

## INTEGRAÇÃO DAS TÉCNICAS DE TRIAGEM VIRTUAL E TRIAGEM BIOLÓGICA AUTOMATIZADA EM ALTA ESCALA: OPORTUNIDADES E DESAFIOS EM P&D DE FÁRMACOS

Rafaela S. Ferreira, Glaucius Oliva e Adriano D. Andricopulo\*

Instituto de Física de São Carlos, Universidade de São Paulo, 13560-970 São Carlos – SP, Brasil

Recebido em 16/1/11; aceito em 2/6/11; publicado na web em 26/7/11

INTEGRATING VIRTUAL AND HIGH-THROUGHPUT SCREENING: OPPORTUNITIES AND CHALLENGES IN DRUG RESEARCH AND DEVELOPMENT. High-throughput screening (HTS) and virtual screening (VS) are useful methods employed in drug discovery, allowing the identification of promising hits for lead optimization. The efficiency of these approaches depends on a number of factors, such as the organization of high quality databases of compounds and the parameterization of essential components of the screen process. This brief review presents the basic principles of the HTS and VS methods, as well as a perspective of the utility and integration of these drug design approaches, highlighting current opportunities and future challenges in medicinal chemistry.

Keywords: high-throughput screening; virtual screening; drug design.

### INTRODUÇÃO

A descoberta de novos fármacos é um grande desafio para a indústria farmacêutica moderna. Os altos investimentos na área de pesquisa e desenvolvimento (P&D) contrastam com o número de novos medicamentos que têm chegado ao mercado nos últimos anos.<sup>1</sup> Este complexo panorama tem forçado a adoção de novas estratégias com o objetivo de aumentar a eficiência do processo de P&D, tendo como alicerces as inovações científicas, tecnológicas e empresariais.

Considerando-se o expressivo número de alvos biológicos (proteínas alvo) promissores para o planejamento de fármacos, as técnicas de triagem biológica automatizada em alta escala (HTS - *high throughput screening*) e de triagem virtual (VS - *virtual screening*) têm ocupado papel de destaque entre as estratégias modernas exploradas na identificação de novas substâncias bioativas. Desde o seu surgimento, a HTS tem sido amplamente utilizada pela indústria farmacêutica, propiciando a avaliação de milhões de compostos contra proteínas alvo. Embora esta estratégia tenha despertado a esperança de uma revolução na descoberta de substâncias bioativas, em pouco tempo as suas limitações ficaram evidentes.<sup>2</sup> A ocorrência de um número elevado de falso-positivos, identificados inicialmente como *hits* (termo em inglês para designar um novo ligante ou composto bioativo) nos ensaios bioquímicos em alta escala pode ser destacada como um dos grandes problemas.<sup>3</sup> Os múltiplos modos de interação ou mecanismos de ação apresentados pelos diferentes *hits* selecionados podem também ser mencionados como outros importantes obstáculos.<sup>4</sup> Além disso, o número possível de compostos com propriedades fármaco-similar (*drug-like*)<sup>5</sup> é várias ordens de magnitude maior do que o número de compostos que pode ser analisado nos processos de HTS. Portanto, o planejamento das bases de dados de compostos é essencial, sendo que fatores como propriedades moleculares e físico-químicas, ou ainda a diversidade química, devem ser considerados.<sup>6-8</sup>

Com os notáveis avanços da biologia estrutural e a disponibilidade de recursos computacionais de alto desempenho, a VS surgiu como importante alternativa à HTS. A VS é um método *in silico* (“computacional”) empregado no estudo de grandes bases de dados

de compostos, com o objetivo de identificar moléculas pequenas com elevado potencial de interagirem com proteínas alvo, para posterior avaliação bioquímica *in vitro*. As estratégias de VS têm como base a estrutura de ligantes bioativos (LBVS - *ligand-based virtual screening*; em português, triagem virtual baseada na estrutura do ligante) ou, a estrutura do receptor biológico (SBVS - *structure-based virtual screening*; em português, triagem virtual baseada na estrutura do receptor).<sup>9</sup> Nestes casos, centenas de milhares de moléculas podem ser avaliadas com base tanto na similaridade de ligantes conhecidos (LBVS), quanto no acoplamento a sítios de ligação específicos de proteínas alvo (SBVS). Este último processo, conhecido como *docking molecular* (*molecular docking*), possibilita a seleção de uma pequena fração de moléculas apresentando um conjunto favorável de interações intermoleculares (SBVS), com base em um determinado mecanismo e modo de ligação (*e.g.*, inibidores enzimáticos que atuam de maneira reversível através de um mecanismo do tipo competitivo).

Com características marcantes como elevado desempenho, versatilidade, baixo custo e facilidade de uso, o sucesso do emprego da VS pode ser verificado pelos inúmeros compostos bioativos descritos na literatura.<sup>9</sup> No entanto, existem algumas limitações inerentes ao método que devem ser entendidas. Na avaliação de compostos por SBVS, por exemplo, os modelos utilizados são simplificados e não representam de forma completa as interações intermoleculares entre ligantes (micromoléculas) e proteínas alvo (macromoléculas).<sup>10</sup> Conseqüentemente, além de uma fração de compostos que possui boa pontuação (*i.e.*, predição de alta afinidade pela proteína alvo) e atividade biológica, há também uma fração de falso-positivos (*i.e.*, compostos que obtêm boa pontuação, mas não se ligam com alta afinidade ao alvo) que não apresenta atividade biológica. Além disso, existe o problema dos falso-negativos, ou seja, compostos com boa atividade biológica que não obtêm boa pontuação, sendo descartados durante a análise.<sup>11</sup>

Os métodos de HTS e VS, bem como todas as tecnologias modernas, apresentam vantagens e desvantagens. Para que os melhores resultados possam ser alcançados, uma questão crucial é o planejamento da estratégia a ser empregada, fundamentado em um entendimento sólido da informação disponível e das potencialidades e limitações dos métodos.<sup>9</sup> Este aspecto ganha maior importância quando se trata de um número muito grande de moléculas (*e.g.*, da ordem de cente-

\*e-mail: aandrico@ifsc.usp.br

nas de milhares, ou até de milhões) que deve ser avaliado para um determinado alvo macromolecular. A decisão sobre qual técnica é a mais apropriada passa diretamente pela avaliação do alvo molecular e da base de dados. Em geral, o número de *hits* identificado pelas duas técnicas é extremamente alto,<sup>3,4,12</sup> frequentemente observando-se atividade inicial para mais de 1% do total de compostos, que devem então ser avaliados mais detalhadamente. Observa-se também que ambas abordagens permitem a descoberta de diversas classes de compostos, sendo de comparável magnitude a sua diversidade estrutural. Desta maneira, a integração destas e de outras técnicas de planejamento de fármacos tem sido tema de grande interesse na última década, e embora diversos trabalhos tenham sido descritos com abordagens interessantes,<sup>4,12-17</sup> raras são as informações conclusivas de seu impacto. Por outro lado, é essencial enfatizar a natureza complementar destas técnicas e o seu elevado valor nos estágios iniciais de projetos de descoberta de novas substâncias bioativas.

Nesta revisão será discutido o estado da arte do uso destas tecnologias modernas, através de uma abordagem crítica envolvendo aspectos experimentais e computacionais e seus fundamentos em Química e Biologia. Serão também discutidas oportunidades e desafios em P&D de fármacos na aplicação e integração destas técnicas por laboratórios de pesquisa acadêmicos e de indústrias farmacêuticas.

## SELEÇÃO DO ALVO MOLECULAR

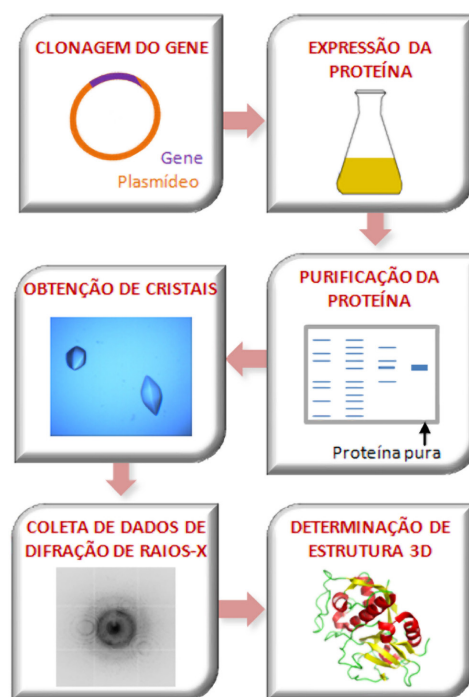
Os fármacos contemporâneos são planejados para atuar, principalmente, através da modulação seletiva da resposta biológica de proteínas alvo. Uma etapa inicial de extrema relevância neste processo é a seleção do alvo molecular, que engloba diversos fatores, como o entendimento da fisiopatologia da doença em questão, o detalhamento da via bioquímica envolvida, a seleção de uma ou mais proteínas alvo e, a validação e o estudo do envolvimento do alvo no tratamento ou cura da disfunção ou doença humana.<sup>18</sup> Os fármacos têm como alvos moleculares enzimas, receptores acoplados à proteína G (GPCRs - *G protein-coupled receptors*), receptores nucleares e canais iônicos, entre outros.

Diversas estratégias podem ser utilizadas na identificação e seleção de alvos biológicos. A análise do nível de expressão de proteínas de um mesmo organismo (*e.g.*, normais, superexpressas, deficientes) pode relevar características estruturais e funcionais que permitam diferenciar tecidos patológicos dos saudáveis, levando à seleção de alvos para o planejamento de fármacos.<sup>19</sup> No caso de doenças humanas causadas por micro-organismos, uma investigação das principais vias bioquímicas possibilita a identificação de proteínas essenciais para o crescimento e desenvolvimento do micro-organismo alvo (*e.g.*, bactéria, parasita, fungo). Uma análise comparativa dos organismos pode levar, preferencialmente, à seleção de proteínas ausentes no hospedeiro (*i.e.*, específicas para o patógeno) ou então com importantes diferenças estruturais em relação às proteínas que desempenham função análoga em humanos.<sup>20</sup> Assim, é possível o planejamento de ligantes seletivos para a(s) proteína(s) do micro-organismo em questão, minimizando problemas como inespecificidade e toxicidade.

A partir do processo de seleção de um alvo molecular, a validação se torna um fator crucial. Com o uso de técnicas de biologia molecular, como silenciamento de genes (*knockout*) ou RNA de interferência (RNAi),<sup>21</sup> ou ainda através da utilização de compostos químicos (*e.g.*, ligantes de receptores, inibidores enzimáticos),<sup>22</sup> é possível estudar os efeitos causados pela modulação da proteína alvo, simulando o efeito de um fármaco.

Após a validação do alvo molecular, ou seja, diante da confirmação do envolvimento da proteína com o estado de disfunção ou doença, se inicia o trabalho de identificação de novos ligantes. Para isso, é necessária a obtenção de quantidades apropriadas da proteína

com elevado grau de pureza para a realização de ensaios biológicos e de cristalização. Este processo envolve várias técnicas de biologia molecular que permitem a clonagem do gene de interesse em um vetor e a sua expressão em bactérias ou leveduras, seguido da purificação e caracterização da proteína (Figura 1).



**Figura 1.** Etapas gerais da preparação do alvo biológico, desde a clonagem do gene de interesse até a determinação da estrutura tridimensional

Para a aplicação de métodos de planejamento de fármacos baseado na estrutura do receptor (SBDD - *structure-based drug design*), é necessária a determinação da estrutura tridimensional da proteína, frequentemente, por difração de raios-X, ou ainda, em muitos casos, por ressonância magnética nuclear (RMN). A coleta de dados de difração de raios-X é precedida pela obtenção de cristais da proteína, possibilitando a determinação da estrutura em alta resolução.

## ORGANIZAÇÃO DE BASES DE DADOS DE COMPOSTOS

Após a definição do alvo molecular de interesse, a etapa subsequente é a organização de uma ou mais bases de dados de compostos. Este aspecto é fundamental, pois o número disponível de moléculas é da ordem de vários milhões, sendo muito superior ao que pode ser testado em condições reais, o que torna imperiosa a preparação de coleções dirigidas (*focused libraries*). A partir de um universo tão imenso de moléculas, pode-se facilmente perceber que é muito pequena a probabilidade de se identificar compostos ativos promissores através de uma avaliação aleatória (*e.g.*, tentativa e erro). Vários estudos indicam que os compostos de bases de dados empregados na HTS possuem fortes características fármaco-similar devido à presença de esqueletos moleculares privilegiados, ao que se atribui em parte o sucesso desta técnica.<sup>23</sup>

A geração de bases de dados é influenciada diretamente por diversos fatores, dentre os quais se destacam as propriedades moleculares e físico-químicas. Um marco importante nesta área foi o estudo desenvolvido por Lipinski e colaboradores, em 1997, sobre algumas propriedades de fármacos administrados por via oral, que resultou na elaboração da *regra dos cinco* (*rule of five*), amplamente utilizada como filtro molecular em projetos de planejamento de fármacos devido a sua simplicidade conceitual e facilidade de aplicação.<sup>5</sup>

**Tabela 1.** Compostos com características fragmento-similar, líder-similar e fármaco-similar

Propriedade	Fragmento-similar	Líder-similar	Fármaco-similar
Peso molecular	< 250	< 350	< 500
Log <i>P</i>	2,5	3,5	< 5
Número de doadores de ligação de hidrogênio	< 3	< 4	< 5
Número de aceptores de ligação de hidrogênio	< 6	< 8	< 10

De acordo com a *regra dos cinco*, a ocorrência de baixa absorção após administração por via oral, ou de baixa permeabilidade, é mais provável para compostos que apresentam uma combinação de duas ou mais propriedades, tais como, peso molecular > 500, número de doadores de ligação de hidrogênio > 5 (soma de grupos OH e NH), número de aceptores de ligação de hidrogênio > 10 (soma de átomos O e N) e,  $\text{clog } P > 5$  (coeficiente de participação calculado para o sistema 1-octanol/água).<sup>5</sup>

Por outro lado, surgiu também o conceito de líder-similar (*lead-like*), empregado para moléculas com propriedades semelhantes às de compostos líderes. Este conceito encontrou grande aplicabilidade, principalmente em estágios iniciais de P&D, onde as substâncias bioativas devem apresentar propriedades similares às de outros compostos que serviram como ponto de partida para o desenvolvimento de novas entidades químicas (NCEs - *new chemical entities*). Posteriormente, foram investigadas as principais alterações nas propriedades moleculares e físico-químicas durante o processo de otimização de compostos líderes (*e.g.*, melhorias nas propriedades farmacodinâmicas e farmacocinéticas).<sup>7</sup> Mais especificamente, as tendências de alterações no peso molecular, número de aceptores ou doadores de ligações de hidrogênios e  $\text{clog } P$  tornaram possível a definição das características líder-similar (Tabela 1), que são muito empregadas na geração de coleções dirigidas de compostos.<sup>6,8</sup>

Outra tendência atual é o uso de bases de dados com características fragmento-similar (*fragment-like*).<sup>24</sup> Os fragmentos moleculares destas bases exibem, geralmente, baixa afinidade pelo sítio de ligação da proteína alvo. Contudo, em função de sua menor complexidade estrutural e boa complementaridade ao alvo de interesse, os fragmentos (*i.e.*, compostos de baixo peso molecular) são alvos atrativos para estudos de química medicinal. Apesar das bases de dados de fragmentos serem relativamente pequenas, contando com alguns milhares de compostos, a diversidade estrutural que pode ser explorada no processo de planejamento molecular é significativa.<sup>24,25</sup> A otimização de fragmentos, sejam estes obtidos por triagens experimentais ou computacionais, está diretamente associada à biologia estrutural. O modo de ligação do fragmento na cavidade apropriada do alvo biológico é utilizado como padrão de reconhecimento molecular para a aplicação de métodos de planejamento de fármacos baseado em fragmentos (FBDD - *fragment-based drug design*). Diversos exemplos enfatizam o uso bem sucedido dessa estratégia na identificação e otimização de ligantes bioativos contribuindo, assim, para a sua rápida difusão e popularização.<sup>26,27</sup>

Além das propriedades físico-químicas, outros fatores podem ser levados em conta na construção de bases de dados, como a avaliação da reatividade química e da diversidade estrutural dos compostos através de descritores bi- (2D) ou tridimensionais (3D).<sup>6,8,28,29</sup> Nos casos em que se pretende avaliar as bases de dados contra diferentes alvos moleculares, a diversidade estrutural é considerada de extrema importância para aumentar as chances de identificação de novos ligantes. Já o desenvolvimento de bases de dados dirigidas, seja para um dado alvo molecular ou determinada classe de proteínas, envolve a seleção de coleções de compostos com base em estruturas moleculares ou propriedades físico-químicas privilegiadas, relacionadas às de ligantes com propriedades conhecidas.<sup>30</sup> Nestes casos, as bases geradas com

o auxílio da informação desses ligantes possuem maior chance de sucesso na identificação de novos compostos bioativos.<sup>31</sup> Em todos os casos, fica evidente o papel crucial da seleção criteriosa do conjunto de compostos a ser avaliado por triagens virtuais ou experimentais.

### TRIAGEM BIOLÓGICA AUTOMATIZADA EM ALTA ESCALA

A triagem biológica automatizada em alta escala (HTS) propicia a avaliação biológica *in vitro* de milhares ou até mesmo milhões de compostos contra uma proteína alvo ou sistema celular definido. Os processos de HTS evoluíram consideravelmente nas últimas décadas graças aos progressos da tecnologia da robotização e de equipamentos automatizados que tornaram possível a realização de ensaios miniaturizados de forma rápida, eficiente e confiável. Além disso, cresce cada vez mais o número de coleções de compostos, fato que pode ser atribuído ao desenvolvimento de metodologias mais eficientes de síntese e aos avanços da química combinatória.

As grandes indústrias farmacêuticas do mundo investiram seriamente em infraestrutura de alta qualidade para realização da HTS, tendo como objetivo principal a descoberta de *hits* nos estágios iniciais de P&D de fármacos. O anticoagulante rivaroxaban (Xarelto®, Bayer), que atua através da inibição do fator Xa,<sup>32</sup> e o inibidor não peptídico da protease do vírus da imunodeficiência humana (HIV - *human immunodeficiency virus*) tipranavir (Aptivus®, Pfizer/Boehringer Ingelheim)<sup>33</sup> são medicamentos desenvolvidos a partir de *hits* identificados por HTS.

Apesar dos sucessos alcançados com o emprego desta técnica, algumas limitações tornaram-se aparentes, como, por exemplo, a elevada porcentagem de falso-positivos identificada na triagem de inibidores enzimáticos. O processo de HTS, invariavelmente, leva à identificação de um número muito grande de *hits*, sendo que apenas uma fração modesta é de ligantes verdadeiros da enzima alvo. Neste caso, o resultado encontrado para uma porção expressiva de *hits* não corresponde à inibição da enzima por mecanismos específicos (*e.g.*, inibição reversível do tipo competitiva), mas pode estar associado a fatores como agregação, inibição promíscua ou artefatos decorrentes do ensaio utilizado (*e.g.*, alta fluorescência, absorvância dos compostos testados).<sup>3,4,12,34,35</sup> É importante ressaltar que a triagem inicial representa o primeiro passo de um processo de HTS. Após a identificação dos *hits*, se faz necessária a realização de ensaios secundários para a confirmação e validação dos resultados, permitindo a priorização de compostos (Figura 2).

A seleção de compostos que atuam por mecanismos distintos (*i.e.*, compostos que podem se ligar a cavidades diferentes do sítio ativo da proteína alvo), o que é uma característica da HTS, proporciona a caracterização de novos sítios alostéricos. Esta pode ser uma vantagem no planejamento de ligantes, pois estes sítios tendem a ser menos conservados do que o sítio ativo entre diferentes espécies.

### TRIAGEM VIRTUAL BASEADA NA ESTRUTURA DO LIGANTE

Os métodos de LBVS e SBVS são alternativas eficientes e



Figura 2. Esquema da triagem biológica automatizada em alta escala (HTS)

versáteis à HTS. Estes possibilitam, a partir da escolha de uma proteína alvo, a seleção de um sítio ativo ou alostérico para o estabelecimento de estratégias com base no mecanismo de ação ou modo de ligação. Além disso, as bases de dados podem ser preparadas de forma a conter compostos com um determinado perfil de propriedades moleculares e físico-químicas. Na aplicação da LBVS, compostos bioativos conhecidos podem ser utilizados como referência para comparações por similaridade química com outras moléculas das bases de dados, ou ser empregados na validação de modelos, juntamente com compostos inativos. As principais técnicas de LBVS incluem o estudo das relações quantitativas entre a estrutura e atividade (QSAR - *quantitative structure-activity relationships*),<sup>36-39</sup> o uso de impressões digitais moleculares (*molecular fingerprints*),<sup>40,41</sup> o emprego de farmacóforos (*pharmacophores*)<sup>42</sup> e, a utilização de inteligência artificial na aprendizagem de máquinas (*machine learning*), como as máquinas de vetor suporte (SVM - *support vector machines*).<sup>43</sup>

Para a geração de modelos de QSAR 2D e 3D, é indispensável a organização de conjuntos padrões de dados. Os compostos devem pertencer à mesma classe química ou possuir um farmacóforo comum. Além disso, devem se ligar a uma única cavidade de interação da proteína e atuar pelo mesmo mecanismo de ação. O parâmetro biológico deve ser quantitativo (e.g.,  $IC_{50}$ ,  $EC_{50}$ ,  $K_i$ ) e comparável para a série de compostos. O objetivo principal dos métodos de QSAR é investigar como a variação da estrutura química afeta a potência biológica ou afinidade de ligação pelo receptor alvo. O ideal é o uso de parâmetros quantitativos *in vitro* contra a proteína alvo, já que dados obtidos em ensaios celulares são dependentes de várias outras propriedades (e.g., permeabilidade celular, solubilidade, estabilidade). Por outro lado, os dados provenientes de diferentes fontes devem ter sido obtidos com o mesmo protocolo e sob as mesmas condições experimentais, assegurando a confiabilidade dos resultados. Bases como a MDDR (MDL - *drug data report*) e a WOMBAT (*world of molecular bioactivity*) reúnem dados padronizados de alta qualidade para diversos alvos biológicos, facilitando a organização de conjuntos de dados para a aplicação de métodos de LBVS.

A descoberta de novas classes de compostos (quimicamente distintas) é um dos maiores desafios dos métodos de LBVS. Estratégias de VS por similaridade química ou modelos farmacofóricos permitem a investigação de conjuntos de dados com diversidade estrutural mais ampla do que aquela encontrada em estudos de QSAR.<sup>37,42,44</sup> Alguns exemplos ilustram a aplicação destas técnicas na identificação de novos esqueletos químicos (*scaffold hopping*).<sup>40,45</sup>

## Etapas de LBVS

Em geral, os métodos de LBVS se fundamentam em estudos de similaridade ou em modelos construídos a partir de conjuntos de ligantes e não ligantes.<sup>40,43</sup> As buscas por similaridade química envolvem a geração de impressões digitais moleculares para todas as moléculas das bases de dados. Esta etapa evolui para a comparação de cada molécula das bases com um ligante de referência (*i.e.*, um composto com atividade biológica conhecida), o que permite a classificação dos compostos em relação a sua similaridade química (Figura 3a). Embora a opção mais simples consista no emprego de apenas um ligante de referência, o desempenho desta estratégia pode ser melhorado quando diversos ligantes são utilizados.<sup>46</sup> Para a classificação das bases de dados pode ser utilizada a análise comparativa da maior similaridade, ou da média das similaridades de cada composto em relação aos ligantes de referência. Por outro lado, para a construção de modelos para a predição da propriedade biológica, o conjunto de dados pode ser dividido em conjunto treinamento, usado para a geração estatística dos modelos, e conjunto teste (contendo novas moléculas que não fizeram parte da geração dos modelos), empregado para a validação externa do modelo. Neste contexto, diversos modelos são avaliados e o mais robusto pode ser utilizado na LBVS, empregando métodos de QSAR e SVM (Figura 3b).

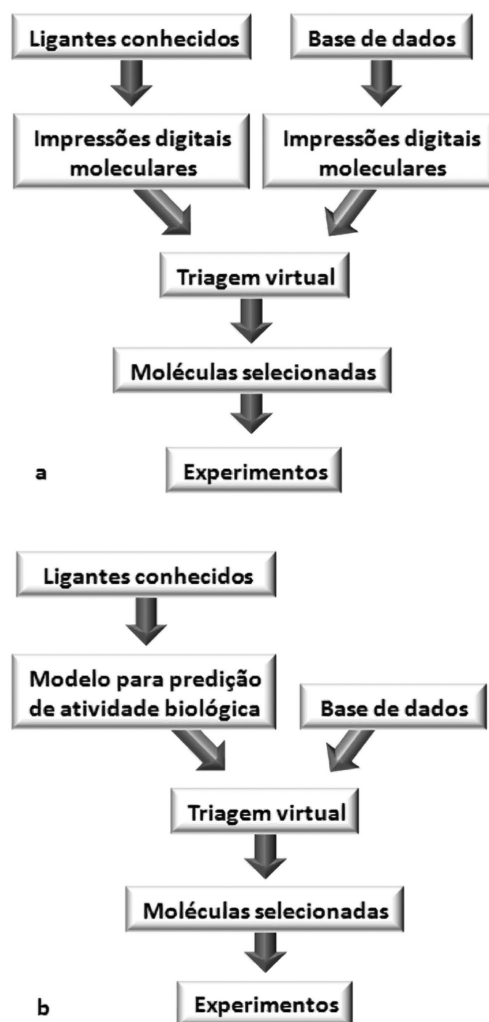


Figura 3. Estratégias de LBVS: a) LBVS baseada em similaridade química; b) LBVS empregando modelos para predição de atividade biológica. No caso de modelos construídos por SVM, informações sobre compostos inativos também são incorporadas ao modelo



As técnicas de LBVS têm encontrado grande versatilidade de emprego e notável qualidade na integração com métodos de SBVS ou HTS, levando à identificação de moléculas bioativas com elevada diversidade estrutural para um conjunto amplo de alvos moleculares.<sup>47</sup> Em particular, a partir de um conjunto de ligantes selecionados por HTS ou SBVS, a LBVS pode ser utilizada na busca de novos compostos bioativos. Outrossim, esta técnica é útil na análise prévia de grandes bases de dados, permitindo a geração de coleções dirigidas para avaliações posteriores, experimentais ou computacionais.<sup>15,17,41,45</sup>

### TRIAGEM VIRTUAL BASEADA NA ESTRUTURA DO RECEPTOR

A SBVS abrange a avaliação de bases de dados a partir da simulação de interações intermoleculares entre moléculas pequenas (ligantes) e receptores biológicos (proteínas). Os métodos de SBVS permitem a descoberta de novos ligantes com diversidade estrutural maior em comparação àqueles identificados na LBVS. Em contrapartida, um requerimento essencial é o conhecimento prévio da estrutura tridimensional da proteína, que pode ser determinada experimentalmente, ou representada por um modelo construído a partir de proteínas homólogas com alta identidade. A exigência do conhecimento da estrutura tridimensional do alvo pode ser apontada como uma das principais limitações da SBVS, já que esta informação não está disponível para diversos alvos terapêuticos validados, cuja cristalização ainda constitui grande desafio. São exemplos importantes os receptores transmembrana, que incluem os diversos GPCRs. No entanto, os frequentes avanços da biologia estrutural têm proporcionado novas oportunidades e desafios para a aplicação da SBVS.

A aplicação harmônica e eficaz dos métodos de SBVS exige diversos cuidados. Primeiramente, a preparação dos ligantes da base de dados requer o cálculo de cargas atômicas e a representação correta de estereoisômeros, tautômeros e estados de protonação.<sup>48,49</sup> Por outro lado, a preparação da proteína envolve fatores importantes como a escolha do modelo tridimensional, a definição do estado de protonação de resíduos e a presença de moléculas de água na cavidade de interação (Figura 4).

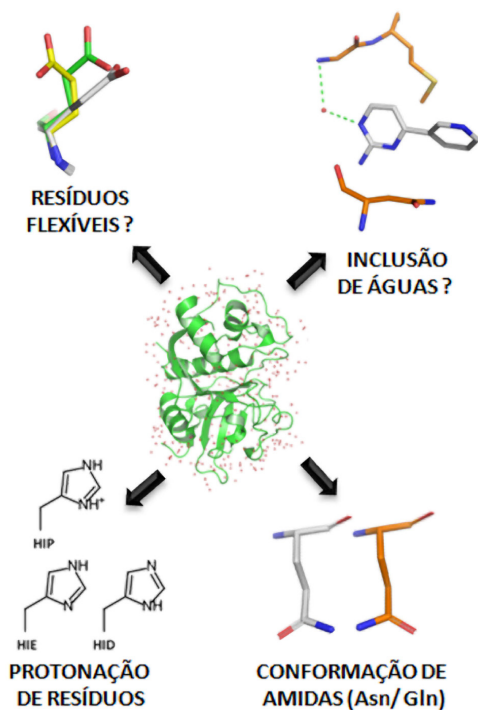


Figura 4. Cuidados na preparação do receptor para SBVS

A estrutura tridimensional da proteína alvo, determinada por técnicas experimentais como a cristalografia de raios-X ou RMN, ou ainda, por métodos computacionais como a modelagem por homologia, deve apresentar informações úteis e de alta qualidade. No caso da disponibilidade de diversas estruturas do mesmo alvo molecular, a escolha do modelo a ser empregado deve levar em conta a resolução e a flexibilidade das estruturas, bem como a presença de ligantes e as variações conformacionais, principalmente no sítio de ligação alvo.

Devido ao grande número de conformeros possíveis, a proteína é geralmente tratada de forma rígida (*i.e.*, em uma única conformação observada por cristalografia) ou com flexibilidade limitada. Esta restrição, associada ao ajuste induzido frequentemente observado na formação de complexos proteína-ligante, faz com que a SBVS seja mais bem sucedida quando são utilizadas estruturas da proteína em complexo com ligantes.<sup>50</sup> A flexibilidade dos resíduos de aminoácido do sítio alvo deve ser avaliada de forma criteriosa, principalmente, se houver várias estruturas disponíveis. Assim, é possível discriminar as regiões de maior flexibilidade e priorizar os resíduos a serem tratados como flexíveis, se o programa de docagem molecular possuir esta funcionalidade. Diversas estratégias levam em conta a flexibilidade proteica, como a utilização de várias estruturas para a docagem,<sup>51-53</sup> o emprego de múltiplos rotâmeros para determinados resíduos,<sup>54,55</sup> a simulação do efeito indutivo de ligantes<sup>56</sup> e o emprego de métodos de minimização da estrutura de complexos proteína-ligante após a docagem.<sup>57</sup>

A presença de moléculas de água conservadas deve ser investigada nas estruturas dos complexos, uma vez que podem participar de interações intermoleculares cruciais para a formação do complexo proteína-ligante. Por sua vez, os resultados de estudos de docagem molecular são influenciados diretamente pela presença (ou ausência) de moléculas de água, levando a padrões distintos de reconhecimento molecular.<sup>58-60</sup> Para uma mesma proteína, ligantes díspares podem causar o deslocamento de diferentes moléculas de água,<sup>26</sup> devendo-se estimar para cada ligante quais as moléculas de água presentes no sítio de interação. No entanto, os altos custos computacionais, envolvidos para estimar as contribuições entálpicas e entrópicas no deslocamento de moléculas de água do sítio para o solvente, tornam inviável esta avaliação para um grande número de compostos. Alguns progressos nesta área têm sido observados, como o uso de programas que estimam posições mais favoráveis para moléculas de água no sítio de interesse,<sup>59,60</sup> ou que consideram flexíveis algumas moléculas de água, de modo que cada ligante possa ser docado na sua presença ou ausência.<sup>58,61</sup> Uma alternativa simples que pode ser adotada para a identificação de moléculas de água conservadas é a comparação de estruturas da proteína alvo em complexo com diferentes ligantes. Em geral, devido ao seu reconhecido papel na mediação de interações intermoleculares entre ligantes e receptores alvos, o processo de docagem é beneficiado pela inclusão de moléculas de água conservadas.

Outras considerações importantes são o estado de protonação dos resíduos da proteína, a posição dos hidrogênios em átomos polares de cadeias laterais e a conformação da amida dos resíduos glutamina e asparagina presentes no sítio de ligação. Como as coordenadas cristalográficas obtidas do PDB geralmente não contêm hidrogênios, os mesmos devem ser adicionados ao receptor previamente ao processo de docagem molecular. Neste procedimento, leva-se em conta a presença de doadores ou aceptores de ligações de hidrogênio tanto para posicionamento dos hidrogênios polares (*e.g.*, resíduos de cisteína, serina, tirosina e treonina), quanto para o estabelecimento da conformação da glutamina e asparagina. Para o estabelecimento do estado de protonação das cadeias laterais dos resíduos tituláveis (Tabela 2), o ambiente químico e o  $pK_a$  devem também ser analisados. Desta maneira, os resíduos de aspartato e glutamato são modelados geralmente com carga negativa, ao passo que os de lisina e arginina,

com carga positiva. Vale ressaltar que a histidina sempre requer atenção especial. Por apresentar  $pK_a$  próximo ao do pH fisiológico ( $pK_a \approx 6$  para o imidazol da cadeia lateral da histidina), a histidina pode ou não estar carregada positivamente, dependendo do ambiente químico ao seu redor. O imidazol, uma vez protonado, fica com dois grupos  $-NH$  e uma carga positiva que é distribuída nos dois nitrogênios, sendo representado por duas estruturas de ressonância. Além disso, nos casos em que a cadeia lateral da histidina se encontra na forma neutra, dois tautômeros podem estar presentes, diferindo na ligação de um hidrogênio a  $N_\delta$  ou  $N_\epsilon$  (Figura 4).

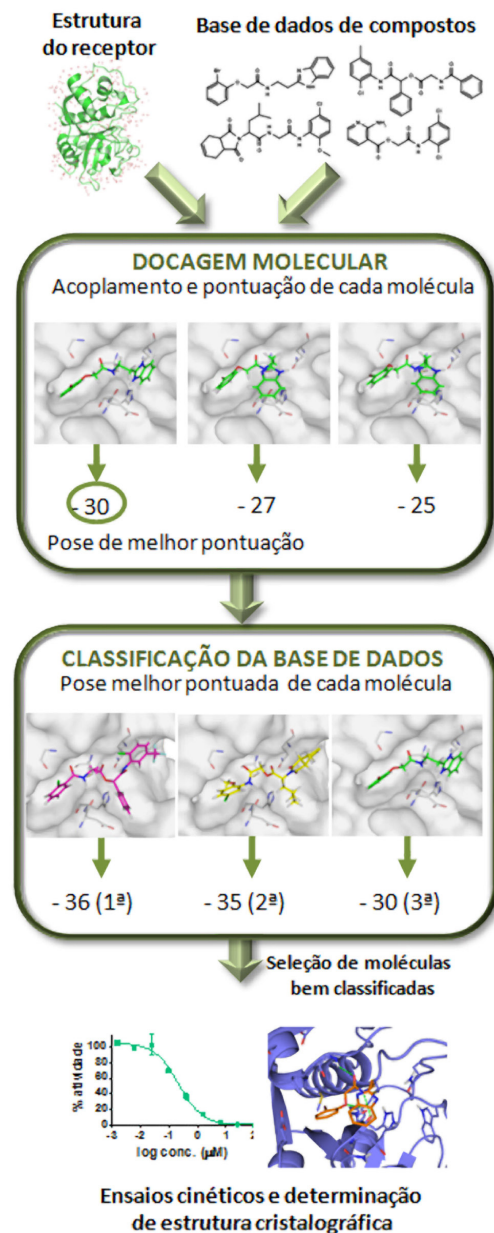
**Tabela 2.**  $pK_a$  e estado de protonação de cadeias laterais de resíduos tituláveis em pH fisiológico

Resíduo	$pK_a$ da cadeia lateral <sup>62</sup>	Carga usual em pH fisiológico
Aspartato	3,9	-1
Glutamato	4,3	-1
Arginina	12,0	+1
Lisina	11,1	+1
Histidina	6,4	0 ou +1
Cisteína	8,5	0
Tirosina	10,0	0

Mesmo os aminoácidos que se encontram carregados em pH fisiológico podem ter seu estado de protonação alterado, dependendo do ambiente químico no sítio da proteína.<sup>62</sup> São exemplos as proteases aspárticas, que possuem dois aspartatos catalíticos no sítio ativo. Apesar do baixo  $pK_a$  da cadeia lateral desse aminoácido ( $pK_a \approx 4$ ), apenas um dos resíduos se encontra carregado negativamente, enquanto o outro está na forma neutra (não ionizado).<sup>63</sup> Assim, a definição do estado de protonação dos resíduos proteicos envolve a apreciação de dados experimentais ou a predição através de programas computacionais desenvolvidos para este fim.<sup>62,64,65</sup>

### Etapas da SBVS

Depois de finalizadas as etapas de preparação dos ligantes e do receptor, o próximo passo da SBVS é a docagem molecular, na qual ocorre o acoplamento (*pose*) dos ligantes da base de dados de acordo com suas interações intermoleculares com os resíduos de aminoácidos do sítio de ligação da proteína alvo. Neste processo, as várias conformações e orientações geradas para cada molécula são classificadas de acordo com uma função de pontuação (*scoring function*) (Figura 5). Diversos métodos de acoplamento e funções de pontuação são disponíveis nos programas de docagem molecular mais comumente empregados.<sup>66-68</sup> De forma sucinta, o acoplamento de ligantes pode ser realizado a partir de conformações pré-calculadas, por construção incremental ou através de métodos em que são exploradas variações aleatórias de conformações e orientações (*e.g.*, algoritmos genéticos). Já as funções de pontuação podem ser físico-químicas, experimentais ou com base no conhecimento. As funções físico-químicas são compostas por termos derivados de princípios físicos, como o cálculo de interações eletrostáticas e de van der Waals, bem como por penalidades de dessolvatação, sem ajustes a dados experimentais. Nas funções experimentais, os diversos termos entálpicos e entrópicos envolvidos na associação proteína-ligante têm seu peso ajustado para reproduzir um conjunto experimental de dados. Por outro lado, as funções com base no conhecimento utilizam dados estatísticos provenientes de estruturas de complexos proteína-ligante para estabelecer parâmetros moleculares úteis, como distâncias, ângulos e interações favoráveis entre certos padrões atômicos.



**Figura 5.** Esquema da triagem virtual baseada na estrutura do receptor (SBVS)

Independentemente do programa de docagem molecular escolhido, ou da estratégia de planejamento de química medicinal adotada, busca-se identificar ligantes com reconhecimento molecular privilegiado em relação ao receptor alvo. As moléculas de melhor classificação são inspecionadas (*e.g.*, ocupação do sítio, complementaridade multifuncional, interações intermoleculares: ligações de hidrogênio, interações iônicas e de van der Waals<sup>69</sup> com o auxílio de programas de visualização molecular, sendo que a intuição, experiência e talento do pesquisador são elementos essenciais na seleção do conjunto final que será submetido à avaliação experimental (*e.g.*, ensaio *in vitro*). Dentro deste contexto, dois pontos particulares se mostram mais sensíveis: as moléculas selecionadas devem ser comercialmente disponíveis ou facilmente obtidas por síntese orgânica e, a avaliação experimental, os ensaios *in vitro* devem ser padronizados e validados. Deste modo, os *hits* computacionais podem se tornar ligantes reais da proteína alvo, tendo propriedades importantes determinadas, como potência, afinidade e seletividade. A investigação do mecanismo de ação é também parte importante destes estudos, agregando informações muito úteis.

O processo de validação experimental pode evoluir para a determinação de estruturas cristalográficas de complexos proteína-ligante, permitindo a elucidação do modo de ligação.<sup>70</sup> Estes elementos fornecem bases sólidas em química medicinal para o planejamento de novos ligantes com propriedades otimizadas.

Diversos exemplos de sucesso da aplicação da SBVS são descritos na literatura, demonstrando sua versatilidade, alto desempenho e grande utilidade em programas de descoberta fármacos.<sup>12,25,45,71-73</sup> É interessante observar a rica variedade de classes químicas identificadas para uma igualmente abastada diversidade de alvos biológicos.

### TRIAGEM EXPERIMENTAL OU VIRTUAL: OS NOVOS PARADIGMAS EM P&D DE FÁRMACOS

Os métodos de HTS e VS têm como objetivo comum a descoberta de moléculas bioativas com grande potencial de otimização molecular. Dentre os vários aspectos que devem ser levados em conta na escolha do método, destacam-se a informação, a mão de obra especializada, o investimento e a infraestrutura disponíveis. Por exemplo, a infraestrutura requerida para a aplicação de métodos experimentais ou computacionais é bastante distinta. A HTS envolve altos investimentos iniciais para a aquisição e preparação de soluções dos compostos, além da automatização dos ensaios para execução em alta escala.<sup>74</sup> Desta forma, esta técnica foi aplicada por muito tempo quase que exclusivamente pelas grandes indústrias farmacêuticas. Mais recentemente, as universidades ganharam acesso a esta estratégia através de algumas iniciativas públicas, podendo-se destacar os centros de triagem experimental do programa *NIH Roadmap* (*National Institutes of Health*, dos EUA),<sup>4,12,74-76</sup> No entanto, essa técnica permanece inacessível para grande parte da comunidade científica, especialmente nos países em desenvolvimento. Por sua vez, a VS exige investimentos e infraestrutura mais modestos e está, portanto, ao alcance de universidades, centros de pesquisa e pequenas empresas de todo mundo.

Assumindo-se um cenário ideal em que não houvesse limitações importantes de investimento, mão-de-obra e infraestrutura, uma questão fundamental seria como e porque escolher esta ou aquela técnica. Este tem sido um tema amplamente debatido através da comparação do desempenho das duas técnicas em uma série de casos.<sup>4,12-17</sup> Uma das vantagens do emprego da HTS é a seleção de compostos bioativos que, por sua vez, não seriam identificados na VS (caracterizando-se como falso-negativos). Em função de sua natureza não baseada no mecanismo, a HTS possibilita a identificação de compostos que se ligam a outros sítios da proteína alvo (além do sítio ativo),<sup>77,78</sup> o que não é possível no caso da VS que, por ser uma técnica baseada no mecanismo, envolve a avaliação de um único sítio de ligação específico do alvo. Em contrapartida, uma das grandes desvantagens da HTS é o número expressivo de falso-positivos, uma vez que os *hits* devem ser validados através de ensaios *in vitro* secundários, aumentando os custos e o tempo do processo.<sup>4,12</sup> Outra limitação é que classes químicas pouco representativas podem não ser priorizadas para ensaios secundários.<sup>12</sup> Por outro lado, uma das principais vantagens da VS

é a descoberta de um percentual maior de compostos bioativos em relação ao número de compostos avaliados experimentalmente.<sup>13,16</sup> Contudo, uma desvantagem desta técnica reside no fato de que classes interessantes de *hits* podem não se identificadas devido a sua baixa pontuação, ou ainda porque os compostos se ligam a cavidades distintas do sítio de ligação selecionado para a VS (Tabela 3).

Embora algumas análises comparativas entre as duas técnicas tenham relevado informações úteis, deve ser observado que na maioria dos casos as condições de comparação não foram as mais apropriadas. Por exemplo, em muitos casos a base de dados de compostos investigada era distinta para cada uma das técnicas.<sup>13,16</sup> Desta forma, os resultados refletiam não somente potencialidades e limitações inerentes às técnicas empregadas, mas, sobretudo, as características das bases de dados.

### INTEGRAÇÃO DAS TÉCNICAS DE TRIAGEM VIRTUAL E EXPERIMENTAL

Antes de mais nada, é importante salientar que o uso de uma técnica não exclui, necessariamente, o emprego da outra. A aplicação conjunta das técnicas de HTS e VS permite experimentar na prática a sinergia entre as duas abordagens em projetos de planejamento de fármacos. Existem basicamente duas estratégias que podem ser exploradas: o uso da VS como filtro precedendo a aplicação da HTS e, a aplicação em paralelo das duas técnicas, com a análise integrada de etapas e resultados (Figura 6). No primeiro caso, o emprego da VS proporciona a seleção de uma fração de compostos da base de dados para a avaliação experimental *in vitro*, levando a uma considerável otimização (*e.g.*, custos, tempo) na identificação de moléculas bioativas.<sup>15,17</sup> Por outro lado, um ponto negativo deste procedimento é que compostos mal classificados na VS não são avaliados experimentalmente, podendo levar à não identificação de possíveis *hits*. No segundo caso, os métodos podem ser realizados em paralelo, sendo os *hits* comuns priorizados para ensaios secundários *in vitro*.<sup>12,14</sup> Este procedimento tem como vantagem a possibilidade de identificar *hits* de HTS que normalmente não seriam priorizados para ensaios secundários, devido à baixa representatividade de sua classe química dentre as centenas de *hits* obtidos.<sup>12</sup> Além disso, em alguns casos a aplicação desses métodos em paralelo permite a identificação de falso-negativos de HTS, inicialmente não identificados por essa técnica devido a sua baixa afinidade pelo alvo,<sup>4</sup> ou de falso-negativos da VS, que não seriam considerados no primeiro caso, em que esta técnica é aplicada como filtro para a HTS.

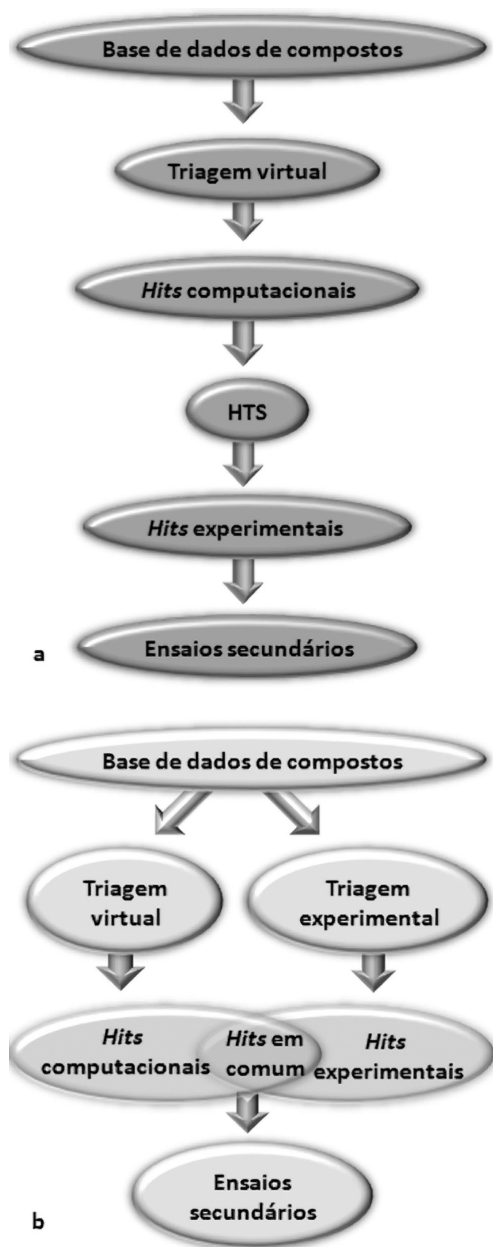
### CONCLUSÕES

A organização de bases de dados de compostos e o emprego de métodos de triagem em larga escala (experimental – HTS, virtual – VS) desempenham um papel essencial no processo de descoberta de novas moléculas bioativas para uma ampla variedade de alvos moleculares. A aplicação criteriosa destas estratégias constitui as bases científicas necessárias para o seu sucesso em química medicinal.

**Tabela 3.** Vantagens e desvantagens dos métodos de HTS e VS

Método	Vantagens	Desvantagens
HTS	Informação experimental sobre toda a base de dados	Alta porcentagem de falso-positivos
	Possibilidade de descoberta de inibidores que atuem em diversos sítios do receptor	Grande exigência de infraestrutura
	Independente de determinação prévia de estrutura tridimensional do alvo	
VS	Número de compostos a serem testados	Necessidade de modelo tridimensional do alvo
	Menor exigência de infraestrutura	Ligantes restritos a um determinado sítio de ligação
	Maior porcentagem de ligantes específicos entre compostos avaliados experimentalmente	Falso-negativos





**Figura 6.** Integração das estratégias de HTS e VS: a) uso da VS como filtro para selecionar compostos para avaliação por HTS; b) aplicação em conjunto das duas técnicas para priorização de hits para ensaios secundários *in vitro*

É preciso entender as características de cada caso a ser estudado e realizar um planejamento cuidadoso de todas as etapas envolvidas, buscando entender as suas limitações e vantagens. O impacto positivo do uso dos métodos de HTS e VS pode ser notado no número expressivo de trabalhos desenvolvidos ao longo das últimas décadas, que levaram à identificação e ao desenvolvimento de diversas moléculas com comprovado potencial terapêutico. A integração destes métodos deverá ser cada vez mais explorada pelos laboratórios de universidades e indústrias farmacêuticas, proporcionando novas oportunidades e desafios nesta fascinante área de descoberta de fármacos.

#### AGRADECIMENTOS

À FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) e ao CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) pelo apoio financeiro.

#### REFERÊNCIAS

- Munos, B.; *Nat. Rev. Drug Discov.* **2009**, *8*, 959.
- Lombardino, J. G.; Lowe III, J. A.; *Nat. Rev. Drug Discov.* **2004**, *3*, 853.
- Thorne, N.; Auld, D. S.; Inglese, J.; *Curr. Opin. Chem. Biol.* **2010**, *14*, 315.
- Babaoglu, K.; Simeonov, A.; Irwin, J. J.; Nelson, M. E.; Feng, B.; Thomas, C. J.; Cancian, L.; Costi, M. P.; Maltby, D. A.; Jadhav, A.; Inglese, J.; Austin, C. P.; Shoichet, B. K.; *J. Med. Chem.* **2008**, *51*, 2502.
- Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J.; *Adv. Drug Deliv. Rev.* **1997**, *23*, 3; Moda, T. L.; Torres, L. G.; Carrara, A. E.; Andricopulo, A. D.; *Bioinformatics* **2008**, *24*, 2270; Moda, T. L.; Montanari, C. A.; Andricopulo, A. D.; *Bioorg. Med. Chem.* **2007**, *15*, 7738.
- Chuprina, A.; Lukin, O.; Demoiseaux, R.; Buzko, A.; Shivanyuk, A.; *J. Chem. Inf. Model.* **2010**, *50*, 470.
- Teague, S. J.; Davis, A. M.; Leeson, P. D.; Oprea, T.; *Angew. Chem., Int. Ed.* **1999**, *38*, 3743.
- Rishton, G. M.; *Drug Discov. Today* **2003**, *8*, 86.
- Guido, R. V. C.; Oliva, G.; Andricopulo, A. D.; *Curr. Med. Chem.* **2008**, *15*, 37; Andricopulo, A. D.; Salum, L. B.; Abraham, D. J.; *Curr. Top. Med. Chem.* **2009**, *9*, 771.
- Leach, A. R.; Shoichet, B. K.; Peishoff, C. E.; *J. Med. Chem.* **2006**, *49*, 5851.
- Shoichet, B. K.; *Nature* **2004**, *432*, 862.
- Ferreira, R. S.; Simeonov, A.; Jadhav, A.; Eidam, O.; Mott, B. T.; Keiser, M. J.; McKerrow, J. H.; Maloney, D. J.; Irwin, J. J.; Shoichet, B. K.; *J. Med. Chem.* **2010**, *53*, 4891.
- Doman, T. N.; McGovern, S. L.; Witherbee, B. J.; Kasten, T. P.; Kurumbail, R.; Stallings, W. C.; Connolly, D. T.; Shoichet, B. K.; *J. Med. Chem.* **2002**, *45*, 2213.
- Polgar, T.; Baki, A.; Szendrei, G. I.; Keseru, G. M.; *J. Med. Chem.* **2005**, *48*, 7946.
- Bologa, C. G.; Revankar, C. M.; Young, S. M.; Edwards, B. S.; Arterburn, J. B.; Kiselyov, A. S.; Parker, M. A.; Tkachenko, S. E.; Savchuck, N. P.; Sklar, L. A.; Oprea, T. I.; Prossnitz, E. R.; *Nat. Chem. Biol.* **2006**, *2*, 207.
- Paiva, A. M.; Vanderwall, D. E.; Blanchard, J. S.; Kozarich, J. W.; Williamson, J. M.; Kelly, T. M.; *Biochim. Biophys. Acta* **2001**, *1545*, 67.
- Edwards, B. S.; Bologa, C.; Young, S. M.; Balakin, K. V.; Prossnitz, E. R.; Savchuck, N. P.; Sklar, L. A.; Oprea, T. I.; *Mol. Pharmacol.* **2005**, *68*, 1301.
- Overington, J. P.; Al-Lazikani, B.; Hopkins, A. L.; *Nat. Rev. Drug Discov.* **2006**, *5*, 993.
- Benson, J. D.; Chen, Y. N. P.; Cornell-Kennon, S. A.; Dorsch, M.; Kim, S.; Leszczyniecka, M.; Sellers, W. R.; Lengauer, C.; *Nature* **2006**, *441*, 451.
- Renslo, A. R.; McKerrow, J. H.; *Nat. Chem. Biol.* **2006**, *2*, 701; Castilho, M. S.; Postigo, M. P.; Pereira, H. M.; Oliva, G.; Andricopulo, A. D.; *Bioorg. Med. Chem.* **2010**, *18*, 1421; Dias, L. C.; Dessoy, M. A.; Silva, J. J. N.; Thiemann, O. H.; Oliva, G.; Andricopulo, A. D.; *Quim. Nova* **2009**, *32*, 2444.
- Zambrowicz, B. P.; Sands, A. T.; *Nat. Rev. Drug Discov.* **2003**, *2*, 38.
- Nomura, D. K.; Dix, M. M.; Cravatt, B. F.; *Nat. Rev. Cancer* **2010**, *10*, 630.
- Hert, J.; Irwin, J. J.; Laggner, C.; Keiser, M. J.; Shoichet, B. K.; *Nat. Chem. Biol.* **2009**, *5*, 479.
- Carr, R. A. E.; Congreve, M.; Murray, C. W.; Rees, D. C.; *Drug Discov. Today* **2005**, *10*, 987.
- Teotico, D. G.; Babaoglu, K.; Rocklin, G. J.; Ferreira, R. S.; Giannetti, A. M.; Shoichet, B. K.; *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 7455.
- Murray, C. W.; Carr, M. G.; Callaghan, O.; Chessari, G.; Congreve, M.; Cowan, S.; Coyle, J. E.; Downham, R.; Figueroa, E.; Frederickson, M.;



- Graham, B.; McMenamin, R.; O'Brien, M. A.; Patel, S.; Phillips, T. R.; Williams, G.; Woodhead, A. J.; Woolford, A. J. A.; *J. Med. Chem.* **2010**, *53*, 5942; Salum, L. B.; Andricopulo, A. D.; *Mol. Divers.* **2009**, *13*, 277; Salum, L. B.; Andricopulo, A. D.; *Expert Opin. Drug Discov.* **2010**, *5*, 405.
27. Chessari, G.; Woodhead, A. J.; *Drug Discov. Today* **2009**, *14*, 668.
28. Rishton, G. M.; *Drug Discov. Today* **1997**, *2*, 382.
29. Roche, O.; Schneider, P.; Zuegge, J.; Guba, W.; Kansy, M.; Alanine, A.; Bleicher, K.; Danel, F.; Gutknecht, E. M.; Rogers-Evans, M.; Neidhart, W.; Stalder, H.; Dillon, M.; Sjögren, E.; Fotouhi, N.; Gillespie, P.; Goodnow, R.; Harris, W.; Jones, P.; Taniguchi, M.; Tsujii, S.; von der Saal, W.; Zimmermann, G.; Schneider, G.; *J. Med. Chem.* **2002**, *45*, 137.
30. Andricopulo, A. D.; Yunes, R. A.; Nunes, R. J.; Savi, A. O.; Cruz, A. B.; Cechinel, V.; *Quim. Nova* **1998**, *21*, 573; Lima, E. O.; Queiroz, E. F.; Andricopulo, A. D.; Nunes, R. J.; Yunes, R. A.; Correa, R.; Cechinel, V.; *Bol. Soc. Chil. Quim.* **1999**, *44*, 185.
31. Orry, A. J. W.; Abagyan, R. A.; Cavasotto, C. N.; *Drug Discov. Today* **2006**, *11*, 261; Deng, Z.; Chuaqui, C.; Singh, J.; *J. Med. Chem.* **2006**, *49*, 490.
32. Roehrig, S.; Straub, A.; Pohlmann, J.; Lampe, T.; Pernerstorfer, J.; Schlemmer, K. H.; Reinemer, P.; Perzborn, E.; *J. Med. Chem.* **2005**, *48*, 5900.
33. Skulnick, H. I.; Johnson, P. D.; Howe, W. J.; Tomich, P. K.; Chong, K. T.; Watenpaugh, K. D.; Janakiraman, M. N.; Dolak, L. A.; McGrath, J. P.; Lynn, J. C.; Horng, M. M.; Hinshaw, R. R.; Zipp, G. L.; Ruwart, M. J.; Schwende, F. J.; Zhong, W. Z.; Padbury, G. E.; Dalga, R. J.; Shiou, L. H.; Possert, P. L.; Rush, B. D.; Wilkinson, K. F.; Howard, G. M.; Toth, L. N.; Williams, M. G.; Kakuk, T. J.; Cole, S. L.; Zaya, R. M.; Lovasz, K. D.; Morris, J. K.; Romines, K. R.; Thaisrivongs, S.; Aristoff, P. A.; *J. Med. Chem.* **1995**, *38*, 4968.
34. Feng, B. Y.; Simeonov, A.; Jadhav, A.; Babaoglu, K.; Inglese, J.; Shoichet, B. K.; Austin, C. P.; *J. Med. Chem.* **2007**, *50*, 2385.
35. Jadhav, A.; Ferreira, R. S.; Klumpp, C.; Mott, B. T.; Austin, C. P.; Inglese, J.; Thomas, C. J.; Maloney, D. J.; Shoichet, B. K.; Simeonov, A.; *J. Med. Chem.* **2010**, *53*, 37.
36. Guido, R. V. C.; Oliva, G.; Montanari, C. A.; Andricopulo, A. D.; *J. Chem. Inf. Model.* **2008**, *48*, 918.
37. Guido, R. V. C.; Castilho, M. S.; Mota, S. G. R.; Oliva, G.; Andricopulo, A. D.; *QSAR Comb. Sci.* **2008**, *27*, 768; Leitão, A.; Andricopulo, A. D.; Oliva, G.; Pupo, M. T.; Marchi, A. A.; Vieira, P. C.; Silva, M. F. G. F.; Ferreira, V. F.; Sa, M. M.; Montanari, C. A.; *Bioorg. Med. Chem. Lett.* **2004**, *14*, 2199.
38. Honorio, K. M.; Polikarpov, I.; Garratt, R.; Andricopulo, A. D.; *J. Mol. Graphics Modell.* **2007**, *25*, 921; Honorio, K. M.; Polikarpov, I.; Garratt, R.; Andricopulo, A. D.; *Lett. Drug Des. Discov.* **2006**, *3*, 261.
39. Salum, L. B.; Dias, L. C.; Andricopulo, A. D.; *J. Braz. Chem. Soc.* **2009**, *20*, 693; Salum, L. B.; Polikarpov, I.; Andricopulo, A. D.; *J. Chem. Inf. Model.* **2008**, *48*, 2243.
40. Eckert, H.; Bojorath, J.; *Drug Discov. Today* **2007**, *12*, 225.
41. Barreiro, G.; Guimaraes, C. R. W.; Tubert-Brohman, I.; Lyons, T. M.; Tirado-Rives, J.; Jorgensen, W. L.; *J. Chem. Inf. Model.* **2007**, *47*, 2416.
42. Weber, K. C.; Salum, L. B.; Honorio, K. M.; Andricopulo, A. D.; da Silva, A. B. F.; *Eur. J. Med. Chem.* **2010**, *45*, 1508.
43. Geppert, H.; Vogt, M.; Bajorath, J.; *Chem. Inf. Model.* **2010**, *50*, 205.
44. Guido, R.; Trossini, G.; Castilho, M.; Oliva, G.; Ferreira, E.; Andricopulo, A.; *J. Enz. Inhib. Med. Chem.* **2008**, *23*, 964.
45. Perez-Pineiro, R.; Burgos, A.; Jones, D. C.; Andrew, L. C.; Rodriguez, H.; Suarez, M.; Fairlamb, A. H.; Wishart, D. S.; *J. Med. Chem.* **2009**, *52*, 1670.
46. Hert, J.; Willett, P.; Wilton, D. J.; Acklin, P.; Azzaoui, K.; Jacoby, E.; Schuffenhauer, A.; *J. Chem. Inf. Model.* **2006**, *46*, 462.
47. Freitas, R. F.; Prokopczyk, I. M.; Zottis, A.; Oliva, G.; Andricopulo, A. D.; Trevisan, M. T. S.; Vilegas, W.; Silva, M. G. V.; Montanari, C. A.; *Bioorg. Med. Chem.* **2009**, *17*, 2476.
48. Irwin, J. J.; Shoichet, B. K.; *J. Chem. Inf. Model.* **2005**, *45*, 177.
49. ten Brink, T.; Exner, T. E.; *J. Chem. Inf. Model.* **2009**, *49*, 1535.
50. McGovern, S. L.; Shoichet, B. K.; *J. Med. Chem.* **2003**, *46*, 2895.
51. Wei, B. Q.; Weaver, L. H.; Ferrari, A. M.; Matthews, B. W.; Shoichet, B. K.; *J. Mol. Biol.* **2004**, *337*, 1161.
52. Claussen, H.; Buning, C.; Rarey, M.; Lengauer, T.; *J. Mol. Biol.* **2001**, *308*, 377.
53. Cavasotto, C. N.; Kovacs, J. A.; Abagyan, R. A.; *J. Am. Chem. Soc.* **2005**, *127*, 9632.
54. Jacobson, M. P.; Friesner, R. A.; Xiang, Z. X.; Honig, B.; *J. Mol. Biol.* **2002**, *320*, 597.
55. Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D.; *Proteins* **2003**, *52*, 609.
56. Sherman, W.; Day, T.; Jacobson, M. P.; Friesner, R. A.; Farid, R.; *J. Med. Chem.* **2006**, *49*, 534.
57. Graves, A. P.; Shivakumar, D. M.; Boyce, S. E.; Jacobson, M. P.; Case, D. A.; Shoichet, B. K.; *J. Mol. Biol.* **2008**, *377*, 914.
58. Huang, N.; Shoichet, B. K.; *J. Med. Chem.* **2008**, *51*, 4862.
59. Young, T.; Abel, R.; Kim, B.; Berne, B. J.; Friesner, R. A.; *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 808.
60. Michel, J.; Tirado-Rives, J.; Jorgensen, W. L.; *J. Phys. Chem. B* **2009**, *113*, 13337.
61. Verdonk, M. L.; Chessari, G.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Nissink, J. W. M.; Taylor, R. D.; Taylor, R.; *J. Med. Chem.* **2005**, *48*, 6504.
62. Li, X.; Jacobson, M. P.; Zhu, K.; Zhao, S. W.; Friesner, R. A.; *Proteins* **2007**, *66*, 824.
63. Friedman, R.; Caflich, A.; *FEBS Lett.* **2007**, *581*, 4120.
64. Vriend, G.; *J. Mol. Graphics* **1990**, *8*, 52.
65. Labute, P.; *Proteins* **2009**, *75*, 187.
66. Moïtessier, N.; Englebienne, P.; Lee, D.; Lawandi, J.; Corbeil, C. R.; *Br. J. Pharmacol.* **2008**, *153*, S7.
67. Gohlke, H.; Klebe, G.; *Angew. Chem., Int. Ed.* **2002**, *41*, 2644.
68. Brooijmans, N.; Kuntz, I. D.; *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 335.
69. Stahl, M.; Rarey, M.; *J. Med. Chem.* **2001**, *44*, 1035.
70. Kolb, P.; Irwin, J. J.; *Curr. Top. Med. Chem.* **2009**, *9*, 755.
71. Postigo, M. P.; Guido, R. V. C.; Oliva, G.; Castilho, M. S.; Pitta, I. D.; de Albuquerque, J. F. C.; Andricopulo, A. D.; *J. Chem. Inf. Model.* **2010**, *50*, 1693.
72. Kolb, P.; Ferreira, R. S.; Irwin, J. J.; Shoichet, B. K.; *Curr. Opin. Biotechnol.* **2009**, *20*, 429.
73. Kolb, P.; Rosenbaum, D. M.; Irwin, J. J.; Fung, J. J.; Kobilka, B. K.; Shoichet, B. K.; *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 6843.
74. Inglese, J.; Johnson, R. L.; Simeonov, A.; Xia, M.; Zheng, W.; Austin, C. P.; Auld, D. S.; *Nat. Chem. Biol.* **2007**, *3*, 466.
75. Michael, S.; Auld, D.; Klumpp, C.; Jadhav, A.; Zheng, W.; Thorne, N.; Austin, C.; Inglese, J.; Simeonov, A.; *Assay Drug Dev. Technol.* **2008**, *6*, 637.
76. Engel, J. C.; Ang, K. K. H.; Chen, S.; Arkin, M. R.; McKerrow, J. H.; Doyle, P. S.; *Antimicrob. Agents Chemother.* **2010**, *54*, 3326.
77. Smithson, D. C.; Lee, J.; Shelat, A. A.; Phillips, M. A.; Guy, R. K.; *J. Biol. Chem.* **2010**, *285*, 16771.
78. Andricopulo, A. D.; Akoachere, M. B.; Krogh, R.; Nickel, C.; McLeish, M. J.; Kenyon, G. L.; Arscott, L. D.; Williams, C. H.; Davioud-Charvet, E.; Becker, K.; *Bioorg. Med. Chem. Lett.* **2006**, *16*, 2283.