_i$ are the predicted pIC_{50} values, and $\bar{y}$ corresponds to the mean pIC_{50} values. The model was validated through leave-one-out cross-validation (statistically evaluated using RMSECV and q^2, defined similarly as above) and external validation (statistically evaluated using RMSEP and r^2_{test}); also, Y-randomization tests (mean of 10 repetitions) were carried out to attest the model robustness (statistically evaluated using $RMSEC_{rand}$ and r^2_{rand}). Additional statistical parameters proposed by Roy et al. [20-23], namely modified r^2 (r^2_m) and corrected penalized r^2 ($^cr^2_P$), were used for validation purposes. While $r^2_m \geq 0.5$ guarantees that not only a good correlation between the experimental and predicted pIC_{50} values is obtained for the test set, but also that the absolute experimental and predicted pIC_{50} values are congruent, the $^cr^2_P$ parameter gives insight about the statistical difference between r^2 and r^2_{rand} (values ≥ 0.5 are acceptable). The activities of the proposed compounds were estimated using the calibration model built. These procedures were all performed using the Chemoface program [24] and Matlab platform [16].

The L-type calcium channel structure from the human cardiac α1 subunit (code CAC1C_HUMAN in the SWISS-PROT database) was modeled in this study. α1 Subunit consists of four repeating motifs, each motif comprising six segments. Recent models of calcium channel [25,26] are based on different alignments of the pore-forming S6 segments with the M2 segment of KcsA. In this study, the alignment scheme was based on that proposed by Zhorov *et al.* [27] (Figura 8). The X-ray structure coordinates of KcsA K+ channel were obtained from the Protein Data Bank (PDB code: 1BLB).

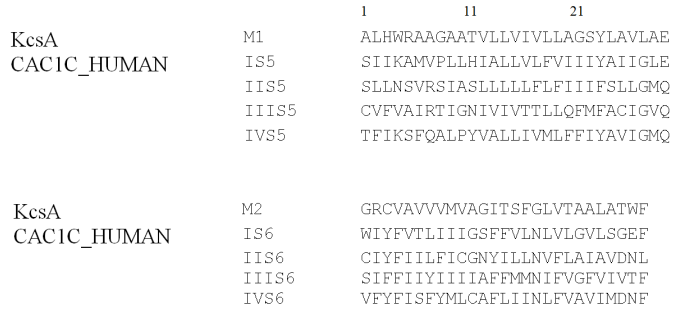


Figure 3 Alignment between KcsA and human calcium channel

Compounds **A-F** (Figure 4) and amlodipine were docked in the calcium channel binding site using the Molegro Virtual Docker (MVD) [28], a program for the prediction of the most likely conformation of how a ligand would bind to a macromolecule. Ligand and protein are considered flexible during the docking simulation. The MolDock scoring function used by MVD is derived from the piecewise linear potential (PLP), a simplified potential whose parameters are fit to protein–ligand structures and binding data scoring functions and further extended in Generic Evolutionary Method for molecular DOCK with a new hydrogen bonding term and new charge schemes. The docking scoring function, E_{score}, is defined by the following energy terms:

$$E_{\text{Score}} = E_{\text{Inter}} + E_{\text{Intra}}$$

where E_{inter} is the ligand–protein interaction energy:

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

The E_{PLP} term is a 'piecewise linear potential' using two different sets of parameters: one set for approximating the steric (Van der Waals) term between atoms, and another stronger potential for hydrogen bonds. The second term

describes the electrostatic interactions between charged atoms. It is a Coulomb potential with a distance-dependent dielectric constant given by: D(r) = 4r. The numerical value of 332.0 adjusts the units of the electrostatic energy to kilocalories per mole [28]. E_{intra} is the internal energy of the ligand.

In summary, these functions are used to automatically superimpose a flexible molecule onto a rigid template molecule. The docking search algorithm used in MVD is based on an evolutionary algorithm, where interactive optimization techniques inspired by Darwinian evolution theory, and a new hybrid search algorithm called guided differential evolution. The guided differential evolution algorithm combines the differential evolution optimization technique with a cavity prediction algorithm during the search process, which allows for a fast and accurate identification of potential binding modes (poses).

3. Results and Discussion

Principal component analysis (PCA) is an useful tool to classify samples with similar properties. PCA [29] creates *p* latent variables (*Y*) as linear combinations of the original *p* variables (*X*), in such a way that new orthogonal axes (principal components, PCs) are built to explain the maximum variance possible in only a few dimensions.

$$Y_i = e^T{}_i X = e_{i1}X_1 + e_{i2}X_2 + \ldots + e_{ip}X_p$$

where the unknown vector e_i establishes the *i*th linear combination, for $i = 1, \ldots, p$.

A biplot of PC3 × PC4 was capable of clustering most of compounds with high bioactivities (positive scores in PC3) from those with moderate and low bioactivities (Figure 3a). When only those images corresponding to the R_1 portion are analyzed ($\mathbf{X_{R1}}$ matrix), a well-defined group in PC1 × PC2 is

observed for the highly bioactive compounds, while PC5 (*eg.* in the biplot of PC4 × PC5) appears to be the component capable for grouping R_2 ($\mathbf{X_{R2}}$ matrix) responsible for the high activity of compounds. Principal component analyses over $\mathbf{X_{R1}}$ and $\mathbf{X_{R2}}$ (Figure 3b,c) indicate that most of R_1 substituents with high bioactivities contains electronegative atoms, while R_2 is preferentially a noncyclic hydrocarbon.

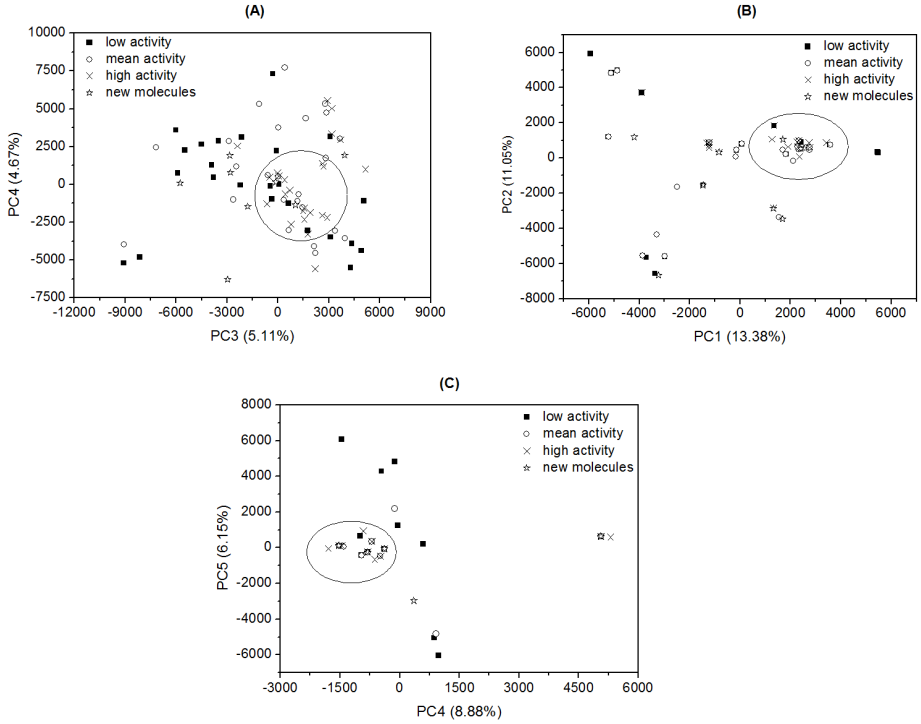


Figura 4 PCAs for the series of calcium channel blockers. (a) Considering the whole chemical structure of the series, most of high activity compounds are clustered as circled; (b) when considering the structural portion comprising the R_1 substituent only, most of the highly active compounds are clustered as indicated; (c) when considering the structural portion comprising the R_2 substituent only, the most active compounds are clustered mainly in negative scores in PC4, as indicated

Based on these chemical insights, a series of novel compounds were proposed (Figure 4) and the corresponding MIA descriptors were included in the PCA; the **E** compound falls within the cluster comprising those molecules with high antagonist effect toward calcium channels. Thus, a quantitative analysis can be performed using MIA-QSAR to estimate the pIC_{50} values of the proposed compounds.

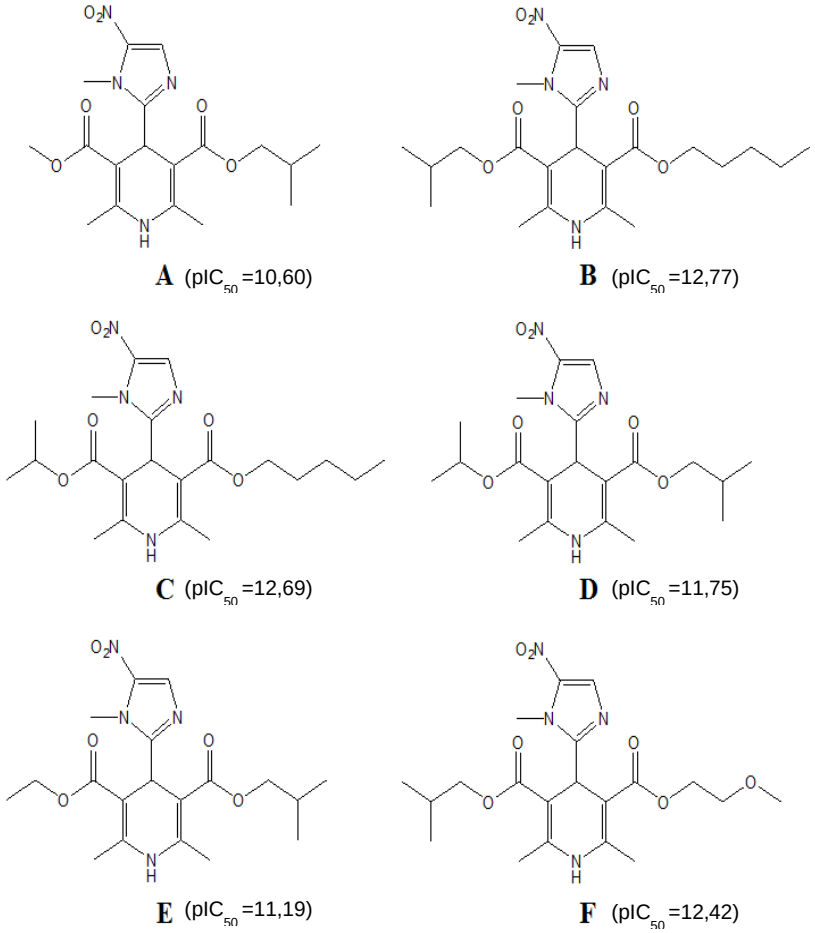


Figure 5 Series of compounds proposed

The optimum number of latent variables (PLS components) used for further analysis was seven, given the decay in the root mean square error of cross-validation (RMSECV) as a function of the number of latent variables. The calibration gives insight about the modeling ability of the MIA descriptors and it was performed over 58 out of 72 compounds; an r^2 of 0.99 (RMSEC = 0.214) attested the high performance of the model (illustrated in Figure 5 and), which is suggested to be achieved when $r^2 > 0.8$. To guarantee that this value was not due to chance correlation, a robustness test based on randomization of the **y** block and subsequent regression with the intact **X** matrix using 7 PLS components gave an r^2_{rand} of 0.74. In order to show that r^2 and r^2_{rand} are statistically very distinct, the ${}^{c}r^2_{p}$ parameter [20] was calculated to be 0.50; then, the good result of calibration is neither a fortuitous correlation nor over fitted.

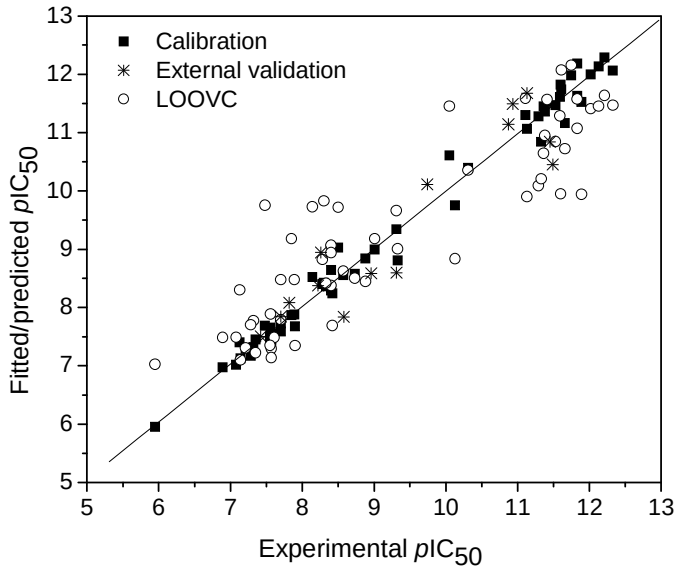


Figure 6 Plot of experimental *vs.* fitted/predicted pIC_{50} using MIA-QSAR, for the series of calcium channel blockers

Table 2 Statistical data for the MIA-QSAR model

Parameter	Calibration	LOOCV	Test set	Y-randomization[a]
r^2	0.99	0.80	0.86	0.74±0.02
RMSE	0.214	0.832	0.539	0.684±0.033

[a] mean of 10 repetitions; r^2_m= 0.81; [c]r^2_p= 0.50

Nevertheless, model validation is required to attest the reliability and prediction power of the QSAR model. A leave-one-out procedure was firstly performed for cross-validation, giving q^2 of 0.80 (RMSECV = 0.832), which is very satisfactory, giving the recommended value of > 0.5. However, external validation is assumed to be the only way to establish a reliable QSAR model [30]; therefore, 14 samples randomly chosen from the overall 72 compounds (attempting to avoid border pIC_{50} values and comprising compounds having broad range of bioactivities) were used and the prediction of their pIC_{50} values fitted very well with the experimental ones, *i.e.* r^2_{test} = 0.86 (RMSEP = 0.539). To prove that not only a good correlation between the experimental and predicted pIC_{50} values was obtained, but also that the absolute experimental and predicted pIC_{50} values are consistent, an r^2_m calculation was performed according to described in the literature [21-23], resulting in the 0.81 value. Thus, the MIA-QSAR is reliable and ready to predict the antagonistic effect of new, congeneric calcium channel blockers.

Before, however, $\mathbf{X_{R1}}$ and $\mathbf{X_{R2}}$ were regressed against the **y** block to evaluate if only a given R substituent could explain the variance in the bioactivities. Indeed, a high correlation (r^2 = 0.96) was found when using $\mathbf{X_{R1}}$, while regression using $\mathbf{X_{R2}}$ resulted in a negligible correlation. This indicates that pIC_{50} is more affected by changes in R_1 than in R_2, probably because the latter refers to nonpolar substituents, while R_1 comprises nonpolar chains and groups containing electronegative heteroatoms. A quick comparison between pIC_{50} values and R_1 reveals that most of compounds with high pIC_{50} contains R_1

with heteroatoms, such as that found by PCA, despite R_1 in the experimentally most active compound is an *n*-pentyl group.

Other QSAR models have been built using DHP derivatives for pharmacophoric identification and classification, as well as to find highly predictive methods to support experimental data [13,31-36]. Our findings were used to propose new drug candidates; novel DHP analogues with expected high bioactivities can be designed by combining substructures of two or more highly active compounds of different series, keeping the common skeleton unaltered. Somewhat surprisingly, this combination has shown to be synergistic, *i.e.* some compounds risen from this methodology have shown calculated bioactivities higher than the parent compounds [37,38], since the best moiety of each original compound was used. The most active compounds of each series are **3**, **50, 53, 70** and **72**, whose substructures were interplayed to give compounds **A**-**F** of Figure 4. In fact, compounds **A** and **E** have already been synthesized and tested elsewhere [39]. The pIC_{50} values of these DHP derivatives were predicted using the MIA-QSAR model (obtained from the **X** matrix), and compounds **B**, **C** and **F** exhibited pIC_{50} values higher than the largest value found in the experimental data, indicating the synergistic effect mentioned above.

In order to validate these predictions, compounds **A**-**F** and amlodipine were docked into calcium channel in order to interact maximally with the residues Thr16-IIIS5, Gln20-IIIS5, Tyr7-IIIS6, Ile8-IIIS6, Ile11-IIIS6, Met15-IIIS6, Met16-IIIS6, Tyr7-IVS6, Met8-IVS6, Ile15-IVS6, Ile16-IVS6 and Asn17-IVS6. The overall docking scores of the proposed compounds correlate very well with the predicted pIC_{50} values obtained by the MIA-QSAR model (R^2 = 0.81, Figure 6). Figure 7 shows the superimposition of the best conformation of compound **B** and amlodipine into the active site of the calcium channel.

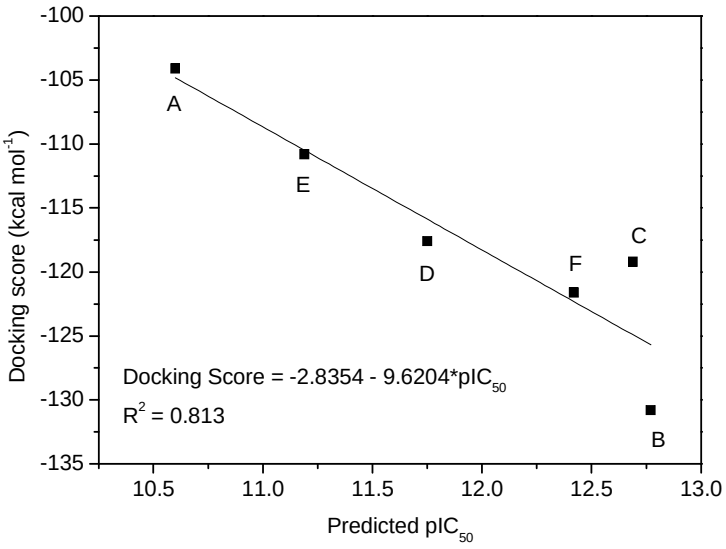


Figure 7 Correlation between the predicted pIC_{50} values of A-F using MIA-QSAR and the docking scores

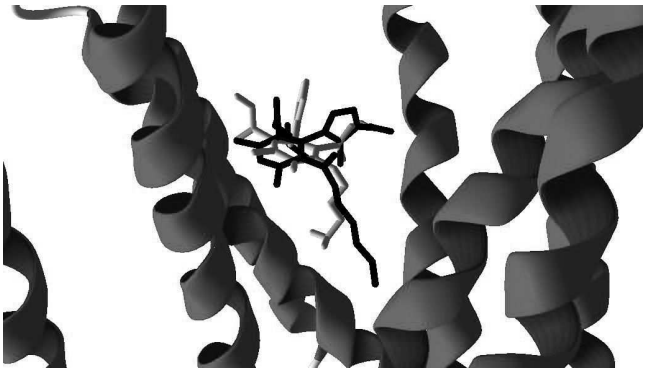


Figure 8 Superimposition of the best conformation of compound **B** (black) and amlodipine (gray) into the active site of the calcium channel

Drug-likeness can be estimated by predicting pharmacokinetics parameters in addition to bioactivity values. This can be achieved by accounting for the Lipinski´s rule of five [40], which states that most "drug-like" molecules have log*P* ≤ 5, molecular weight ≤ 500, number of hydrogen bond acceptors

(n_{ON}) ≤ 10, and number of hydrogen bond donors (n_{OHNH}) ≤ 5; molecules violating more than one of these rules may have problems with bioavailability. These parameters, together with the molecular polar surface area (TPSA, related to oral bioavailability), were readily assessed using the Molinspiration program (www.molinspiration.com) and indicated that the most active, predicted compounds (**B** and **C**) would not have any problem with absorption (Table 3); **F** and a symmetrical compound with two *n*-pentyl substituents presented log*P* or n_{ON} values higher than recommended.

Table 3 Calculated parameters of the Lipinski´s rule of five and TPSA for the proposed compounds (**A**-**F**), and for reference compounds

Compound	log*P*	MW	n_{ON}	n_{OHNH}	TPSA	Violations
Amlodipine	2.58	408.9	7	3	99.9	0
Felodipine	4.42	370.2	5	1	64.6	0
R^1 and R^2 = n-pentyl	5.29	462.5	10	0	128.3	1
A	2.53	392.4	10	1	128.3	0
B	4.47	448.5	10	1	128.3	0
C	4.09	434.5	10	1	128.3	0
D	3.27	420.5	10	1	128.3	0
E	2.90	406.4	10	1	128.3	0
F	2.32	436.5	11	1	137.5	1

Also, the pharmacokinetic data in Table 4 obtained by the Hologram QSAR Technique (www.pkdb.ifsc.usp.br) [41] indicate that the proposed compounds should have comparable properties when compared to the commercial CCBs Amlodipine and Felodipine.

Table 4 Computed pharmacokinetic parameters for the proposed compounds (**A**-**F**), and for reference compounds.[a]

Compound	HIA (%)	BBB (log BB)	Solubility (log S)[b]
Amlodipine	84.58	-1.18	-4.10
Felodipine	100.00	-0.49	-5.16
R^1 and R^2 = n-pentyl	100.00	-0.35	-6.20
A	81.27	-0.28	-1.83
B	84.10	-0.28	-1.69
C	83.53	-0.24	-1.86
D	81.59	-0.27	-1.80
E	81.43	-0.28	-1.83
F	85.50	-0.39	-1.75
Error	(+/- 13)	(+/- 0.30)	(+/- 0.82)

[a] HIA = human intestinal absorption; BBB = blood brain barrier penetration. [b] To obtain solubility in g L^{-1}, multiply logS by the molecular weight of the compound. [b] To obtain solubility in g/L multiply logS by molecular weight of the compound

In summary, some improved DHPs have been designed using MIA-QSAR, as confirmed by docking studies and computational drug-likeness assessment. Compounds **B** and **C** were found to be more promising, given their high predicted activities and pharmacokinetic data comparable to those of known calcium channel blockers. Despite the presence of an R polar group (together with a noncyclic hydrocarbon chain) in dictating enhanced bioactivities, the presence of a *n*-pentyl group seems to play a key role for the development of promising DPH analogues.

Acknowledgements

The authors thank FAPEMIG for the financial support of this research, as well as CAPES for studentships (to E.G.M., D.G.S., and M.C.G.) and CNPq for fellowships (to M.P.F. and E.F.F.C.). Dr. Tiago L. Moda is also gratefully acknowledged for the assistance with the pharmacokinetic calculations.

References

[1] M. Nelson, *Drug treatment of elevated blood pressure,* Aust. Prescr. 33 (2010), pp. 108-112.

[2] M. W. Wolowyk, E. E. Knaus in *Calcium Channel Modulators in Heart and Smooth Muscle: Basic Mechanisms and Pharmacological Aspects* (Eds.: S. Abraham, G. Amital), VCH, Weinheim, 1990.

[3] J. Fossheim, *Crystal structure of the dihydropyridine calcium antagonist felodipine. Dihydropyridine binding prerequisites assessed from crystallographic data,* J. Med. Chem. 29 (1986), pp. 305-307.

[4] G. H. Scholz, S. Vieweg, M. Uhlig, M. Thormann, P. Klossek, S. Goldmann, and H. J. Hofmann, *Inhibition of thyroid hormone uptake by calcium antagonists of the dihydropyridine class,* J. Med. Chem. 40 (1997), pp. 1530-1538.

[5] R. S. Kass, J. P. Arena, and S. Chin, *Block of L-type calcium channels by charged dihydropyridines,* J. Gen. Physiol. 68 (1991), pp. 63-75.

[6] R. S. Kass and J. P. Arena, *Influence of pH on calcium channel block by amlodipine, a charged dihydropyridine compound. Implications for location of the dihydropyridine receptor,* J. Gen. Physiol. 93 (1989), pp. 1109-1127.

[7] Y. W. Kwan, R. Bangalore, M. Lakitsh, H. Glossmann, and R. S. Kass, *Inhibition of cardiac L-type calcium channels by quaternary amlodipine: implications for pharmacokinetics and access to dihydropyridine binding site,* J. Mol. Cell. Cardiol. 27 (1995), pp. 253-262.

[8] R. Miri, C. A. McEwen, and E. E. Knaus, *Synthesis and calcium channel modulating effects of modified Hantzsch nitrooxyalkyl 1,4-dihydro-2,6-dimethyl-3-nitro-4-(pyridinyl or 2-trifluoromethylphenyl)-5-pyridinecarboxylates,* Drug Dev. Res. 51 (2000), pp. 225-232.

[9] R. Miri, S. E. Howlett, and E. E. Knaus, *Synthesis and calcium channel modulating effects of isopropyl 1,4- dihydro- 2,6- dimethyl-3-nitro-4-(thienyl)-5-pyridinecarboxylates,* Arch. Pharm. 330 (1997), pp. 290-294.

[10] A. Shafiee, R. Miri, A. R. Dehpour, and F. Solimani, *Synthesis and calcium channel antagonist activity of nifedipine analogues containing nitroimidazolyl substituent in guinea-pig ileal smooth muscle,* Pharmaceut. Sci. 2 (1996), pp. 541-543.

[11] R. Miri, H. Niknahad, G. H. Vesal, and A. Shafiee, *Synthesis and calcium channel antagonist activities of 3-nitrooxyalkyl, 5-alkyl 1,4-dihydro-2,6-dimethyl-4-(1-methyl-5-nitro-2-imidazolyl)-3,5-pyridinedicarboxylates,* Il Farmaco 57 (2002), 123-128.

[12] D. Sarsero, T. Fujiwara, P. Molenaar, and J. A. Angus, *Human vascular to cardiac tissue selectivity of L- and T-type calcium channel antagonists,* Br. J. Pharmacol. 145 (1998), pp. 109-119.

[13] M. A. Safarpour, B. Hemmateenejad, R. Miri, and M. Jamali, *Quantum chemical QSAR study of some newly synthesized calcium channel blockers,* QSAR Comb. Sci. 22 (2003), pp. 997-1005.

[14] M. P. Freitas, S. D. Brown, and J. A. Martins, *MIA-QSAR: A simple 2D image-based approach for quantitative structure–activity relationship analysis,* J. Mol. Struct. 738 (2005), pp. 149-154.

[15] ACD/ChemSketch Version 10.02, Advanced Chemistry Development, Inc., Toronto, 2006.

[16] Matlab 7.5, Mathworks Inc., Natick, 2007.

[17] M. P. Freitas, *Multivariate QSAR: From classical descriptors to new perspectives,* Curr. Comp.-Aid. Drug Des. 3 (2007), pp. 235-239.

[18] R. A. Cormanich, M. P. Freitas, and R. Rittner, *2D Chemical drawings correlate to bioactivities: MIA-QSAR modelling of antimalarial activities of*

2,5-diaminobenzophenone derivatives, J. Braz. Chem. Soc. 22 (2011), pp. 637-642.

[19] Pirouette 3.0.1, Infometrix Inc., Gainessville, 2006.

[20] I. Mitra, A. Saha, and K. Roy, *Exploring quantitative structure–activity relationship studies of antioxidant phenolic compounds obtained from traditional Chinese medicinal plants*, Mol. Simul. 36 (2010), pp. 1067-1079.

[21] K. Roy, I. Mitra, S. Kar, P. K. Ojha, R. N. Das, and H. Kabir, *Comparative studies on some metrics for external validation of QSPR models,* J. Chem. Inf. Model. 52 (2012), pp. 396-408.

[22] P. P. Roy, S. Paul, I. Mitra, and K. Roy, *On two novel parameters for validation of predictive QSAR models,* Molecules 14 (2009), pp. 1660-1701.

[23] P. K. Ojha, I. Mitra, R. N. Das, and K. Roy, *Further exploring r^2_m metrics for validation of QSPR models,* Chemom. Intell. Lab. Sys. 107 (2011), pp. 194-205.

[24] C. A. Nunes, M. P. Freitas, A. C. M. Pinheiro, and S. C. Bastos, *Chemoface: A novel free user-friendly interface for chemometrics*, J. Braz. Chem. Soc. 23 (2012), pp. 2003-2010.

[25] I. Huber, E. Wappl, A. Herzog, J. Mitterdorfer, H. Glossmann, T. Langer, and J. Striessnig, *Conserved Ca^{2+} antagonist binding properties and putative folding structure of a recombinant high affinity dihydropyridine binding domain*, Biochem. J. 347 (2000), pp. 829-836.

[26] G. Barreiro, C. R. W. Guimarães, and R. B. de Alencastro, *A molecular dynamics study of an L-type calcium channel model*, Protein Engin. 5 (2002), pp. 109-122.

[27] B. S. Zhorov, E. V. Folkman, and V. S. Ananthanarauanan, *Homology model of dihydropyridine receptor: implications for L-type Ca(2+) channel modulation by agonists and antagonists*, Arch. Biochem. Biophys. 393 (2001), pp. 22-41.

[28] R. Thomsen and M. H. Christensen, *MolDock: A new technique for high-accuracy molecular docking,* J. Med. Chem. 49 (2006), pp. 3315-3321.

[29] K. Pearson, *On lines and planes of closest fit to systems of points in space,* Philos. Mag. 6 (1901), pp. 559-572.

[30] A. Golbraikh and A. Tropsha, *Beware of q^2!* J. Mol. Graph. Model. 20 (2002), pp. 269-276.

[31] N. Geçen, E. Saripinar, E. Yanmaz, and K. Şahin, *Application of electron conformational–genetic algorithm approach to 1,4-dihydropyridines as calcium channel antagonists: pharmacophore identification and bioactivity prediction,* J. Mol. Model. 18 (2012), pp. 65–82.

[32] A. Mohajeri, B. Hemmateenejad, A. Mehdipour, and R. Miri, *Modeling calcium channel antagonistic activity of dihydropyridine derivatives using QTMS indices analyzed by GA-PLS and PC-GAPLS,* J. Mol. Graph. Model. 26 (2008) 1057–1065.

[33] P. Kahraman and M. Turkay, *Classification of 1,4-dihydropyridine calcium channel antagonists using the hyperbox approach,* Ind. Eng. Chem. Res. 46 (2007), pp. 4921–4929.

[34] X. Yao, H. Liu, R. Zhang, M. Liu, Z. Hu, A. Panaye, J. P. Doucet, and B. Fan, *QSAR and classification study of 1,4-dihydropyridine calcium channel antagonists based on least squares support vector machines,* Mol. Pharm. 2 (2005), pp. 348–356.

[35] Y. Takahata, M. C. A. Costa, and A. C. Gaudio, *Comparison between neural networks (NN) and principal component analysis (PCA): structure activity relationships of 1,4-dihydropyridine calcium channel antagonists (nifedipine analogues),* J. Chem. Inf. Comput. Sci. 43 (2003), pp. 540–544.

[36] K. J. Schleifer and E. Tot, *CoMFA, CoMSIA and GRID/GOLPE studies on calcium entry blocking 1,4-dihydropryridines,* Quant. Struct. Act. Relat. 21 (2002), pp. 239–248.

[37] J. R. Pinheiro, M. Bitencourt, E. F. F. da Cunha, T. C. Ramalho, and M. P. Freitas, *Novel anti-HIV cyclotriazadisulfonamide derivatives as modeled by ligand- and receptor-based approaches,* Bioorg. Med. Chem. 16 (2008), pp. 1683-1690.

[38] J. E. Antunes, M. P. Freitas, E. F. F. da Cunha, T. C. Ramalho, and R. Rittner, *In silico prediction of novel phosphodiesterase type-5 inhibitors derived from Sildenafil, Vardenafil and Tadalafil,* Bioorg. Med. Chem. 16 (2008), pp. 7599-7606.

[39] R. Miri, K. Javidnia, A. Kebriaie-Zadeh, H. Nkhnahad, N. Shaygani, S. Semnanian, and A. Shafiee, *Synthesis and evaluation of pharmacological activities of 3,5-dialkyl 1,4-dihydro-2,6-dimethyl-4-nitroimidazole-3,5-pyridine dicarboxylates,* Arch. Pharm. 336 (2003), pp. 422-428.

[40] C. A. Lipinski, F. Lombardo, B. W. Dominy, and P. J. Feeney, *Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings,* Adv. Drug. Delivery Rev. 23 (1997), pp. 3-25.

[41] T. L. Moda, L. G. Torres, A. E. Carrara, and A. D. Andricopulo, *PK/DB: database for pharmacokinetic properties and predictive in silico ADME models,* Bioinformatics 24 (2008), pp. 2270-2271.

ARTIGO 3

EXPLORING STRUCTURE-BASED DRUG DESIGN FOR THE DEVELOPMENT OF MULTI-TARGET ANTIHYPERTENSIVES

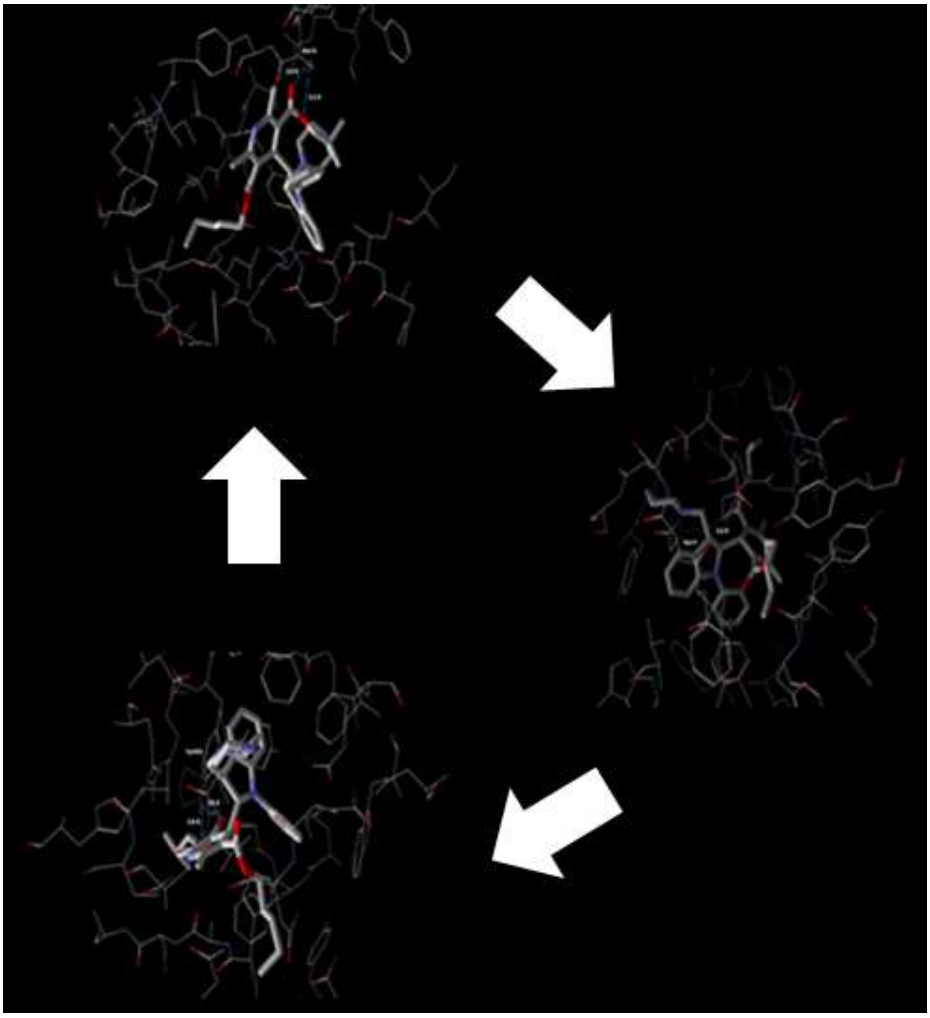
Artigo redigido conforme norma da revista Letters in Drug Design & Discovery.

Estella G. da Mota, Elaine F. F. da Cunha, Matheus P. Freitas

Abstract

This work reports the docking studies of compounds designed from the combination of substructures of three types of antihypertensives: angiotensin converting-enzyme (ACE) inhibitors, calcium channel blockers and renin inhibitors. Consequently, multi-target compounds are expected to be obtained. Indeed, a few purposes showed both docking scores and intermolecular ligand-enzyme interaction towards all three targets higher than the reference compounds Captopril (ACE inhibitor), Amlodipine (calcium channel blocker) and Aliskiren (renin inhibitor). These results, which were discussed in terms of ligand-enzyme interactions (especially hydrogen bonding between amino acid residues and ligands), indicate promising perspectives for synthesis and biological tests.

Graphical Abstract:



Keywords: antihypertensives; angiotensin converting-enzyme; calcium channel; renin; docking studies; intermolecular interactions

1. Introduction

Hypertension, or high blood pressure, is associated with cardiovascular diseases and accounts for 9.4 million deaths worldwide every year [1]. It can be prevented by means of healthy lifestyles and controlled using antihypertensives.

There is a variety of antihypertensive types, including angiotensin converting-enzyme (ACE) inhibitors, calcium channel blockers and renin inhibitors. Captopril (Figure 1) is a widely antihypertensive that acts by inhibiting the activity of ACE, an enzyme responsible for the conversion of angiotensin I to angiotensin II, a potent vasoconstrictor [2]. Recently, some Captopril analogues were virtually modeled and the corresponding calculated bioactivities for some of them were found to be higher than for the parent compound [3]. Similarly, the predicted bioactivities of some 1,4-dihydropyridines as calcium channel blockers were higher than the calculated value for the reference drug Amlodipine (Figure 1) [4]. Lastly, some indole-3-carboxamide derivatives have shown to be promising candidates as renin inhibitors [5,6], whose only marketed drug is Aliskiren (Figure 1).

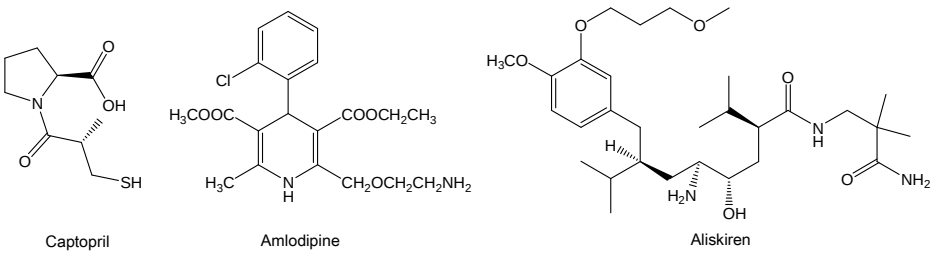


Figure 1 Chemical structures of Captopril (an ACE inhibitor), Amlodipine (a calcium channel blocker) and Aliskiren (a renin inhibitor)

Some drugs are administered simultaneously, like in some anti-HIV cocktails, in order to inhibit different enzymatic targets and to improve the

efficacy in the treatment. A single drug capable of inhibiting more than one enzymatic target, a multi-target drug, would benefit from a less exhaustive and more efficient treatment. Thus, this work reports the docking study of new molecules originated from the combination of substructures of chemical compounds computationally known as promising candidates either as ACE inhibitors or calcium channel blockers or renin inhibitors. Consequently, pharmacophoric moieties responsible for each type of activity are combined in a single compound to hopefully give a promising multi-target anti-hypertensive. The three compound classes have structural similarity (Figure 2), enabling superposition during the docking alignment.

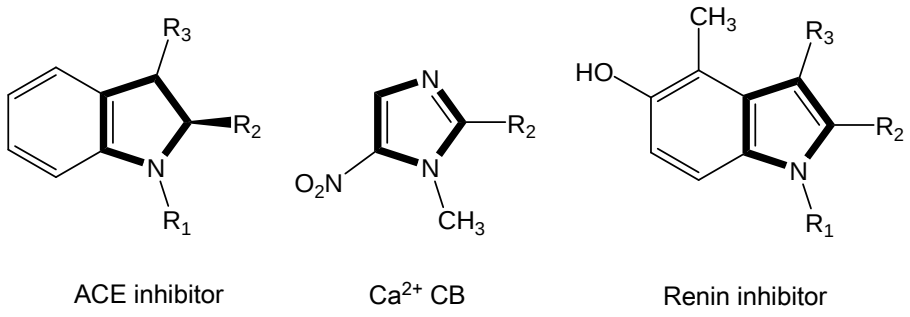
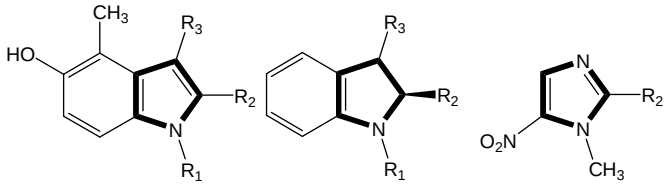


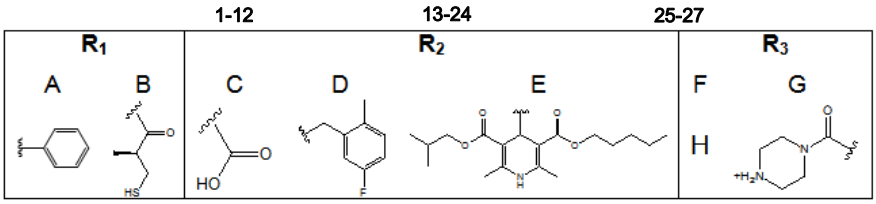
Figure 2 General frameworks of ACE inhibitors, calcium channel blockers and renin inhibitors used in this study. The common substructure used for docking alignment is bold faced

2. Materials and Methods

The most promising renin inhibitor, ACE inhibitor and calcium channel blocker (**4**, **19** and **27**, respectively) within the corresponding series of compounds available in the literature [3,4,6] were used to give rise to new multi-target drug candidates, whose chemical structures were derived from the combination of the substructures in these parent compounds. The explored multi-target drug candidates are given in Table 1.

Table 1 Proposed multi-target antihypertensives **1-27**





Compound	R_1	R_2	R_3
1	A	C	F
2	A	C	G
3	A	D	F
4[a]	A	D	G
5	A	E	F
6	A	E	G
7	B	C	F
8	B	C	G
9	B	D	F
10	B	D	G
11	B	E	F
12	B	E	G
13	A	C	F
14	A	C	G

Table 1, continuation

Compound	R_1	R_2	R_3
15	A	D	F
16	A	D	G
17	A	E	F
18	A	E	G
19[b]	B	C	F
20	B	C	G
21	B	D	F
22	B	D	G
23	B	E	F
24	B	E	G
25	-	C	-
26	-	D	-
27[c]	-	E	-

[a] Parent compound with high inhibitory activity against renin [6]. [b] Parent compound with high inhibitory activity against ACE [3]. [c] Parent compound highly active as calcium channel blocker [4]

The docking studies were performed for the proposed compounds **1-27** and also for the reference ones, *i.e.* Captopril, Amlodipine and Aliskiren (Figure 1). The crystal structures with the active sites of ACE (Captopril inside) and renin (Aliskiren inside) were obtained from the Protein Data Bank (PDB codes 1UZF and 2V0Z, respectively). In this study, the L-type calcium channel structure from the human cardiac α1-subunit (code CAC1C_HUMAN in the SWISS-PROT database) was modeled similarly as previously reported [4]. Because of recent alignments based on the pore-forming S6 segments with the M2 segment of KcsA [7,8], we used the scheme proposed by Zhorov et al. [9], taking into account the structure coordinates of KcsA Kþ channel available in the Protein Data Bank (code: 1BLB). The docking procedure was carried out using the Molegro Virtual Docker (MVD) program [10]. Ligand and protein were considered flexible during the docking simulation. According to the MolDock scoring function available in MVD, a piecewise linear potential are fit

to protein-ligand structures and binding data scoring functions, further extended in genetic evolutionary method including an hydrogen bonding term and charge schemes. The docking scoring function, E_{score}, is defined by the following energy terms:

$$E_{Score} = E_{Inter} + E_{Intra}$$

where E_{Inter} is the ligand–protein interaction energy:

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

E_{PLP} is the 'piecewise linear potential' that uses parameters to account for the steric term between atoms (Van der Waals) and hydrogen bonds, in addition to another term to describe electrostatic interactions between charged atoms. Further details can be found elsewhere [10]. E_{Intra} is the internal energy of the ligand. In this work, E_{Score} was used to account for the drug-likeness of the proposed molecules, since this parameter has shown to be highly correlated with bioactivity values in a variety of studies [3,4,6,11-13]. However, the E_{Inter} term, which corresponds to the ligand-enzyme interaction energy, will also be considered, because E_{Intra} can vary significantly from one compound another due to their structural diversity. In addition, the contribution from hydrogen bond, which is an important parameter describing ligand-enzyme interaction, was also computed.

3. Results and Discussion

Compounds **4**, **19** and **27** have been computationally found to be promising antihypertensives because of their high predicted activities against renin, ACE and calcium channel, respectively [3,4,6]. Since all three classes of compounds contain a congruent pharmacophoric group (Figure 2), the

substituents R_1, R_2 and R_3 of **4**, **19** and **27** can be exchanged in order to give a single molecule with properties related to renin, ACE and calcium channel inhibition/blockage. This can be estimated by docking all proposed compounds **1**-**27** in all three enzymatic targets and then comparing the results with the reference compounds, namely Aliskiren (renin inhibitor), Captopril (ACE inhibitor) and Amlodipine (calcium channel blocker).

The first approach compared the docking scores of **1**-**27** with the corresponding values of the reference compounds in their respective enzymatic targets (Table 2). Some prospectively sounding drug candidates appeared with docking score energies below (more negative) the reference compounds in their respective active sites, but only compound **18** was found to be more promising than the reference compounds in all three targets. Thus, compound **18** is most likely the best multi-target drug candidate amongst those of Table 2. Despite the good correlation between docking scores (E_{Score}) and the intermolecular enzyme-ligand interactions (E_{Inter}), *i.e.* R^2 = 0.87 for renin, 0.85 for ACE and 0.88 for calcium channel, the intramolecular energy computed within the docking score values can vary significantly from one ligand another, due to the high structural diversity within the series. Thus, the E_{Inter} was further evaluated to confirm the best inhibitor towards all three targets according to the intermolecular enzyme-ligand interaction.

The intermolecular interaction energies were consistent with the docking scores, since, in general, compounds with high docking scores were also those with the highest intermolecular energies. The bold faced data in Table 2 show compounds with both docking score and intermolecular energy more negative than the reference compounds. Again, compound **18** exhibited better theoretical performance than all three reference compounds in the respective enzymatic targets, except for the intermolecular energy in the renin active site. Even though, the calculated intermolecular energy between renin and compound **18**

was the highest within the series of proposed compounds and comparable to that of Aliskiren (-201.4 against -205.8 kcal mol^{-1}). Compound **18** carries a R_2 group which is typical of highly active calcium channel blockers, while the R_1 and R_3 substituents are present in known renin inhibitors. Consequently, **18** is expected to strongly target ACE even not having substituents similar to those found in inhibitors of this enzyme.

Table 2 Results from the docking modeling (in kcal mol^{-1})

Compound	ACE			Ca^{2+} channel			Renin		
	E_{Score}	E_{Inter}	E_{HB}	E_{Score}	E_{Inter}	E_{HB}	E_{Score}	E_{Inter}	E_{HB}
Captopril	-68.5	-76.1	-6.6						
Amlodipine				-65.8	-63.0	-0.1			
Aliskiren							-153.0	-205.8	-15.6
1	-71.4	-71.8	-7.0	-63.8	-62.6	-2.7	-93.4	-93.2	-2.2
2	**-79.2**	**-81.3**	-15.4	**-79.8**	**-76.1**	-3.9	-107.7	-113.8	-0.5
3	**-76.1**	**-83.4**	-2.4	**-88.4**	**-92.2**	-2.5	-71.6	-68.3	-1.0
4	-1.3	-21.2	2.2	**-98.6**	**-97.1**	0.0	**-182.3**	-169.1	-2.7
5	-82.9	-65.6	-5.0	**-105.1**	**-82.1**	-2.1	-150.9	-149.1	-8.0
6	-19.0	-1.0	-5.8	**-125.6**	**-104.8**	-0.9	-72.2	-78.6	-5.5
7	**-74.4**	**-78.3**	-10.8	-53.6	-50.7	0.0	-81.1	-82.3	-2.4
8	**-98.6**	**-94.9**	-10.9	-74.6	-63.0	0.1	-91.9	-84.4	-11.4
9	-33.9	-52.6	0.0	-54.8	-63.7	-2.7	-120.4	-131.0	0.0
10	-48.4	-68.0	-4.8	**-82.3**	**-87.9**	-0.1	-144.8	-147.3	-8.3
11	-4.5	-2.6	-2.2	**-83.5**	**-74.3**	-2.8	-114.7	-115.5	-4.4
12	-53.9	-46.7	-5.0	**-86.0**	**-90.6**	-1.2	-72.8	-88.4	-2.8
13	**-88.8**	**-91.4**	-5.8	-48.2	-49.5	0.0	-82.1	-86.4	-2.7
14	**-97.8**	**-101.6**	-7.5	-65.0	-68.8	-4.7	-132.8	-132.2	-11.0
15	**-72.2**	**-96.2**	0.0	**-72.9**	**-75.2**	0.0	-110.4	-115.6	0.0
16	-62.1	-68.0	-1.9	**-102.6**	**-95.3**	-2.5	-121.6	-150.3	-1.5
17	**-74.1**	**-92.5**	-2.5	**-94.6**	**-102.6**	-0.3	-152.4	-172.6	-1.8
18	**-88.7**	**-106.3**	2.5	**-130.4**	**-121.0**	-4.0	**-162.8**	-201.4	-2.2
19	**-76.2**	**-75.2**	-3.8	-32.9	-44.5	0.0	-69.6	-74.6	-3.1
20	**-95.1**	**-100.9**	-3.1	-48.0	-57.1	0.0	-98.9	-111.2	-7.6
21	-51.0	-66.1	0.0	-26.7	-53.9	0.0	-79.2	-94.3	-3.7
22	-68.2	-69.9	-8.5	**-84.8**	**-89.5**	-2.5	-122.2	-157.3	-5.5
23	**-121.8**	**-125.4**	-9.4	**-95.4**	**-103.4**	-2.5	-103.9	-144.2	-5.8
24	**-101.5**	**-141.0**	-6.9	**-120.1**	**-113.2**	-2.1	-141.9	-163.0	-7.9
25	-54.1	-62.5	-7.5	-27.6	-31.0	0.0	-45.7	-49.7	-3.4
26	-68.2	-62.9	-4.0	-46.7	-43.3	-0.1	-87.7	-85.0	-2.5
27	-69.5	-57.2	-2.0	**-84.2**	**-74.8**	-1.0	-115.7	-116.3	-3.4

Hydrogen bonding (HB) is an important descriptor of ligand affinity towards the active site of an enzyme. Indeed, at least for Aliskiren, HB seems to rule its affinity towards renin, since this interaction is calculated to be responsible by *ca.* 15 kcal mol^{-1} of ligand:enzyme stabilization. The most promising multi-target compound **18** was calculated to present strong interaction with amino acid residues in the active site of calcium channel due to HB (Table 2). On the other hand, this interaction does not appear to play a determinant role for its activity against ACE, while it stabilizes the **18**:renin interaction only modestly (by *ca.* 2 kcal mol^{-1}). Figure 3 shows compound **18** inside the active sites of each enzyme, where HB is depicted as dotted lines (especially relevant for the calcium channel), while other interactions, such as hydrophobic and electrostatic contributions, should operate.

A

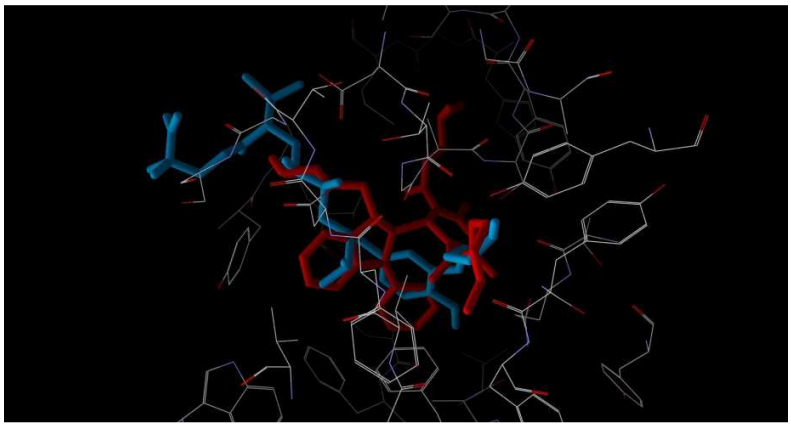
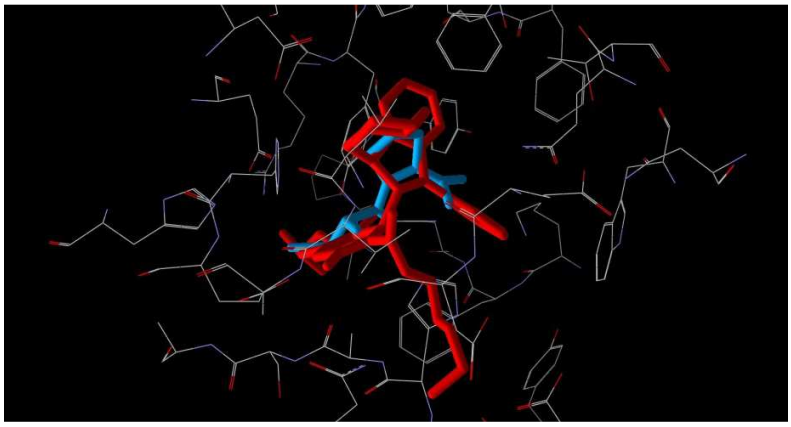


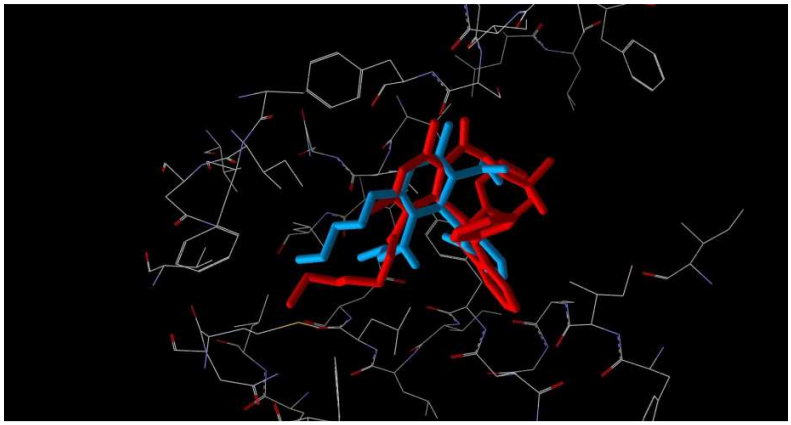
Figure 3 Compound **18** (red) superposed with the reference compounds (blue) Aliskiren, Captopril and Amlodipine in the respective active sites of docked inside the active sites of renin (**A**), ACE (**B**) and calcium channel (**C**)

(...continua...)

B



C



Overall, some proposed compounds showed better calculated performance in comparison to the reference compounds, while compound **18** is particularly suggested for synthesis as multi-target antihypertensive, because of its expected simultaneous affinities towards renin, ACE and calcium channel, which are higher than those affinities of the reference, marketed drugs Aliskiren, Captopril and Amlodipine, respectively. An important contribution for this behavior as calcium channel blocker comes from hydrogen bonding between the substituents in the molecule with the active site amino acid residues, while other interactions appear to rule its affinity towards renin and ACE.

Acknowledgements

Authors are thankful to FAPEMIG for the financial support of this research, as well as to CAPES for the studentship (E.G.M.) and to CNPq for fellowships (to E.F.F.C. and M.P.F.).

References

[1] Lim, S. S. et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study. *Lancet*, **2010**, *380*, 2224-2260.

[2] Paul, M.; Mehr, A. P.; Kreutz, R. Physiology of local renin-angiotensin systems. *Physiol. Rev.*, **2006**, 86, 747-803.

[3] Silva, D. G.; Freitas, M. P.; da Cunha, E. F. F.; Ramalho, T. C.; Nunes, C. A. Rational design of small modified peptides as ACE inhibitors. *Med. Chem. Commun.*, **2012**, *3*, 1290-1293.

[4] da Mota, E. G.; Silva, D. G.; Guimarães, M. C.; da Cunha, E. F. F.; Freitas, M. P. Computer-assisted design of novel 1,4-dihydropyridine calcium channel blockers. *Mol Simul.*, **2014**, *40*, 959-965.

[5] Jing, T.; Feng, J.; Zuo, Y.; Ran, B.; Liu, J.; He, G. Exploring the substructural space of indole-3-carboxamide derivatives binding to renin: a novel active-site spatial partitioning approach. *J. Mol. Model.*, **2012**, *18*, 4417–4426.

[6] da Mota, E. G.; Duarte, M. H.; da Cunha, E. F. F.; Freitas, M. P. Private communications (paper submitted elsewhere).

[7] Huber, I.; Wappl, E.; Herzog, A.; Mitterdorfer, J.; Glossmann, H.; Langer, T.; Striessnig, J. Conserved Ca^{2+} antagonist binding properties and putative

folding structure of a recombinant high affinity dihydropyridine binding domain. *Biochem. J.*, **2000**, *347*, 829–836.

[8] Barreiro, G.; Guimarães, C. R. W.; de Alencastro, R. B. A molecular dynamics study of an L-type calcium channel model. *Protein Eng.*, **2002**, *5*, 109–122.

[9] Zhorov, B. S.; Folkman, E. V.; Ananthanarayanan, V. S. Homology model of dihydropyridine receptor: implications for L-type Ca(2+) channel modulation by agonists and antagonists. *Arch. Biochem. Biophys.*, **2001**, *393*, 22–41.

[10] Thomsen, R.; Christensen, M. H. MolDock: A new technique for high-accuracy molecular docking. *J. Med. Chem.*, **2006**, *49*, 3315-3321.

[11] Guimarães, M. C.; Silva, D. G.; da Mota, E. G.; da Cunha, E. F. F.; Freitas, M. P. Computer-assisted design of dual-target anti-HIV-1 compounds. *Med. Chem. Res.* **2014**, *23*,1548-1558.

[12] Antunes, J. E.; Freitas, M. P.; da Cunha, E. F. F.; Ramalho, T. C.; Rittner, R. In silico prediction of novel phosphodiesterase type-5 inhibitors derived from Sildenafil, Vardenafil and Tadalafil. *Bioorg. Med. Chem.*, **2008**, *16*, 7599-7606.

[13] Pinheiro, J. R.; Bitencourt, M.; da Cunha, E. F. F.; Ramalho, T. C.; Freitas, M. P. Novel anti-HIV cyclotriazadisulfonamide derivatives as modeled by ligand- and receptor-based approaches. *Bioorg. Med. Chem.*, **2008**, *16*, 1683-1690.

CONSIDERAÇÕES FINAIS

Este trabalho possibilitou a construção de dois modelos QSAR, um para os inibidores da renina e outro para os bloqueadores dos canais de cálcio, ambos estatisticamente testados, demonstrando robustez e capacidade de predição, fato que foi confirmado pelos estudos de *docking.* Esses modelos foram utilizados para a predição de novas moléculas com atividade inibidora da renina e também novos bloqueadores dos canais de cálcio. Os novos compostos propostos apresentaram elevados valores calculados de pIC_{50}, sendo alguns deles mais ativos do que os relatados na literatura.

Estudos de *docking* de compostos concebidos a partir da combinação de subestruturas dos três tipos de anti-hipertensivos em questão (inibidores da renina, bloqueadores do canal de cálcio e inibidores da enzima conversora de angiotensina) também foram realizados. No geral, alguns compostos propostos apresentaram melhor desempenho calculado em comparação com os compostos de referência. O composto **18** é particularmente indicado para a síntese, como anti-hipertensivo multialvo, em razão de suas afinidades simultâneas com o sítio ativo da renina, da ECA e do canal de cálcio, que são superiores às apresentadas pelos compostos comercializados alisquireno, captopril e amlodipina, respectivamente.

A capacidade de modelagem e as informações químicas fornecidas pelos modelos QSAR podem conduzir à síntese de novos e mais potentes inibidores da renina e bloqueadores dos canais de cálcio. Os resultados animadores obtidos para o desenvolvimento de anti-hipertensivos multialvo abrem possibilidades para que futuros trabalhos sejam realizados, contribuindo para a elaboração de novos fármacos.