

CONSTRUÇÃO, OTIMIZAÇÃO E ANCORAGEM MOLECULAR DE SUBSTÂNCIAS BIOATIVAS EM BIOMACROMOLÉCULAS: UM TUTORIAL PRÁTICOVictor S. Batista^{a, b}, Renan L. Farias^b, Leonardo P. M. Simões^a e Nailton M. Nascimento-Júnior^{a, *}^aDepartamento de Bioquímica e Química Orgânica, Instituto de Química, Universidade Estadual Paulista, 14800-060 Araraquara – SP, Brasil^bDepartamento de Química Analítica, Físico-Química e Inorgânica, Instituto de Química, Universidade Estadual Paulista, 14800-060 Araraquara – SP, Brasil

Recebido em 27/05/2021; aceito em 23/08/2021; publicado na web em 15/09/2021

CONSTRUCTION, OPTIMIZATION AND MOLECULAR DOCKING OF BIOACTIVE SUBSTANCES IN BIOMACROMOLECULES: A PRACTICAL TUTORIAL. In the last two decades, increasing advances in molecular biology and instrumental analysis for solving macromolecular structures have amplified the applicability of cheminformatics in drug discovery. In particular, molecular docking, an *in silico* structural-based method, has prospered as an efficient tool for understanding molecular interactions that drives formation of stable receptor-ligand systems. Brazil affords many research groups working hard on this subject; however, there is a lack of material in the Portuguese language, teaching how to apply molecular docking methodologies. Herein, we define and discuss a simple and low-cost workflow for molecular docking evaluation, comprising software installation (supplemental material) and how to use them for construction of small-molecules, perform docking calculations, analysis of the results and the preparation of quality figures. Given that, we have used free software and web-servers, as well as *in silico* tools supported by national funding agencies.

Keywords: molecular docking tutorial; Biovia DSV; MOPAC2016; CCDC GOLD suit; PyMol.

INTRODUÇÃO

A interface entre os protocolos de expressão e purificação de proteínas, proteômica e técnicas de resolução estrutural (difração de raios X, ressonância magnética nuclear, microscopia eletrônica criogênica, espalhamento de raios X de baixo ângulo e laser de raios X de elétrons livres), tem contribuído significativamente com a identificação e resolução estrutural de alvos terapêuticos biomoleculares.¹⁻³ Muitos desses alvos contemplam os depósitos de suas coordenadas cartesianas atômicas em bancos de dados com acesso público (ex. *Protein Data Bank*).⁴

Como consequência, o uso de rotinas *in silico* nos estágios iniciais da descoberta de fármacos, estruturalmente específicos, tem recebido máxima relevância. Visualizar micromoléculas ligantes interagindo com seus alvos biológicos é parte fundamental da estratégia SBDD (*structure-based drug design*).⁵⁻⁸ Esses complexos fornecem detalhes acerca das variáveis que governam o processo de reconhecimento molecular nos sítios de ligação da biomacromolécula receptora. Além disso, quando predições computacionais complementam dados experimentais, obtêm-se caminhos mais sustentáveis no processo de descoberta de fármacos, tendo em vista a redução do uso de recursos, assim como a economia de tempo na triagem e identificação de compostos líderes.^{6,9-11}

A ancoragem molecular (A.M.), também designada como atracamento ou docagem, é uma metodologia *in silico* essencial nessa perspectiva, uma vez que fornece o prognóstico dos modos de ligação de micromoléculas (ou macromoléculas) ao interagirem com um sítio receptor.¹² As poses (conformações e orientações) do ligante são encontradas por algoritmos de busca, enquanto que a classificação dessas soluções é conduzida em termos de afinidade pelo sítio de ligação via funções de pontuação (ou funções de *score*).^{2,12,13}

Os valores numéricos obtidos pelas funções de pontuação são decorrentes da modelagem matemática das interações, que ocorrem durante a formação de complexos estáveis.^{2,14-16} Por sua vez, as funções de pontuação podem ser empíricas, baseadas em campos de força e baseadas no conhecimento. Contudo, as simplificações matemáticas que deixam os cálculos de A.M. mais acelerados tornam as predições de afinidade desses complexos menos acuradas.¹⁷ Consequentemente, para investigar detalhes energéticos das interações/ligações do sistema receptor-ligante, outros métodos computacionais devem ser combinados. Uma prática usual, e confiável, consiste na junção entre a A.M. e a Dinâmica Molecular (D.M.), aplicada ao complexo exportado.^{2,3,9,18}

Em relação aos algoritmos de busca, se tem pelo menos duas categorias básicas: sistemática e estocástica.¹⁹ Os algoritmos de busca “sistemáticos” são determinísticos, modelam o espaço de busca em intervalos limitados (predefinidos). Os algoritmos estocásticos amostram o espaço de busca por alterações aleatórias nas variáveis de estado até que o critério de término seja atendido.^{12,19} Atualmente, os algoritmos estocásticos são considerados pela comunidade científica como mais hábeis na obtenção de soluções representativas, todavia Santos e colaboradores³ afirmam que a comparação entre diferentes algoritmos de busca em *softwares* de A.M. não é uma tarefa linear, haja vista que diferentes métodos podem ser mais eficientes para uma classe de receptores que outras.

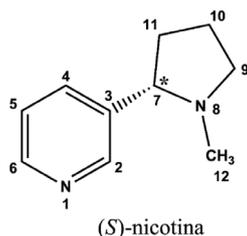
Entre os métodos estocásticos, podem ser destacados os algoritmos genéticos (A.G.).^{15,17,19-21} Um A.G. leva essa alcunha por mimetizar o processo de evolução natural. Primeiramente, uma pose é qualificada como cromossomo, o que descreve a posição do ligante no espaço de busca. Uma população inicial de cromossomos é gerada aleatoriamente. Então, a população é submetida a uma série de operações genéticas: mutações (mudanças aleatórias nos cromossomos) e recombinações (mistura entre cromossomos), produzindo novas gerações de poses, as quais são inspecionadas por funções de pontuação que selecionam as mais aptas ao sítio receptor. A rotina é concluída após o desenvolvimento de um número adequado de gerações, parâmetro que é pré-definido pelo usuário.

*e-mail: nailton.monteiro@unesp.br

^aLaboratório de Química Medicinal, Síntese Orgânica e Modelagem Molecular (LaQMedSOMM)

A A.M. pode ser inserida em diversos seguimentos: processo de descoberta de fármacos, incluindo estudos de QSAR, triagem virtual em larga escala, elaboração de bibliotecas combinatorias, hipóteses de ligação para pesquisas em mutagenese, cristalografia de raios X no ajuste de substratos e antagonistas, dentre outros. No Brasil, existem muitos grupos de pesquisa importantes dentro dessa temática e que utilizam essas ferramentas em suas investigações. Apesar disso, nota-se que são insuficientes os materiais, no idioma português, sobre o *modus operandi* de como aplicar metodologias de A.M.

Portanto, nesta nota técnica, está elaborado um fluxo de trabalho simples para avaliação de um complexo receptor-ligante, por A.M., utilizando como modelo a estrutura do receptor nicotínico de acetilcolina humano $\alpha 4\beta 2$ (PDB ID: 5kxi), resolvida por difração de raios X (3,9 Å), complexada à molécula de nicotina ($K_i = 18,2 \pm 5,08 \text{ nmol L}^{-1}$).²² Esquema 1.



Esquema 1. Representação esquemática 2D da estrutura química da nicotina em forma de linhas e identificação numérica dos átomos

Os receptores nicotínicos de acetilcolina (nAChRs) são receptores do neurotransmissor endógeno acetilcolina, sendo divididos em dois grupos: musculares e neuronais.²² O receptor neuronal subtipo $\alpha 4\beta 2$ é majoritário no cérebro de mamíferos, sendo altamente sensível a exposição da molécula de nicotina (agonista). Outrossim, devido as suas funções biológicas estarem relacionadas com diversas desordens neurológicas, como Alzheimer, Parkinson, esquizofrenia e principalmente dependência química à nicotina, o entendimento das propriedades de interação receptor-ligante é fundamental na busca por formas terapêuticas.²²

Visando um público de novos usuários, o texto compreende a instalação dos softwares utilizados (ver seção material suplementar 1S), obtenção e preparo da biomacromolécula e do ligante, realização dos cálculos de A.M., análise de resultados e o preparo de figuras para publicação. Para execução desse trabalho foram empregados softwares e servidores online de livre acesso e/ou versões acadêmicas. Não obstante, o pacote CSDS SYSTEM, contendo os softwares privados Mercury,²³ Hermes e GOLD,^{15,17} foi obtido no domínio (dotlib.com.br), segundo a política de acesso vinculada ao Portal Periódicos CAPES.

É importante salientar que esse tutorial prático também possui o objetivo de atingir profissionais de outras áreas do conhecimento, debruçando-se em uma narrativa mais abrangente. Consequentemente, não serão abordados à fundo os aspectos teóricos, acerca dos problemas associados ao processo químico computacional. Advertimos, ainda, que a metodologia foi elaborada para utilização no sistema operacional Windows 10 (64 bits). Por fim, é seguramente aconselhado que, apesar da metodologia elaborada ser útil como um guia inicial, existe espaço para quaisquer aprimoramentos e/ou modificações, que visem atender, de forma mais adequada, outros problemas de A.M.

PARTE EXPERIMENTAL

Aquisição de estrutura 3D da biomacromolécula a partir do servidor online Protein Data Bank (PDB ID: 5kxi)

Ao acessar o domínio <https://www.rcsb.org/> na barra de busca,

deve-se inserir o código PDB 5kxi.²² Uma vez redirecionado para a página que contém um resumo de informações sobre essa estrutura, o arquivo **PDB FORMAT** na opção **DOWNLOAD FILES** pode ser descarregado. Salve o arquivo (5kxi.pdb) em uma pasta criada especificamente para o trabalho.

Análise das interações na estrutura cristalográfica com o servidor online PDBsum

Para avaliar as interações originais ligante-receptor, presentes na estrutura da biomacromolécula alvo, acesse o domínio <http://www.ebi.ac.uk/pdbsum/>, digite o código PDB 5kxi e clique em **FIND**.²⁴ Em seguida, na barra superior, se deve clicar na aba **LIGANDS** e selecionar o ligante que se deseja obter as interações com o sítio ativo [**NCT 402 (A)**]. Para descarregar o arquivo *.pdf, contendo tais informações, clique com o botão direito do mouse sobre o símbolo PDF, em seguida clique em **salvar link como**, e escolha a pasta de trabalho.

Avaliação da qualidade da estrutura cristalográfica com o servidor online PROCHECK

Acesse o domínio <https://saves.mbi.ucla.edu/>. Faça **upload** da estrutura cristalográfica obtida na etapa anterior (5kxi.pdb) e clique em **RUN PROGRAMS**. Em seguida, selecione a opção **PROCHECK**²⁵ e clique em **START**. Ao término do processo, clique em **RESULTS**. Após exibição dos resultados de análise, clique em **RAMACHANDRAN PLOT**.²⁶ Nesse momento, devem ser observados os resíduos em regiões permitidas ou "outliers", assim como o percentual de resíduos em cada uma dessas regiões (*favoured, allowed and outlier*). É possível fazer o download do arquivo *.pdf contendo estas informações clicando no link **PDF**, localizado acima do gráfico de Ramachandran.

Construção de ligantes utilizando o software BIOVIA Discovery Studio Visualizer

Abra o software Discovery Studio Visualizer (DSV)²⁷ e crie uma nova janela de trabalho **FILE > NEW > MOLECULE WINDOW**. Clique na guia **SMALL MOLECULES** e selecione a opção **SKETCH**, localizada na guia **SKETCH MOLECULES** do painel à esquerda da tela de trabalho.

Obs.: Para instruções/comandos básicos sobre a construção de moléculas, utilizando o DSV, ver seção material suplementar 2S.

Construção da molécula de nicotina

- i) Com a opção **SKETCH** selecionada, desenhe o anel piridínico. Para criar ligações duplas, selecione a ligação desejada e aperte a tecla 2 no teclado (3 para ligações triplas). Alternativamente, com a opção **SKETCH** selecionada, clique sobre a ligação desejada (dois cliques para ligações triplas). Ainda, é possível selecionar todas as ligações simultaneamente com a tecla **shift** e mudar seus atributos em um único comando. Deve-se clicar com o botão direito do mouse sobre a seleção e pressionar a opção **BONDS**, localizando a lista de ligações disponíveis;
- ii) Para mudar um átomo de carbono por nitrogênio, selecione o átomo desejado e pressione a tecla "N". Neste caso, a função **SELECT** deve estar habilitada. Também é possível fazer essa mudança clicando com o botão direito sobre um átomo já selecionado e alterando-o em **ELEMENT**;
- iii) À esquerda da área de edição, clique em **H ADD** para adicionar hidrogênios e em **CLEAN GEOMETRY** para otimizar os ângulos

e as distâncias da molécula. O comando **CLEAN GEOMETRY** pode ser utilizado para otimizar uma subunidade da estrutura em particular;

- iv) Com a opção **SKETCH** ainda selecionada, clique no átomo de hidrogênio da posição 3 do anel piridínico e transmute-o para carbono, permitindo assim a construção do anel pirrolidínico. A seguir, adicione os hidrogênios utilizando a função **H ADD** (Figura 1).
- v) Clique em dois átomos de hidrogênio desse novo átomo de carbono para transmutá-los em mais dois átomos de carbono (seguido da adição dos hidrogênios). Faça isso mais uma vez, até que tenhamos 5 átomos de carbono, além do anel piridínico. Para fechar o ciclo, clique no primeiro átomo desejado e em seguida no segundo (ambos os cliques com a função **SKETCH** selecionada). Em seguida, clique em **CLEAN GEOMETRY**; **Atenção!** O hidrogênio ligado ao carbono terciário (C7) deve estar para cima em relação ao nitrogênio piridínico (N1). Selecione a função **TORSION** e altere o ângulo diedro formado pelos átomos “A”, “B”, “C” e “D”, mantendo o botão esquerdo do mouse pressionado (Figura 2A).
- vi) Selecione o átomo de carbono (sp^3) da posição 8, ligado diretamente ao carbono terciário, e transmute-o em nitrogênio. Com a função **SELECT** habilitada, selecione o nitrogênio criado e clique sobre o mesmo com o botão direito do mouse. Em seguida, vá até a aba **CHARGE** e selecione a opção +1; **Atenção!** Quando um átomo de nitrogênio em configuração eletrônica sp^3 forma três ligações, o mesmo possuirá um par de elétrons livres e carga zero. Contudo, no caso em questão, o átomo de nitrogênio N8 está protonado. Logo, esse átomo contribuiu com seu par isolado para formar uma quarta ligação, essa terá a natureza de ligação covalente coordenada. Como consequência, o N8 deve receber uma carga positiva, embora o mesmo seja mais eletronegativo que o átomo de H e o átomo de carbono C12 diretamente ligados a ele.
- vii) Com a função **SELECT** habilitada, clique no átomo de hidrogênio mais alto, ligado ao nitrogênio pirrolidínico N8, aqui referimo-nos ao átomo de H que está direcionado para o usuário conforme

indicado na Figura 2B, e transmute-o em carbono, seguido da adição de hidrogênios, Figura 2C;

Atenção! Para melhor representar a nicotina, isso é, na mesma orientação e conformação encontrada na estrutura cristalográfica (PDB ID: 5kxi),²² é importante que o átomo de hidrogênio pirrolidínico, ligado ao nitrogênio, esteja direcionado para baixo do anel.

- viii) Sugere-se uma mudança no nome do ligante para “NCT” a fim de facilitar a análise posterior dos resultados de ancoragem molecular. Para isso, clique na seta que abre a janela de hierarquia e clique duas vezes sobre o nome da molécula. Altere o nome para NCT e pressione a tecla **Enter**. Em seguida, selecione toda a estrutura e utilize o comando **CLEAN GEOMETRY**. Salve a molécula em formato *.mol2 (NCT.mol2) em **FILE > SAVE**.

Espeiação *in silico* do estado de protonação do ligante, utilizando o MarvinSketch

Abra o *software* MarvinSketch²⁸ e construa a molécula de nicotina (ver material suplementar, seção 3S). Para simular o estado de protonação da nicotina em pH fisiológico, clique na aba **CALCULATIONS > PROTONATION > pKa**. Não faça nenhuma alteração na janela que se abre e pressione **OK > OK**. Selecione e mantenha pressionada a barra de rolagem até encontrar o pH 7,4, considerado como o valor de pH fisiológico padrão para esse tipo de predição. Observe que a micro-espécie mais significativa é a segunda (Figura 3), que corresponde à nicotina protonada no anel pirrolidínico.

Otimização geométrica do ligante, por meio dos softwares CCDC Mercury e MOPAC2016

O cálculo de otimização geométrica da molécula de nicotina será conduzido em nível semi-empírico, utilizando o pacote MOPAC2016,²⁹ a partir da interface gráfica do *software* CCDC Mercury.²³ Assim, carregue no Mercury, o arquivo NCT.mol2. Para melhor visualização, sugere-se utilizar o estilo “bolas e bastões”,

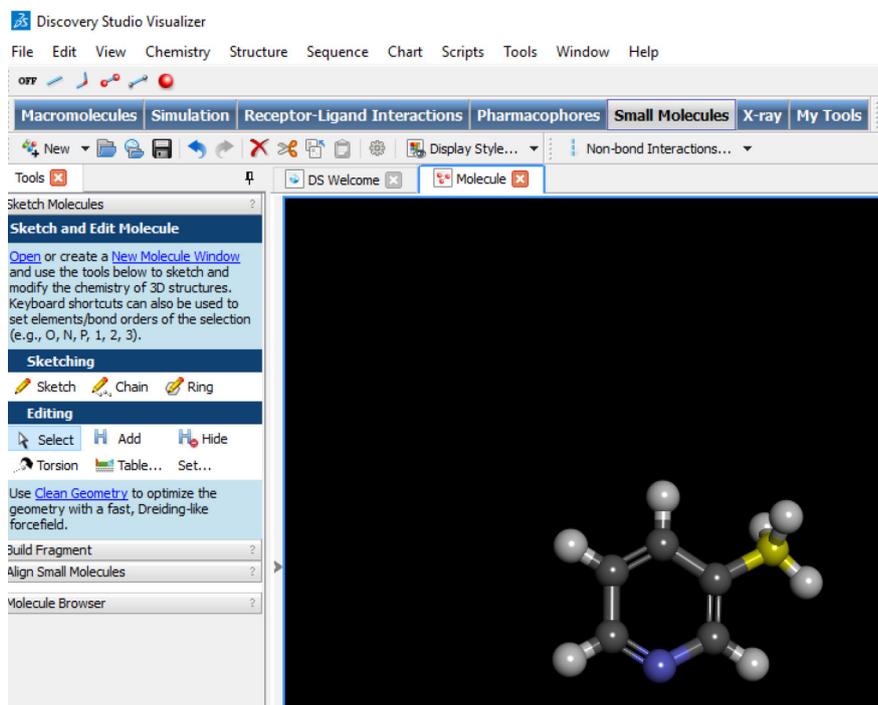


Figura 1. Posição do átomo de hidrogênio, transmutado para carbono (em amarelo, quando selecionado)

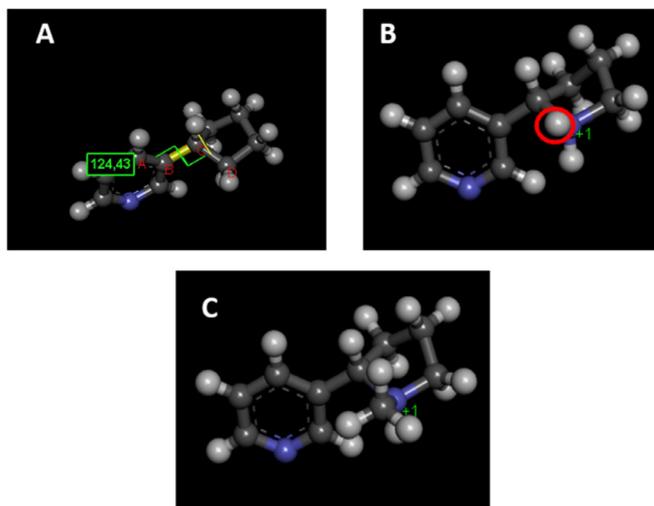


Figura 2. (A) Átomos envolvidos no ângulo diedro (ângulo de torção) entre os anéis da estrutura proposta. (B) Indicação do átomo de hidrogênio do nitrogênio N8, o qual deverá ser transmutado para carbono (C12) com posterior adição de hidrogênios. (C) estrutura da nicotina resultante

seleccione **DISPLAY > STYLES > BALL AND STICK**, na barra de ferramentas. Em seguida, para editar o arquivo de entrada do MOPAC2016, pressione consecutivamente **CSD-MATERIALS > CALCULATIONS > MOPAC** (Figura 4).

Na opção **MOPAC** clique em reticências [...] (Figura 5), e carregue o arquivo executável MOPAC2016.exe, previamente instalado.

Para iniciar os cálculos de otimização geométrica, seleccione o hamiltoniano PM7 na opção **HAMILTONIAN**.³⁰ Marque as opções **OPTIMISE GEOMETRY** e **CALCULATE BOND ORDERS**. Digite os comandos **"CHARGE=+1"** e **"XYZ"** na caixa de texto **EXTRA KEYWORDS** e clique em **SUBMIT**. Escolha um local para salvar

os resultados do cálculo, com extensão *.mol2. Sugere-se uma nova pasta, por exemplo **"ligantes_do_tutorial_PM7"**.

Ancoragem molecular

Redocagem da nicotina por meio do programa CCDC GOLD

- i) Inicialmente, inicie o *software* GOLD.^{15,17} Automaticamente, a interface gráfica Hermes será acionada. Para realizar o carregamento da proteína de interesse siga **WIZARD > PROTEINS > LOAD PROTEIN** e seleccione o arquivo 5kxi.pdb. Nessa mesma janela, marque a opção **FIX ALL PROTEIN ROTABLE BONDS**, para trabalhar com a ancoragem molecular em modo rígido. Em seguida, ainda na mesma janela, clique na aba **5KXI** ao lado de **GLOBAL OPTIONS** e seleccione a opção **PROTONATION & TAUTOMERS**, clicando em seguida no botão **ADD HYDROGENS**, conforme demonstrado na Figura 6.
- ii) Em seguida, na opção **EXTRACT/DELETE WATERS**, delete todas as moléculas de água. (**Obs.:** Em muitos casos, moléculas de água funcionam como ponte, que medeiam interações, entre o ligante e resíduos-chave. Sendo assim, essas "águas" são importantes no cálculo de *docking* e sugerimos fortemente que seja realizada, previamente, uma inspeção visual, para confirmação. Ao confirmar a inexistência de tal condição, pode-se executar o procedimento de exclusão de todas as moléculas de água co-cristalizadas).
- iii) Após esse procedimento, na opção **DELETE LIGANDS**, marque todas as caixas de texto dos ligantes a serem removidos e clique na opção **EXTRACT**. Para cada ligante, o programa solicitará uma identificação e o diretório no qual o mesmo será salvo. Aqui é recomendado usar os nomes originais para os ligantes, salvando o ligante "A_1" em um local de fácil acesso, tendo em vista que ele será usado como referência para os cálculos dos desvios (RMSD), na etapa de redocagem.

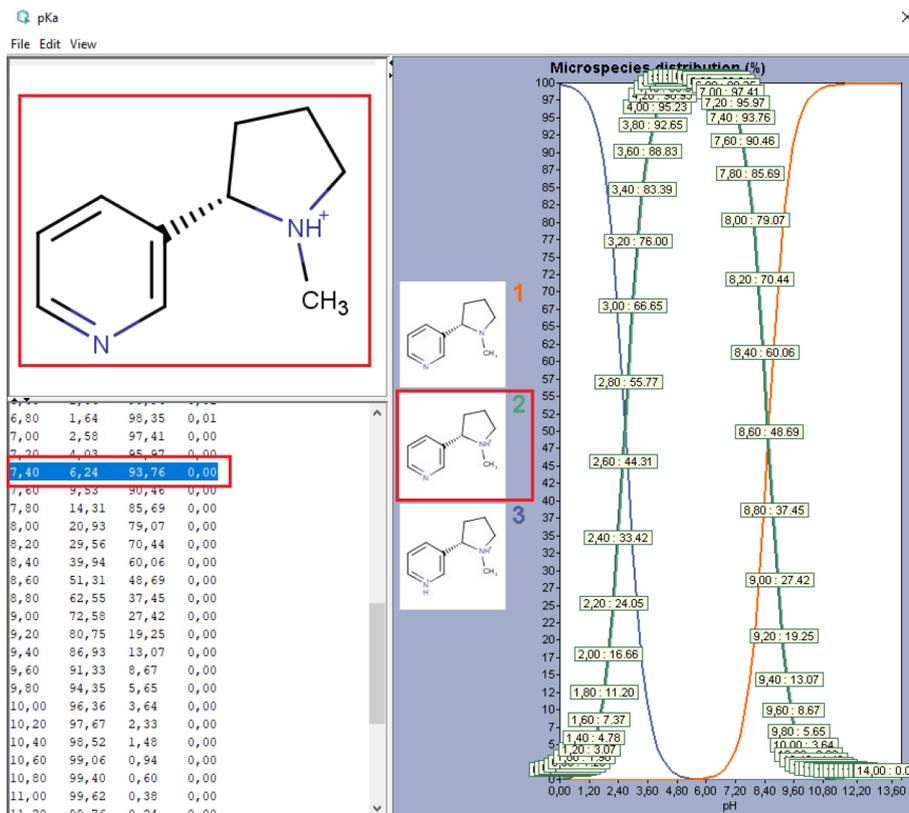


Figura 3. Janela com os resultados do cálculo de pKa

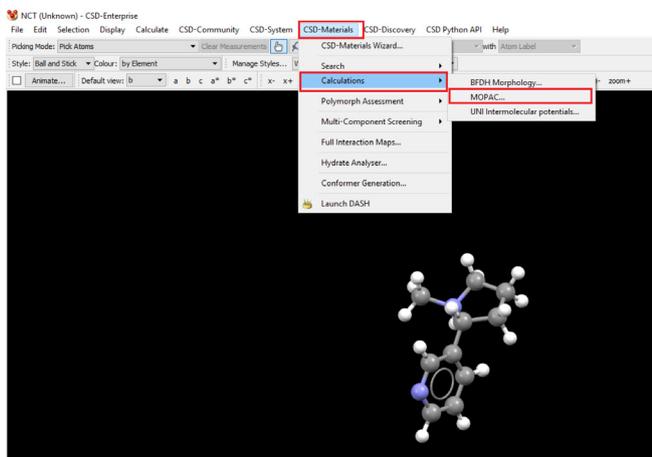


Figura 4. Iniciar o MOPAC pela interface do Mercury

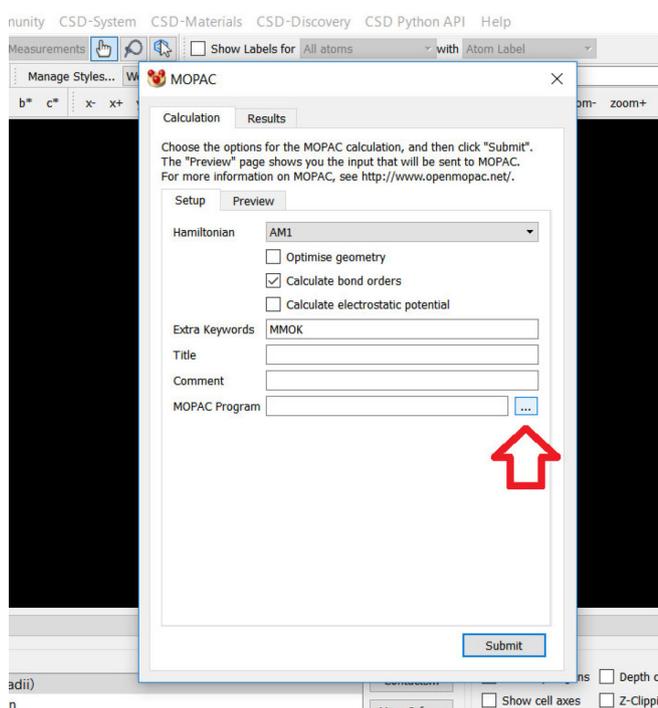


Figura 5. Carregar programa MOPAC pela interface do Mercury

iv) Retorne para a aba GLOBAL OPTIONS e clique na opção DEFINE BINDING SITE. Nessa janela, marque a opção POINT. Digite as coordenadas $x = 67.3680$, $y = -27.2034$ e $z = -39.0781$ e clique em VIEW. (Obs.: As coordenadas x , y , z são definidas como a centro da esfera de simulações, a qual contempla o sítio de ligação. Neste tutorial escolheu-se um ponto que coincide exatamente com a posição cartesiana do átomo de nitrogênio pirrolidínico (N8) do ligante nicotina que está co-cristalizado na estrutura da biomacromolécula PDB ID 5kxi. Todavia, é importante destacar que para cada sistema se deve analisar a melhor posição com base, principalmente, em critérios experimentais). É possível alterar o raio a ser considerado para o sítio de ligação na caixa de texto **SELECT ALL ATOMS WITHIN**. Aqui, recomenda-se o raio de 10 Å para o cálculo. (Obs.: É possível medir o tamanho de uma molécula dentro do próprio Hermes. Para isso, clique na opção **SHOW HYDROGENS** na parte superior do software e em seguida mude o **PICKING MODE** para **MEASURE DISTANCES**. Feito

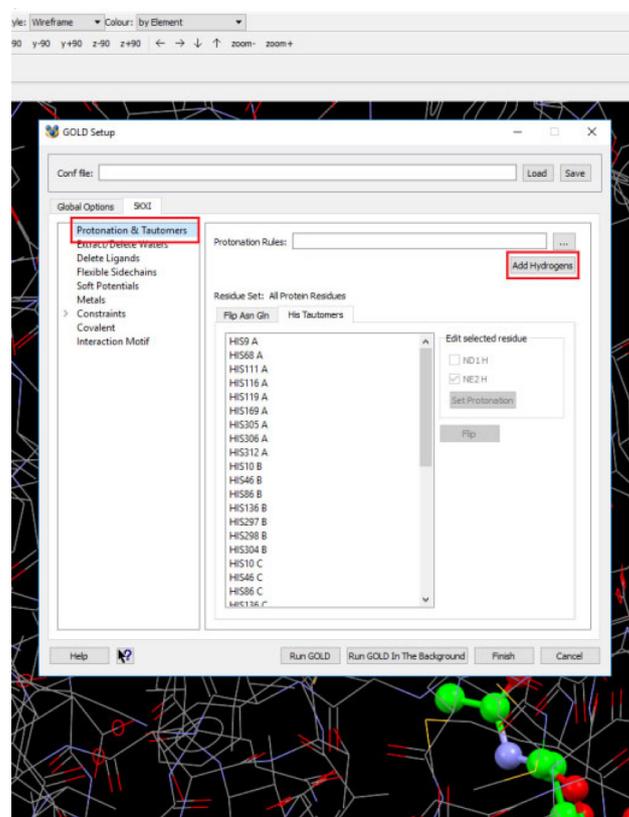


Figura 6. Adicionar hidrogênios à estrutura da proteína

isso, basta clicar sobre os dois átomos cuja distância se quer medir e o valor em angstrom aparecerá acima de uma linha tracejada na tela). O valor da distância medida pode ser verificado na forma de lista em **DISPLAY > MORE INFORMATION > DISTANCE LIST**.

- v) Alternativamente, é possível definir o sítio de ligação com base na posição de algum átomo de referência que se julgue adequado. No caso apresentado, o centro será definido nas coordenadas do nitrogênio pirrolidínico, do ligante de referência (“A_1”). **Atenção!** É importante salientar que uma vez escolhida a referência para uma redocagem, deve-se mantê-la durante as ancoragens subsequentes (por exemplo, se estivermos usando “A_1” como referência é imperativo **NÃO** usar “D_1” para definir o sítio de ligação para outras simulações, dentro da mesma rotina de cálculo). Essa observação é desnecessária para cálculos subsequentes que utilizam o mesmo arquivo *.conf gerado pela redocagem, entretanto, é de extrema importância na ocasião de ancoragens que sejam programadas do início, ou seja, novos arquivos *.conf.
- vi) Na aba **DEFINE BINDING SITE**, com a opção **POINT** selecionada, minimize a janela **GOLD Setup** a fim de deixar visível a estrutura tridimensional da proteína na interface gráfica. Utilize os botões do mouse, esquerdo (rotacionar) e direito (aproximar ou afastar), para fazer com que toda a extensão da proteína fique visível no Hermes, (Figura 7):
- vii) Dentro da janela **MOLECULE EXPLORER**, na aba **DISPLAY**, é possível observar que o *software* ainda exhibe os ligantes removidos. Clique duas vezes na *checkbox* ao lado de **ALL ENTRIES**, para que o Hermes esconda todas as estruturas. Em seguida, clique sobre a *checkbox* ao lado de “A_1”.
- viii) Utilize o mouse para centralizar e aproximar a molécula visualizada, a fim de facilitar a seleção do átomo correto. Em sequência, clique no átomo de nitrogênio do anel pirrolidínico do ligante para defini-lo como centro do sítio de ligação. Feito isso, maxi-

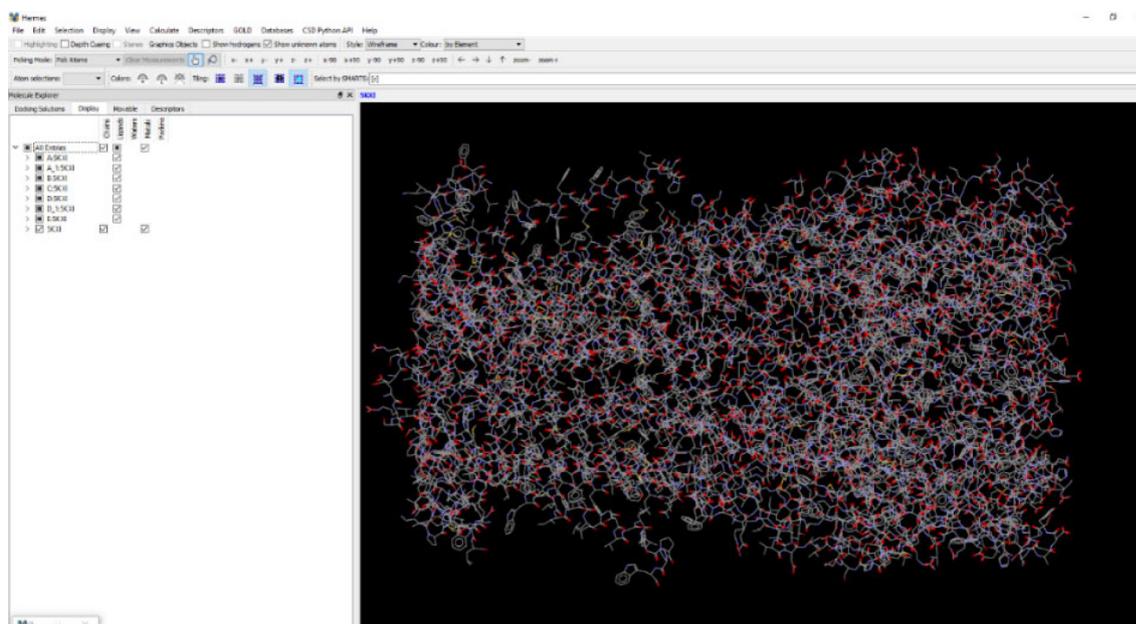


Figura 7. Visualização da estrutura da proteína no GOLD

mize a janela **GOLD setup** que deve estar localizada abaixo do **MOLECULE EXPLORER**. Se a definição for bem sucedida, uma esfera vermelha poderá ser visualizada ao redor do ponto em questão, quando a janela estiver maximizada. Além disso, as coordenadas obtidas devem ser as mesmas já mencionadas neste tutorial, mostradas na Figura 8.

- ix) Após a definição do sítio de ligação, dirija-se à opção **SELECT LIGANDS** e clique em **ADD**, para adicionar a molécula de nicotina, previamente otimizada (NCT.mol2). Idealmente, trabalha-se entre 50 e 100 poses de ancoragem, contudo, para esta nota técnica, selecione 10 poses de ancoragem. Posteriormente, clique nas reticências [...], ao lado de **REFERENCE LIGAND**, e selecione como referência o ligante extraído **A_1**.
- x) Vá para a opção **FITNESS & SEARCH OPTIONS** e desmarque a opção **ALLOW EARLY TERMINATION** e selecione a função de pontuação CHEMPLP, em **SCORING FUNCTION**. O programa

GOLD contém 4 funções básicas de pontuação: GoldScore, ChemScore, CHEMPLP e ASP, sendo possível utilizá-las de forma exclusiva ou em conjunto, dependendo do objetivo do usuário. As funções são parametrizadas distintamente, de modo que é sempre necessário avaliar qual delas será mais eficiente para o sistema estudado.³¹

- xi) Em seguida, clique na opção **RUN GOLD**

Na janela que se abre, selecione a pasta na qual devem ser salvos os resultados (ex.: docking_tutorial) e marque a caixa **PROTEIN** (caso a mesma esteja desmarcada) para salvar a estrutura preparada para o cálculo. Clique em **SAVE** para iniciar o cálculo, conforme mostrado na Figura 9.

Uma vez finalizado o cálculo, clique em **VIEW SOLUTIONS**. Em seguida, clique em **CLOSE** para fechar a janela do cálculo, fazendo o mesmo para a janela de programação do cálculo, deixando apenas a interface, Hermes, aberta. Os resultados

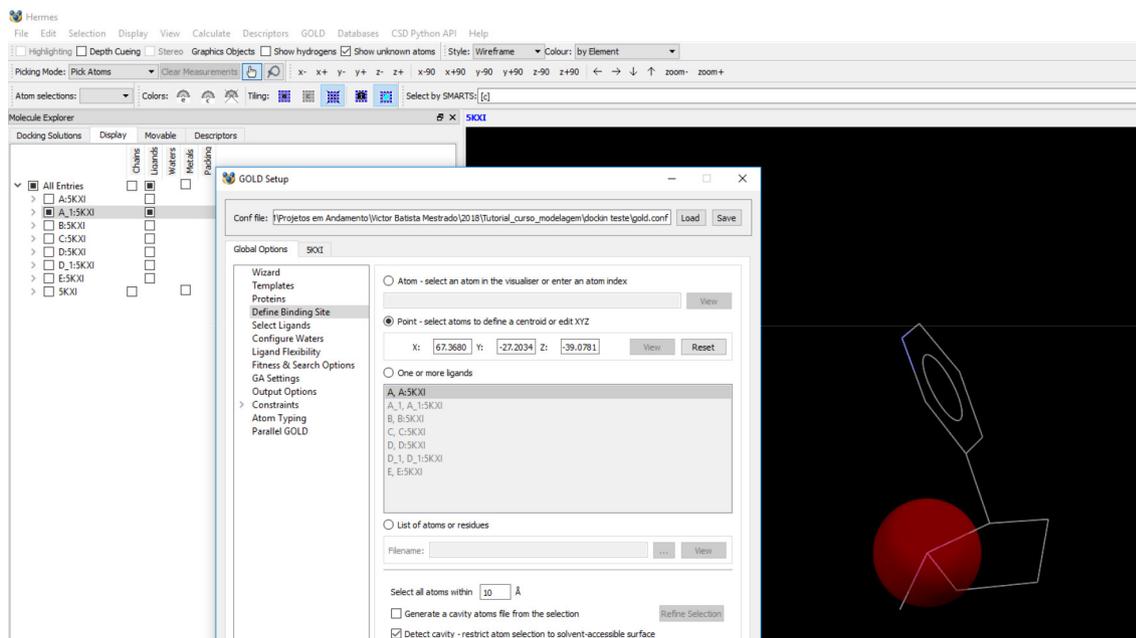


Figura 8. Ponto central da esfera de interação definido como sítio ativo da simulação

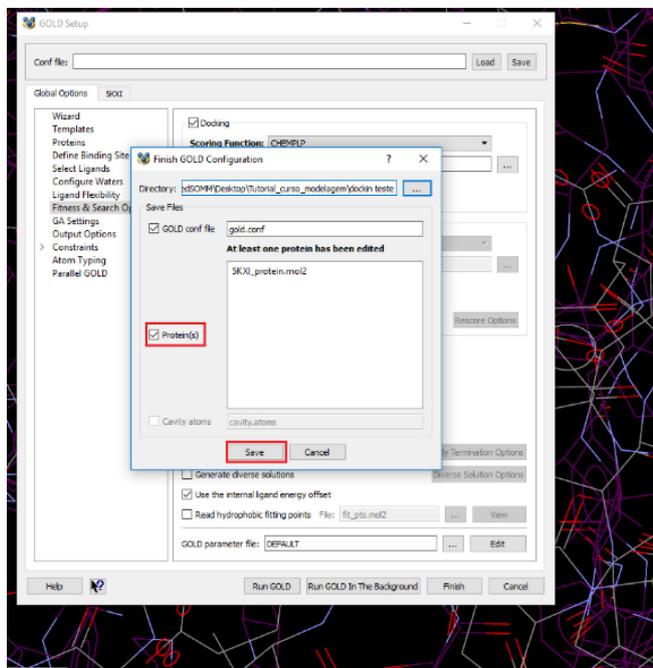


Figura 9. Confirmar início do cálculo das poses de ancoragem

serão mostrados na janela **MOLECULE EXPLORER**, na aba **DOCKING SOLUTIONS**, que deverá ser mostrada, automaticamente, após o término do cálculo. Os valores dos desvios médios quadráticos (em Å) obtidos entre as poses da molécula NTC e a referência **A_1** podem ser visualizados arrastando a barra de rolagem dessa janela para o lado direito, Figura 10.

xii) Seleccione o resultado com a maior pontuação (PLP.Fitness) e clique em **FILE > EXPORT COMPLEX > NOME_DO_ARQUIVO (EX.: COMPLEXOS_EXPORTADOS) > SALVAR**. Todos os arquivos exportados do GOLD, no formato *.mol2, podem ser abertos em qualquer computador, mesmo sem os arquivos de entrada elaborados previamente ao cálculo de ancoragem. Contudo, é importante salientar que outros arquivos obtidos no *.conf não poderão mais ser abertos pelo GOLD se forem copiados para outras pastas, ou se as pastas forem renomeadas.

Por isso, é importante definir pastas exclusivas para os estudos de ancoragem, com nomes precisos, sem caracteres especiais e que não sejam pastas que precisam ter os arquivos transferidos.

RESULTADOS E DISCUSSÃO

Análise dos resultados da ancoragem molecular, utilizando o software DSV

Tão importante quanto a elaboração da metodologia computacional, a etapa de interpretação dos dados exportados, em estudos de ancoragem molecular, consiste na avaliação da orientação e conformação da micromolécula no sítio de ligação.³² Além disso, consiste também em identificar quais são os resíduos chave e as forças intermoleculares que suportaram o reconhecimento complexo receptor-ligante.^{12,32}

Nesse sentido, após o carregamento do arquivo exportado do GOLD, no software DSV, analisam-se tanto as interações intermoleculares quanto os demais parâmetros mencionados, utilizando a ferramenta **LIGAND INTERACTIONS**, a qual está disponível na aba **RECEPTOR-LIGAND INTERACTIONS**, conforme a Figura 11.

Resalta-se que essa ferramenta permite inspecionar os complexos obtidos, por cálculos A.M. e aqueles determinados experimentalmente, desde que haja um arquivo de entrada, contendo as coordenadas cartesianas atômicas, no formato *.pdb.⁴

Na etapa de inspeção, após observar a conformação ativa do ligante, o usuário pode priorizar os contatos não-covalentes mais relevantes. Para tal, habilitando o painel de monitoramento do DSV, a aba **NON-BOND** permite ao usuário caracterizar as interações presentes no complexo, ao mesmo tempo, verificando quais são os parâmetros geométricos (distância e ângulo) associados. Na Tabela 1, podem ser visualizados alguns parâmetros geométricos, tipo de interação, cor atribuída e categoria das interações favoráveis, habitualmente listadas, em complexos receptor-ligante.

No exemplo em questão, o qual apresenta o menor valor para raiz do desvio-médio quadrático (RMSD \approx 1,5 Å), pode ser observado que o ligante simulado (**NCT.mol2**) exibe uma pose (orientação e conformação) estabilizada no bolsão de ligação do receptor nicotínico de acetilcolina $\alpha 4\beta 2$, entre as interfaces α - β , a partir da formação de contatos não-covalentes com resíduos chave, Figura 12.

hbond	PLP.part.metal	PLP.part.nonpols	PLP.part.repulsive	Reference.RMSD	Version
-1.0000	0.0000	-48.0295	0.0280	1.6075	0.0000
-1.0000	0.0000	-49.5015	0.0217	1.6027	0.0000
-1.0000	0.0000	-48.2512	0.0298	1.6046	0.0000
-1.0000	0.0000	-51.1858	0.0271	1.6069	0.0000
-1.0000	0.0000	-50.9417	0.0288	1.5860	0.0000
-1.0000	0.0000	-50.5942	0.0273	1.6024	0.0000
-1.0000	0.0000	-48.4728	0.0248	1.6030	0.0000
-1.0000	0.0000	-49.6116	0.0234	1.6075	0.0000
-1.0000	0.0000	-52.1343	0.0271	1.5973	0.0000
-1.0000	0.0000	-48.6827	0.0242	1.6137	0.0000

Figura 10. Janela de resultados

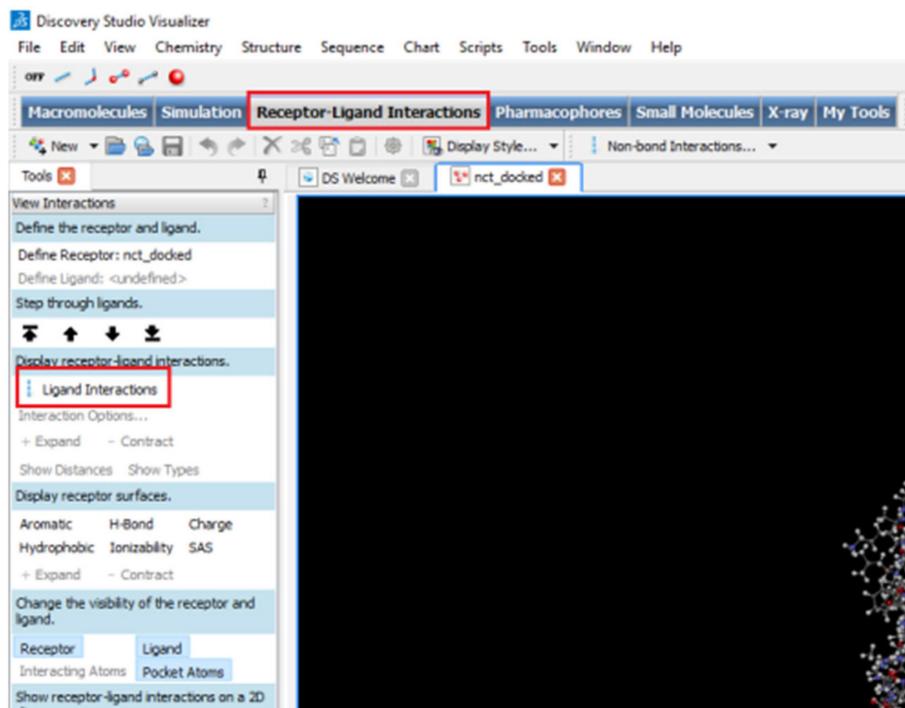


Figura 11. Localização da ferramenta Ligand interactions

Tabela 1. Algumas interações intermoleculares usualmente analisadas em complexos receptor-ligante. Os parâmetros geométricos e cores associadas estão de acordo com a configuração padrão do software DSV.^{27,33}

Tipo de interação	Cor atribuída	Categoria	Parâmetros Geométricos		Descrição
Interações de hidrogênio	Verde Azul claro Laranja	Clássica; Não clássica; Ponte salina;	D-H...A (máx. 3,4 Å); D-H...A (máx. 3,8 Å); D-H...A (máx. 4,0 Å); A...A (máx. 3,0 Å);	$D\hat{H}A$ (mín. 90°; máx. 180°)	D-H...A (A = aceitador); C-H...A/par isolado/ π ; Interações de H entre grupos carregados;
Eletrostática	Laranja Laranja Laranja	Carga...carga; π ...cátion; π ...ânion;	(+)...(-) (máx. 5,6 Å); π ...(+) (máx. 5,0 Å); π ...(-) (máx. 5,0 Å);	- (Máx. 40°); (Máx. 40°);	Interação entre cargas opostas; Interação de nuvem π e grupo (ou átomo) carregado positivamente/negativamente;
Hidrofóbica	Roxa Rosa Rosa	π ... π ; Alquila; π ...alquila; π ...par isolado (PI)	Csp2...Csp2 (máx. 4,5 Å) Csp3...Csp3 (máx. 5,5 Å) Csp2...Csp3 (máx. 4,0 Å) Csp2...PI (máx. 3,0 Å)	(Máx. 90°; mín. 50°, face-à-face); (Máx. 90°, mín. 30°, forma T); - (Máx. 45° ± 20°); (Máx. 45° ± 40°);	Interação entre nuvens eletrônicas π deslocalizadas; Fraca interação σ - σ ; Fraca interação π - σ ; Fraca interação π -PI (Par Isolado);
Interações de Halogênio	Laranja claro Laranja claro	F; Cl, Br e I;	C-F...A (máx. 3,7 Å); Cl, Br e I (X...A máx. $\leq \Sigma \text{vdW}(X, A) + 1$ Å);	$C\hat{F}A$ (min. 120°; máx. 180°); $C\hat{X}A$ (min. 120°; máx. 180°);	Interações envolvendo cavidades- σ

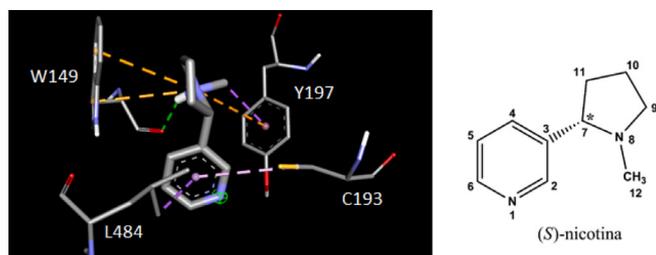


Figura 12. Tela de trabalho do DSV com o complexo nicotina- $\alpha 4\beta 2$ e esquema 2D da molécula de nicotina. As interações intermoleculares estão exibidas como linhas tracejadas, diferenciadas por cores em tipo e natureza (Verde = interação de hidrogênio clássica; Laranja = interação eletrostática; Roxa = Interações hidrofóbicas; Rosa = Interações parcialmente hidrofóbicas). Obs.: Para uma melhor visualização, apenas os hidrogênios polares são mostrados, CHEMISTRY > HYDROGENS > SHOW POLAR

Na tela de trabalho do DSV, as linhas tracejadas representam os contatos não-covalentes, sendo cada cor referente a uma categoria (natureza) de interação (ex. a linha tracejada verde indica a formação de uma interação de hidrogênio clássica do tipo D-H...A, sendo D = doador e A = aceitador). É importante salientar que o software DSV, em configuração padrão, utiliza essa paleta de cores tão-somente para diferenciar as interações. Todavia, o usuário poderá definir a cor da interação, clicando na coluna referente a cor na aba **NON-BOND** e selecionar aquela que melhor represente o sistema de interesse.

Como segue, o complexo nicotina- $\alpha 4\beta 2$ indica que o nitrogênio pirrolidínico (N8), o qual se encontra protonado, estabelece três interações de natureza eletrostática π ... π com os anéis aromáticos das cadeias laterais dos resíduos de W149 e Y197, Figura 12. Além disso, devido a protonação do átomo de N8, ocorre a formação de uma interação de hidrogênio clássica (N-H...O) com o átomo de O carbonílico da cadeia principal do resíduo de W149. Duas interações

hidrofóbicas do tipo $\sigma\cdots\pi$ são observadas entre a metila, C12, e o anel aromático da nicotina com as cadeias laterais dos resíduos de Y197 e L484, respectivamente. Finalmente, um contato do tipo $\pi\cdots\pi$ (par isolado) entre a subunidade aromática da nicotina e o átomo de enxofre presente na cadeia lateral do resíduo de C193 auxilia a estabiliza da nicotina no bolsão de ligação.

O algoritmo genético do *software* GOLD modula as orientações e conformações do ligante, visando encaixá-lo em um ou mais mínimos de energia do sítio ativo do receptor.¹⁵ Prontamente, as populações de cromossomos são criadas e avaliadas pelos operadores genéticos (mutação, recombinação ou *crossover*), mapeando-se os parâmetros de geometria do ligante. Nessa conjuntura, o A.G. combina a sobrevivência entre as melhores soluções com intercâmbio de informações genéticas para formar estruturas heurísticas aperfeiçoadas, ou seja, os ligantes com melhores adaptabilidade (orientação e conformação) no sítio de ligação. Os valores das pontuações, compostas pela modelagem das interações receptor-ligante e afinidade, são conduzidas pela função de pontuação pré-definida pelo usuário.^{2,13-15,19}

No exemplo em questão, foi aplicada a metodologia de redocagem para validação da rotina de cálculo empregada na simulação. É advertido que este protocolo tem maior eficiência em situações nas quais as moléculas otimizadas e o ligante co-cristalizado são estruturalmente similares.^{12,32} Além disso, nessa metodologia é importante manter a estrutura da proteína fixa (rígida), a fim de garantir a conformação obtida experimentalmente, dos resíduos do sítio ativo. Nessa etapa, examina-se a capacidade do A.G. conduzir o ligante otimizado à conformação determinada pelo ligante co-cristalizado (reprodução do dado experimental). Na prática, o usuário deve, ao final do cálculo, observar os valores de RMSD com base na sobreposição das posições cartesianas dos átomos não-hidrogenóides, referentes as moléculas (simulada e original), no sítio de ligação. São plausíveis os protocolos com valores da raiz do desvio-médio quadrático não superiores a 2,0 Å.^{2,12,13,19,34}

Edição de figuras, utilizando o *software* PyMol

Para realizar a inspeção visual, ambos os arquivos (complexo exportado e a estrutura original 5kxi.pdb) são carregados no PyMol.³⁵ É importante destacar que a estrutura da biomacromolécula alvo pode ser obtida por homologia (modelagem molecular comparativa), no caso da ausência do arquivo 3D nos bancos de dados de proteínas (ver material suplementar, seção 4S).

Com o *software* PyMol aberto, as estruturas devem ser suprimidas:

ALL > H > EVERYTHING. Os ligantes de interesse são selecionados na barra de sequência, **DISPLAY > SEQUENCE**, sendo o ligante co-cristalizado denominado como **NCT** (localizado entre as cadeias A e B do receptor) e no complexo exportado o ligante é identificado como **UNK**, e está localizado no início da barra de sequência. Para facilitar a inspeção, na opção (**SELE**) a função S, “show”, é habilitada com posterior clique em **STICKS**. Por conseguinte, faz-se necessário remover os hidrogênios através da função **A “Action” > HYDROGENS > REMOVE**. Ainda, o usuário tem a opção de alterar a cor do plano de fundo para branco, **DISPLAY > BACKGROUND > WHITE**, Figura 13.

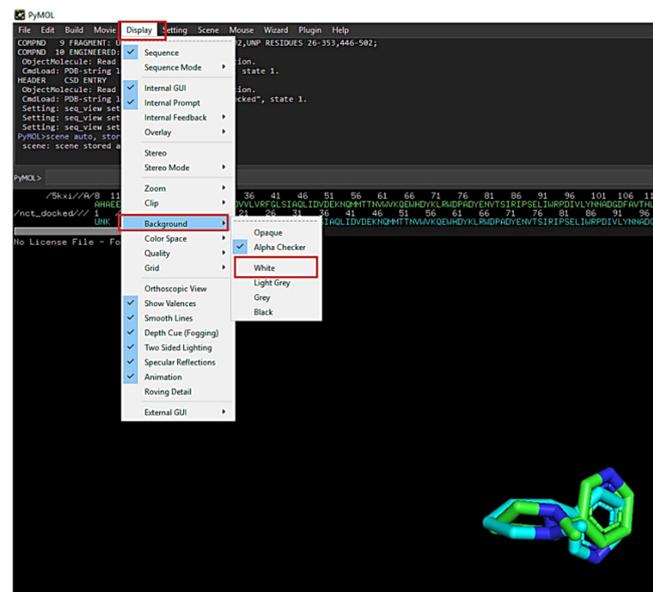


Figura 13. Moléculas de nicotina sobrepostas: simulada (C, em azul claro) e co-cristalizada (C, em verde), apresentadas como “sticks” e com os respectivos átomos de H suprimidos. Além disso, etapa de alteração do plano de fundo

Para expandir a seleção e apresentar as cadeias laterais dos resíduos que compõem o sítio de ligação, basta habilitar a proteína como “cartoon” (**ALL > S “SHOW” > CARTOON**) e clicar com o botão direito do mouse sobre o ligante co-cristalizado, seguindo **ACTIONS > EXPAND > BY 5 Å, RESIDUES**, conforme demonstrado na Figura 14.

Mantendo os resíduos selecionados, deve ser habilitada a função S, à direita da opção (**SELE**), optando pela forma **STICKS**.

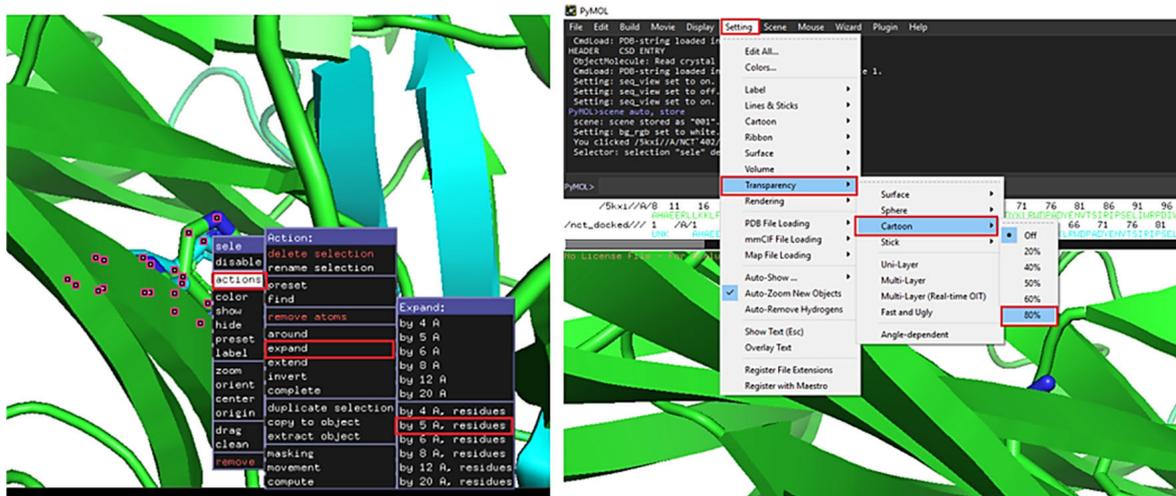


Figura 14. (Esquerda) Seleção e formatação dos resíduos, em um raio de 5 Å do ligante de interesse. (Direita) Aumento de transparência da estrutura em “cartoon”

Uma melhor visualização é adquirida alterando a cor do ligante co-cristalizado em relação aos resíduos previamente expandidos. Nesse sentido, mantenha-o selecionado aplicando em sequência o comando **C > BY ELEMENT**, escolhendo uma das cores do painel. Aqui, foi escolhida a coloração laranja para os átomos de carbono. Ao mesmo tempo, é sugerido o comando **SETTING > TRANSPARENCY > CARTOON > 80%**, para melhorar a visualização, conforme mostrado na Figura 14.

Antes de completar a edição da figura, podem ser coletadas as coordenadas da pose, segundo a visão atual na tela, com a palavra-chave **GET_VIEW**, o qual deve ser digitada pelo usuário na linha de comando em frente ao nome “PyMol>”, pressionando a tecla **Enter** ao final do processo, Figura 15.

Essas coordenadas podem ser utilizadas, posteriormente, para analisar outros ligantes ancorados, com o mesmo receptor, a partir da mesma orientação. Para realizar esse procedimento, é suficiente colar,

Ctrl+V, as coordenadas já coletadas na linha de *input* dos comandos e pressionar a tecla **Enter**.

As rotinas de A.M. devem ser validadas, sendo a redocagem o método mais utilizado. Nessa abordagem, a estrutura do ligante co-cristalizado da proteína é removido do sítio de ligação. Em sequência, o algoritmo deverá recalculer a orientação e conformação do ligante no bolsão de ligação. Serão considerados protocolos aptos, ou bem sucedidos, caso haja concordância entre o modo de ligação experimental e o calculado. Aqui, a pose de maior pontuação apresentou valor de RMSD = 1,5 Å, isso é, RMSD < 2Å, considerada uma rotina apta para o sistema em questão.

Finalmente, para gerar uma figura de alta qualidade, visando relatórios técnicos, artigos, dentre outros, seleciona-se a opção **DRAW/RAY > RAY(slow)**, conforme exemplificado na Figura 16.

Quando o *software* terminar o cálculo, pressione **FILE>SAVE IMAGE TO FILE**, para salvar a imagem no diretório desejado. Para

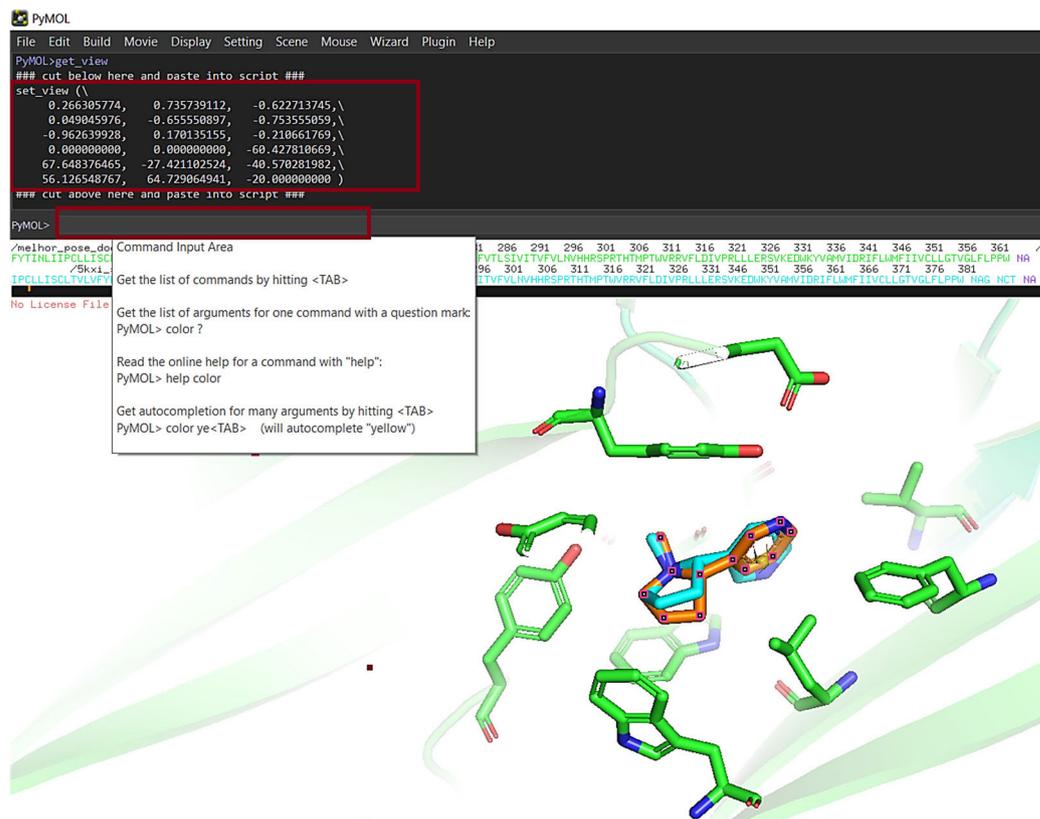


Figura 15. Coordenadas cartesianas da pose de A.M. segundo a visão atual da tela; Linha de comandos do PyMol. Moléculas do ligante nicotina sobrepostas: simulada (C, em azul claro), co-cristalizada (C, em laranja) e resíduos do sítio de ligação (C, em verde) com no máximo 5 Å de distância do ligante. Todas as moléculas estão apresentadas como “sticks” e com os respectivos átomos de H omitidos para melhor visualização

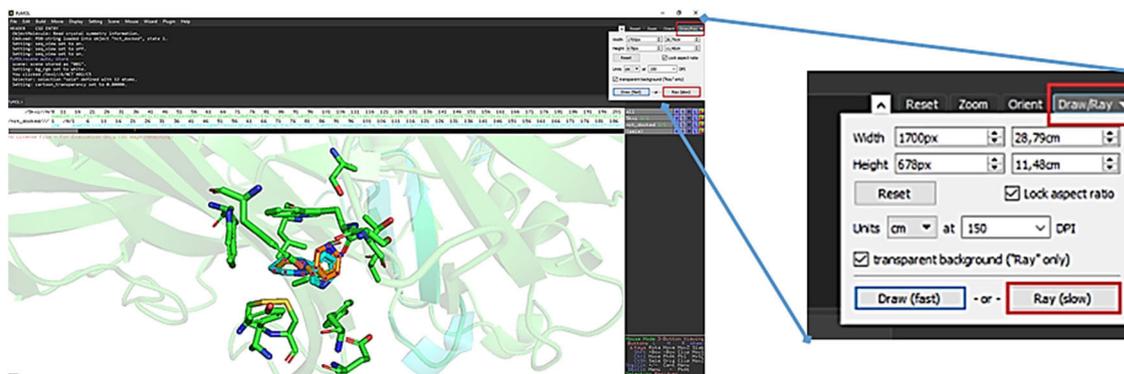


Figura 16. Gerar figuras de alta qualidade com a função Ray

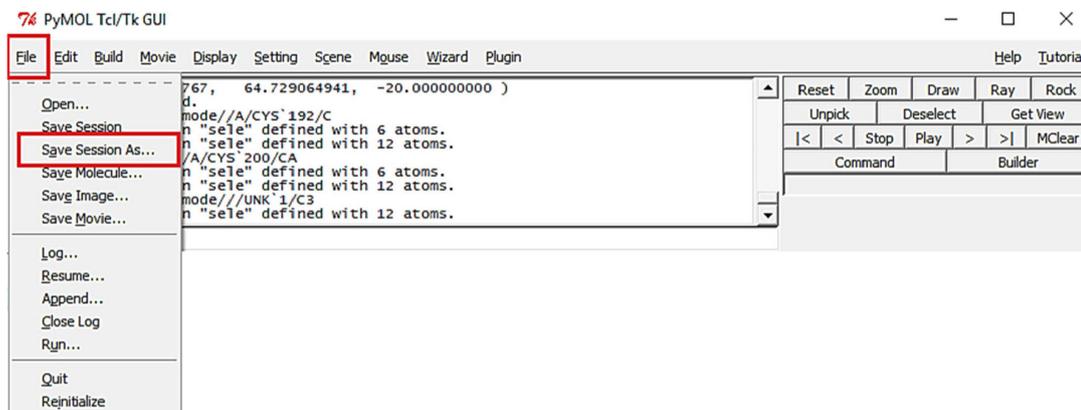


Figura 17. Salvar arquivo de sessão no PyMol

salvar um arquivo de sessão contendo as modificações feitas até o momento, siga os comandos **FILE > SAVE SESSION AS** e escolha o local desejado, Figura 17.

Obs.: A opção de nomear os resíduos de aminoácidos nesse *software* pode resultar em figuras poluídas visualmente. Portanto, é aconselhável identificar os resíduos, posteriormente, com um editor de imagem (ex. *Paint* ou *PowerPoint*).

CONCLUSÕES

No decorrer deste texto, uma rotina computacional de fácil acesso e de baixo custo foi descrita, objetivando apresentar as principais etapas que pertencem a um procedimento de A.M., visando um público de usuários recentes e não especialistas. Nesse sentido, foram abordadas as etapas obtenção do arquivo 3D da biomacromolécula de interesse, construção e otimização do ligante, especificação do estado de protonação em pH fisiológico, condução da simulação de A.M. por meio do *software* GOLD, visualização do complexo receptor-ligante e monitoramento das interações intermoleculares via *software* DSV, concluindo com a edição de figuras de qualidade para publicação, utilizando o *software* PyMol.

Por fim, é fortemente reiterado que essa metodologia deve ser modificada, em algumas etapas, para resolução de problemas individuais de A.M. No entanto, o mesmo pode servir como um guia inicial aos interessados em SBDD.

MATERIAL SUPLEMENTAR

O Material Suplementar citado ao longo do trabalho está disponível em <http://quimicanova.sbq.org.br>, em formato pdf, com acesso livre.

AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. Adicionalmente, agradecemos à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Projeto Regular, Processo 2018/00187-7, pelo apoio financeiro ao LaQMedSOMM, que foi fundamental para o desenvolvimento deste trabalho.

REFERÊNCIAS

1. Anwar, T.; Kumar, P.; Khan, A. U. In *Molecular Docking for Computer-Aided Drug Design*, Coumar, M. S., ed.; Academic Press: Cambridge, 2021; cap. 1.

2. Amutha R.; Sangeetha B.; Muthukumar, R. In *A Theoretical Revisit on Molecular Modeling and Docking Approaches*, Coumar, M. S., ed.; Academic Press: Cambridge, 2021; cap. 1.
3. Santos, L. H. S.; Ferreira, R. S.; Caffarena, E. R. In *Integrating Molecular Docking and Molecular Dynamics Simulations*. Azevedo Jr, W. F., ed.; Humana Press: Totowa, 2019; cap 2.
4. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E.; The Protein Data Bank; *Nucleic Acids Research* **2000**, *28*, 235.
5. Dai, W.; Zhang, B.; Su, H.; Li, J.; Zhao, Y.; Xie, X.; Jin, Z.; Liu, F.; Li, C.; Li, Y.; Bai, F.; Wang, H.; Cheng, X.; Cen, X.; Hu, S.; Yang, X.; Wang, J.; Liu, X.; Xiao, G.; Jiang, H.; Rao, Z.; Zhang, L.-K.; Xu, Y.; Yang, H.; Liu, H.; *Science* **2020**, eabb4489.
6. Sahoo, M.; Jena, L.; Daf, S.; Kumar, S.; *Genomics Inform.* **2016**, *14*, 104.
7. Moorthy, N. S. H. N.; Sousa, S. F.; Ramos, M. J.; Fernandes, P. A.; *J. Enzyme Inhib. Med. Chem.* **2016**, *31*, 1428.
8. Dong, G.; Wu, Y.; Sun, Y.; Liu, N.; Wu, S.; Zhang, W. N.; Sheng, C.; *MedChemComm* **2018**, *9*, 1142.
9. Banerjee, R.; Perera, L. Viranga, T. L. M.; *Drug Discov. Today* **2021**, *26*, 804.
10. Mitsui, T.; Hirayama, K.; Aoki, S.; Nishikawa, K.; Uchida, K.; Matsumoto, T.; Kabuta, T.; Wada, K.; *Neurochem. Int.* **2010**, *56*, 679.
11. Bhutoria, S.; Ghoshal, N.; *J. Mol. Graphics Modell.* **2010**, *28*, 577.
12. Barreiro, E. J.; Fraga, C. A. M.; *Química medicinal: as bases moleculares da ação dos fármacos*, 3ª ed., Artmed: Porto Alegre, 2015.
13. Silakari, O.; Singh, P. K. In *Concepts and Experimental Protocols of Modelling and Informatics in Drug Design*, Silakari, O., Singh, P. K., eds.; Academic Press: Cambridge, 2021; cap 6.
14. Pagadala, N. S.; Syed, K.; Tuszyński, J.; *Biophys. Rev.* **2017**, *9*, 91.
15. Li, Z.; Gu, J.; Zhuang, H.; Kang, L.; Zhao, X.; Guo, Q.; *Appl. Soft Comput.* **2015**, *26*, 299.
16. Schneider, G.; Böhm, H.-J.; *Drug Discov. Today* **2002**, *7*, 64.
17. Chen, Y.-C.; *Trends Pharmacol. Sci.* **2015**, *36*, 78.
18. Bagheri, M.; Fatemi, M. H.; *J. Lumin.* **2018**, *202*, 345.
19. Morris, G. M.; Lim-Wilby, M.; *Methods Mol. Biol.* **2008**, *443*, 365.
20. Kang, L.; Guo, Q.; Wang, X.; *Bioorg. Med. Chem. Lett.* **2012**, *22*, 6568.
21. Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J.; *J Comput. Chem.* **2009**, *30*, 2785.
22. Morales-Perez, C. L.; Noviello, C. M.; Hibbs, R. E.; *Nature* **2016**, *538*, 411.
23. Macrae, C. F.; Sovago, I.; Cottrell, S. J.; Galek, P. T. A.; McCabe, P.; Pidcock, E.; Platings, M.; Shields, G. P.; Stevens, J. S.; Towler, M.; Wood, P. A.; *J. Appl. Crystallogr.* **2020**, *53*, 226.
24. Beer, T. A. P.; Berka, K.; Thornton, J. M.; Laskowski, R. A.; *Nucleic Acids Res.* **2014**, *42*, D292.

25. Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M.; *J. Appl. Cryst.* **1993**, *26*, 283.
26. Lovell, S. C.; Davis, I. W.; Arendall, W. B.; de Bakker, P. I.; Word, J. M.; Prisant, M. G.; Richardson, J. S.; Richardson, D. C.; *Proteins* **2003**, *50*, 437.
27. BIOVIA Dassault Systèmes; *BIOVIA Discovery Studio Visualizer*; San Diego, 2020.
28. ChemAxon; *MarvinSketch* v21.4, disponível em: <http://www.chemaxon.com>, acessada em setembro 2021.
29. Stewart, J. J. P.; *MOPAC2016; Stewart Computational Chemistry*: CO, USA, 2016.
30. Stewart, J. J.; *J. Mol. Model.* **2013**, *19*, 1.
31. Li, Y.; Han, L.; Liu, Z.; Wang, R.; *J. Chem. Inf. Model.* **2014**, *54*, 1717.
32. Batista, V. S.; Nascimento-Júnior, N. M.; *Med. Anal. Chem. Int. J.* **2018**, *2*, 000123.
33. Kemmish, H.; Fasnacht, M.; Yan, L.; *PLoS One* **2017**, *12*, e0177923.
34. Rocha, F. V.; Farias, R. L.; Lima, M. A.; Batista, V. S.; Nascimento-Júnior, N. M.; Garrido, S. S.; Leopoldino, A. M.; Goto, R. N.; Oliveira, A. B.; Beck, J.; Landvogt, C.; Mauro, A. E.; Netto, A. V. G.; *J. Inorg. Biochem.* **2019**, *199*, 110725.
35. Alexander, N.; Woetzel, N.; Meiler, J.; *International Conference on Computational Advances in Bio and Medical Sciences* **2011**, 2011, 13.