



UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO DE CIÊNCIAS DA SAÚDE
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA
BIOINFORMÁTICA



ATRACAMENTO MOLECULAR

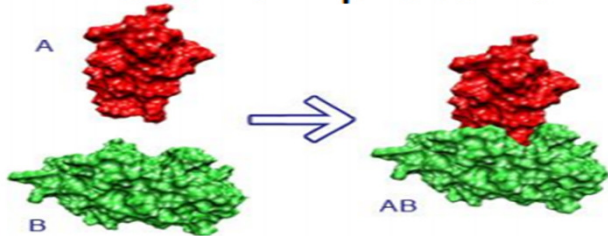
Ivana Alece Arantes Moreno
Júlia de Assis Pinheiro
Pablo Viana Oliveira
Tamires dos Santos Vieira

VITÓRIA
2018

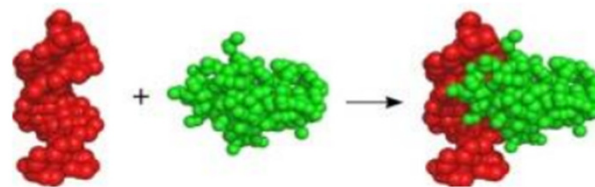
Atracamento Molecular

- Interação entre macromoléculas

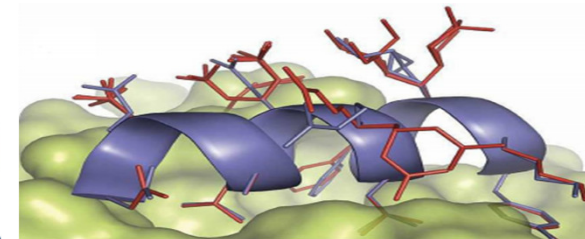
Proteína – proteína



DNA – proteína



Proteína - peptídeo



- Macromolécula e pequena molécula ligante → agonista/antagonista ou substrato/inibidor

Proteína – pequena molécula

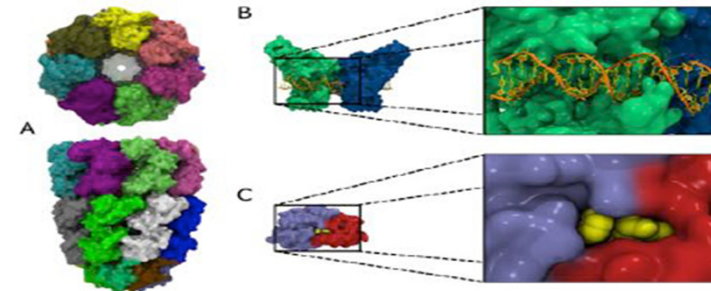
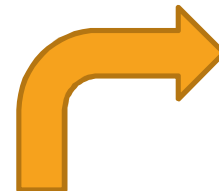
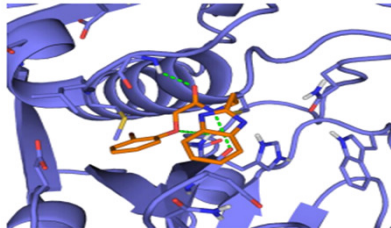


Figura 1-9: Exemplos de complexos moleculares: (A) chaperona GroEL (PDB ID 1AON), (B) complexo DNA com proteína DMT1 (PDB ID 3PT6) e (C) complexo da enzima HIV-1 protease com o inibidor indinavir (PDB ID 1H5G). As versões menores em B e C estão em escala com A.

- Estimar como duas moléculas interagem

A afinidade dessa ligação depende de uma função de avaliação e dos métodos computacionais aplicados [Guedes et al. 2013].

- **Atracamento molecular (*Molecular Docking*)**
 - Técnica computacional dedicada à previsão do **modo de ligação** e dos detalhes do **reconhecimento molecular** proteína-proteína e receptor-ligante.

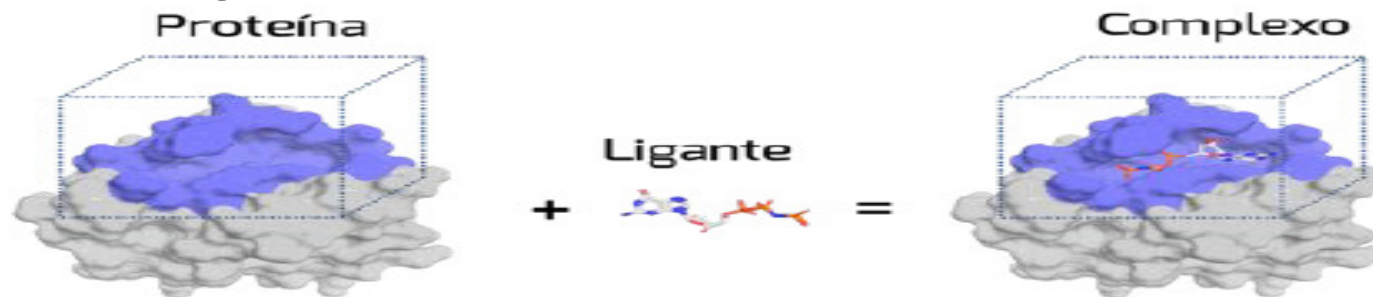
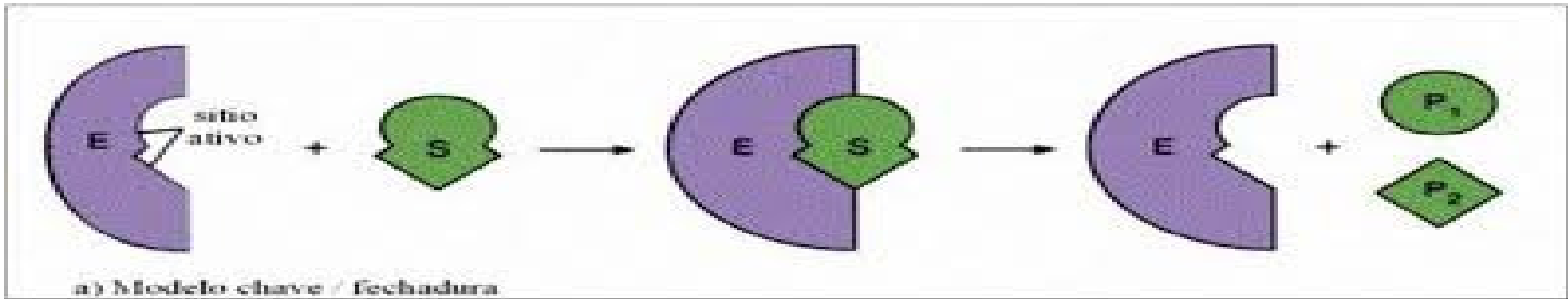


Figura 2-9: Emprego do método de atracamento molecular na predição do modo de ligação do GTP ao seu sítio de ligação na proteína c-H-ras p21.

Pesquisas associadas à saúde e à biotecnologia.

Reconhecimento Molecular

- Metodologias computacionais baseadas no modelo chave-fechadura.



- Tanto o ligante quanto a proteína são flexíveis, podendo modificar a sua conformação durante o processo de formação do complexo receptor-ligante (ajuste induzido).



Reconhecimento Molecular

Flexibilidade da Proteína

catálise de reações enzimáticas

transdução de sinais

transporte através de proteínas de membrana

mudanças conformacionais associadas a formas ativas e não ativas de proteínas

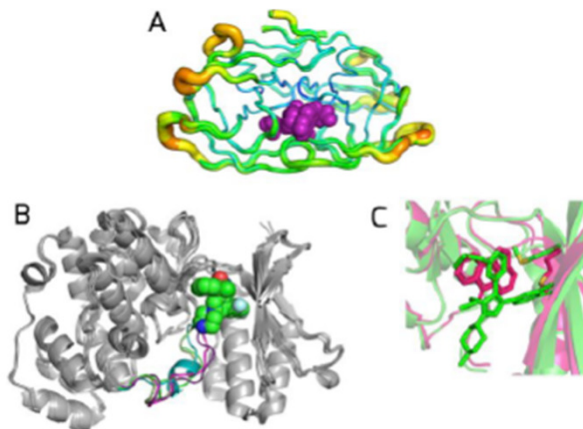
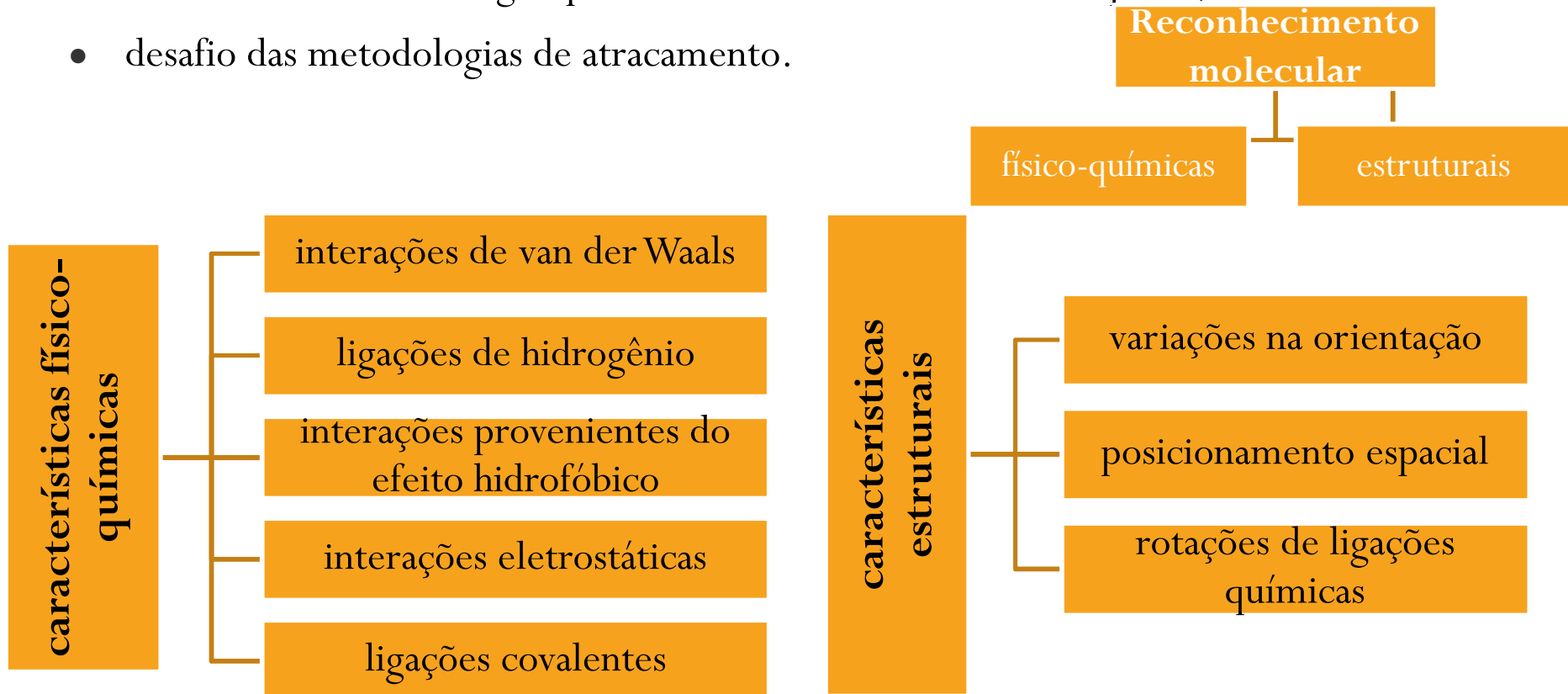


Figura 3-9: Graus de flexibilidade do receptor: (A) mobilidade do esqueleto peptídico da enzima protease do HIV-1, (B) diversas conformações de alça no sítio de ligação do ATP à enzima MAP cinase p38, e (C) mudança conformacional da cadeia lateral de resíduo na enzima cinase JNK3, influenciada por diferentes inibidores.

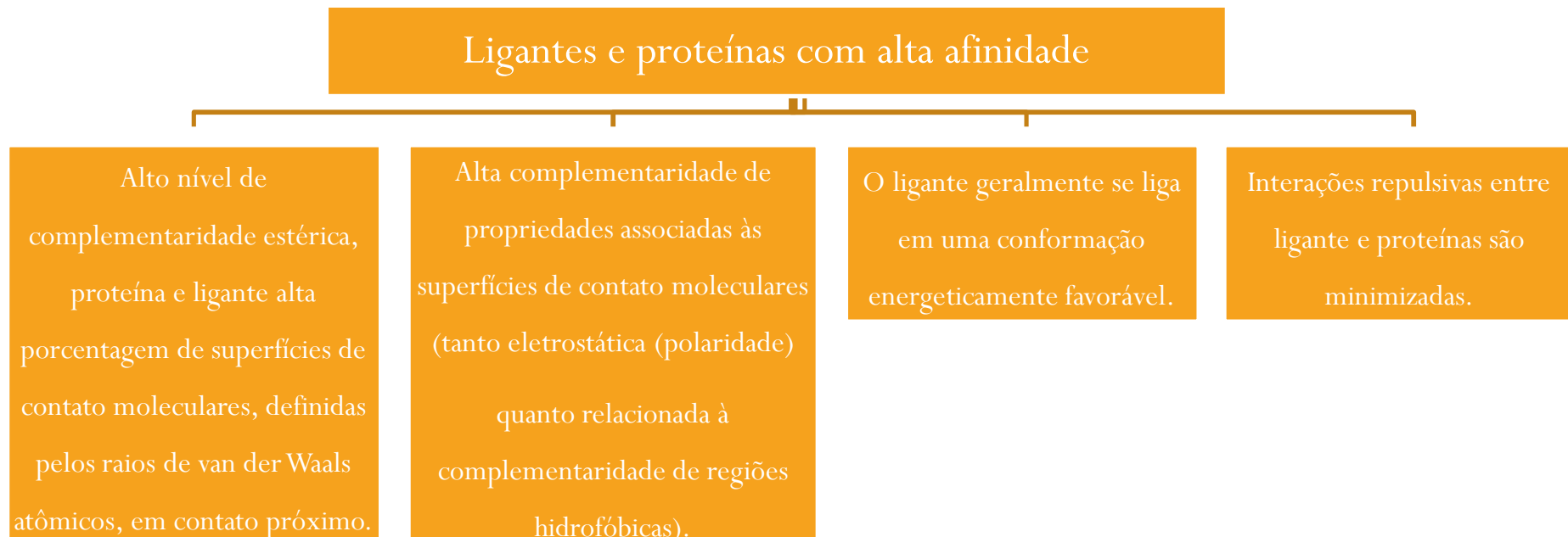
Reconhecimento Molecular

- **Atracamento proteína-ligante:** modelos de encaixe induzido e seleção conformacional;
 - direcionar as metodologias para fornecer um tratamento adequado;
 - desafio das metodologias de atracamento.



Reconhecimento Molecular

- **Características físico-químicas:** definem o grau de afinidade e de especificidade do ligante pela proteína relacionadas com as interações intermoleculares existentes no complexo;
- **Características estruturais:** associadas aos arranjos espaciais moleculares.



Problema de atracamento

- i)* investigação e predição da conformação e orientação de uma molécula ligante no seu sítio de complexação;

- ii)* predição da afinidade em um complexo receptor-ligante.

Scoring

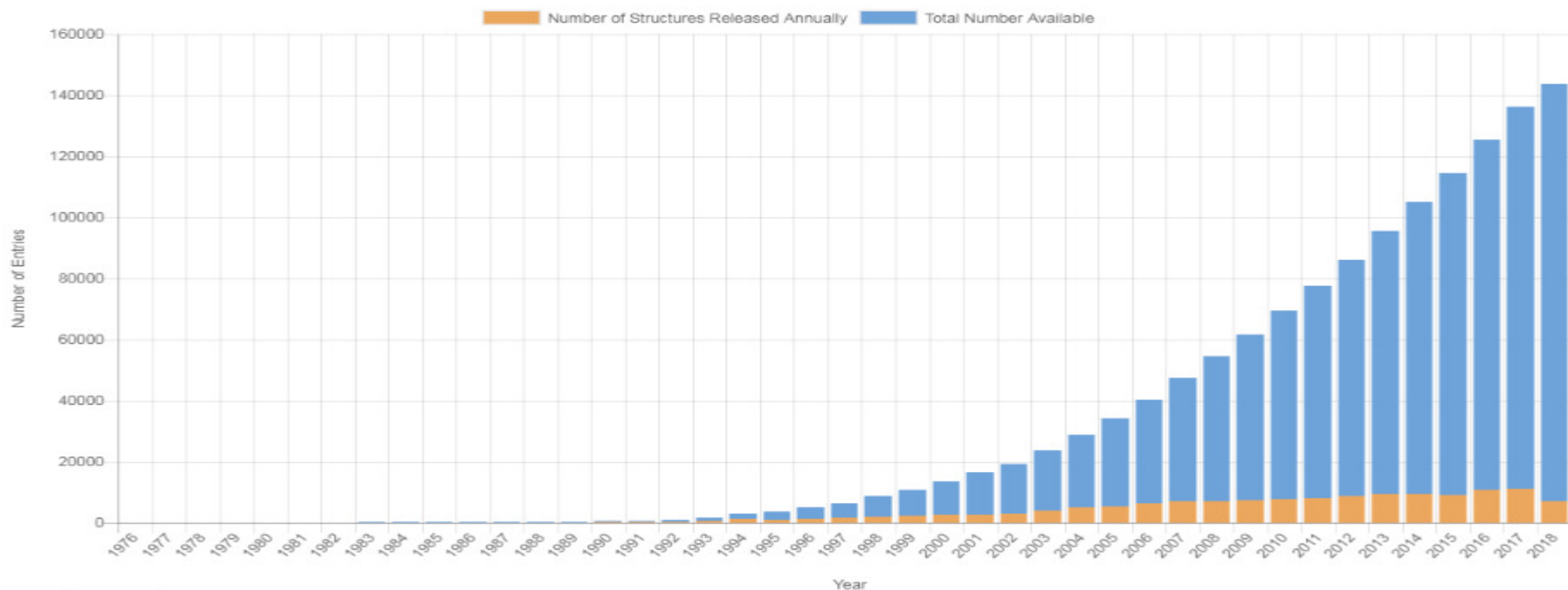
Métodos de Atracamento

Programas

Portal	Programa de atracamento
SwissDock	EADock DSS
DockingServer	AutoDock
DockThor Portal	DockThor
1-Click Docking	AutoDock Vina
DOCK Blaster	DOCK
Docking At UTMB	AutoDock Vina
ParDOCK	Método de Monte Carlo
PATCHDOCK	PatchDock
MEDock	MEDock

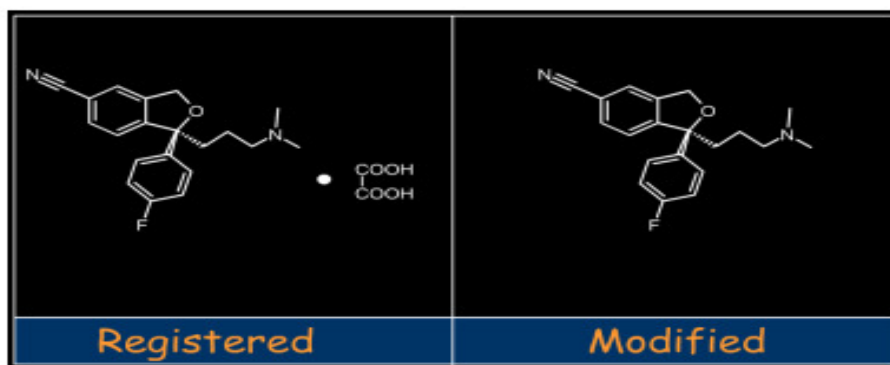
Preparação do sistema

PDB Statistics: Overall Growth of Released Structures Per Year



Preparação do sistema

Salt removing



Optical isomers (enantiomers) generation



Tautomers generation



Double bond cis/trans isomers generation



Preparação do sistema



- CORINA
- CONCORD
- OMEGA
- Balloon
- Multiconf-DOCK.

Preparação do sistema

- **Prever os sítios de ligação:**

- análises GEOGRÁFICAS e VOLUMÉTRICAS;
 - **Cavidades** (FINDSITE, SURFNET e LIGSITE);
 - **Energias de interação** (Q-SITEFINDER e GRID);
 - Uso de propriedades p/ busca de **padrões** (webPDBinder);

- **Procura de erros:**

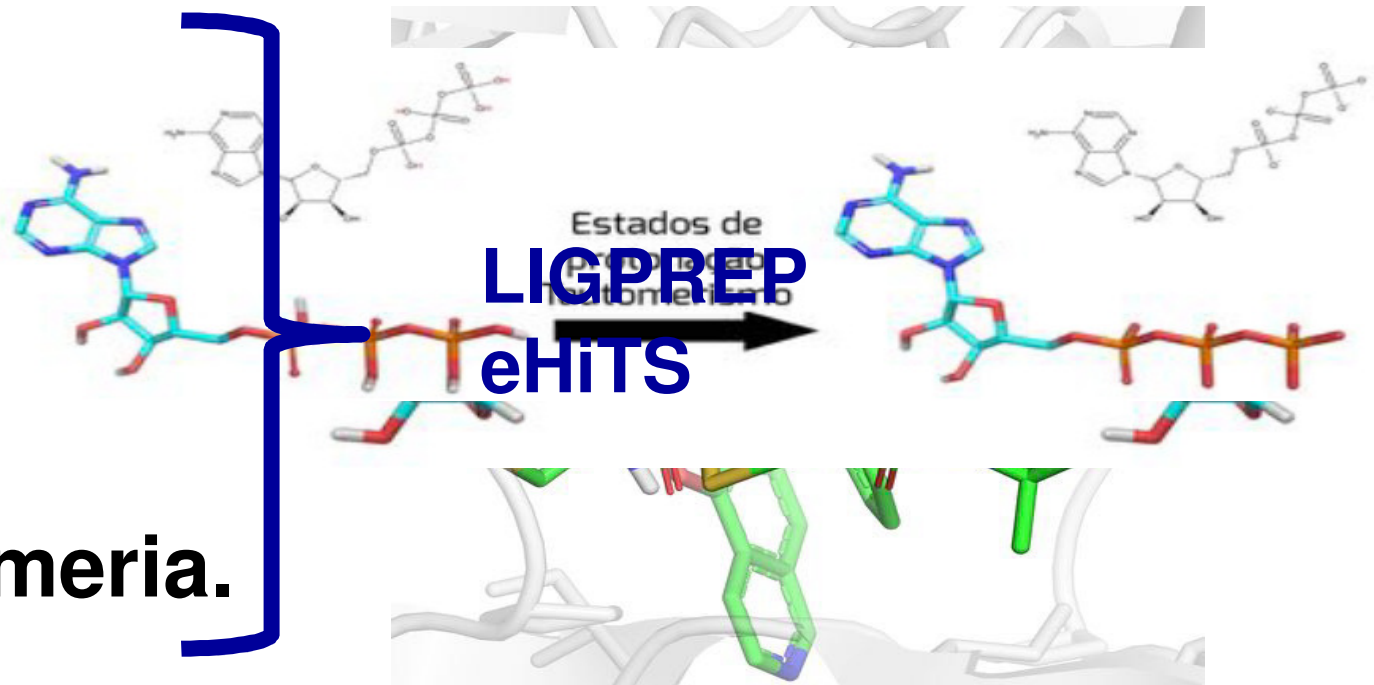
- WHAT_IF, MOLPROBITY e PROCHECK.

Preparação do sistema

- Possíveis erros:
 - ausência de átomos e/ou resíduos;
 - mau posicionamento de cadeias laterais;
 - presença de duas ou mais conformações para um resíduo ou conjunto;
 - conformações não nativas.

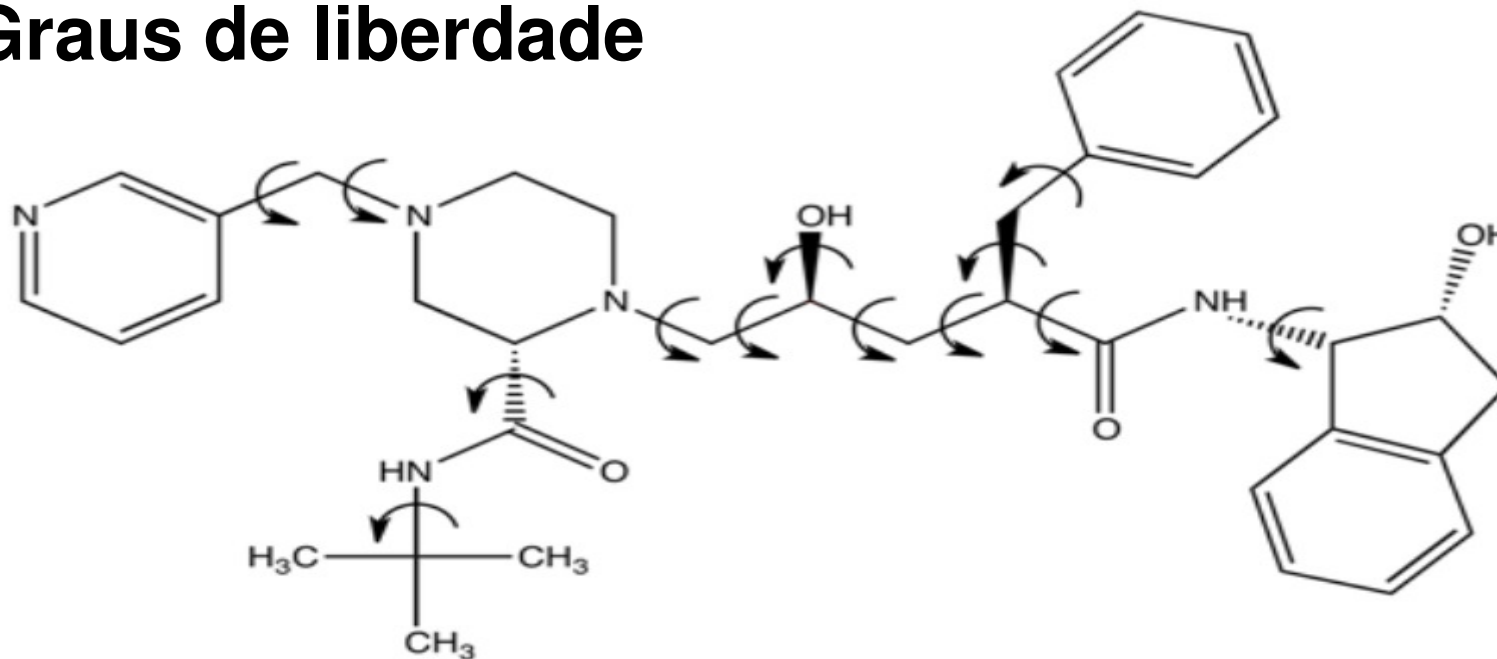
Preparação do sistema

- Estado de protonação
- Isomeria
- Tautomeria.



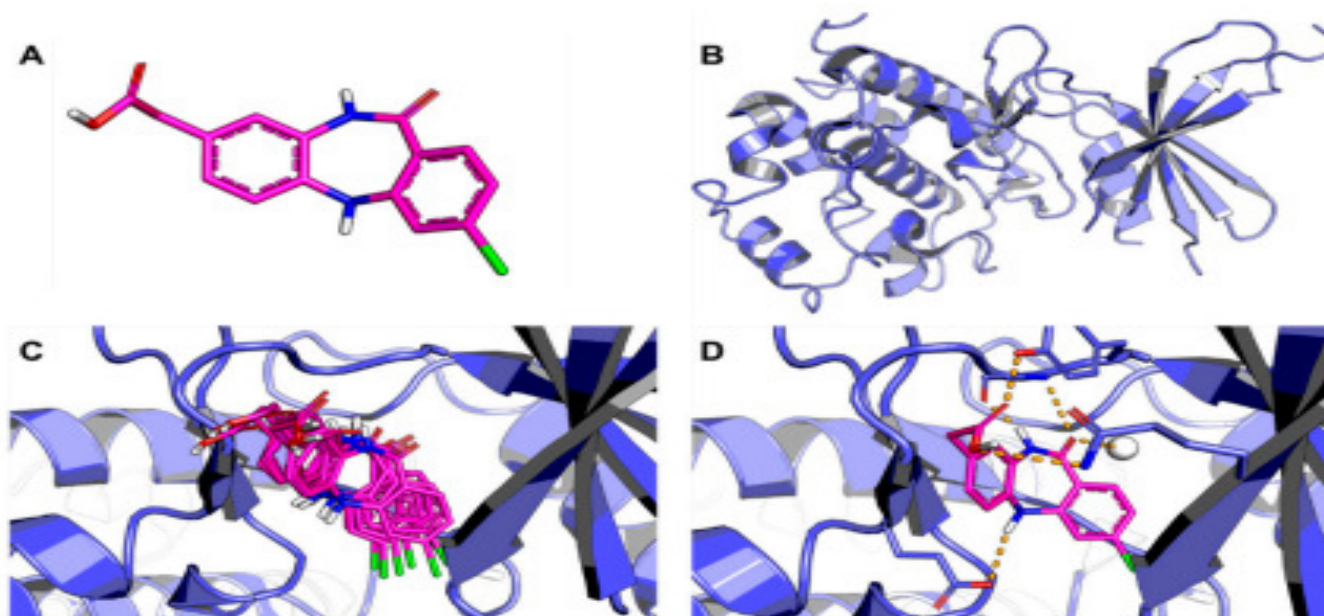
Preparação do sistema

Graus de liberdade



Preparação do sistema

Esboço do processo de acoplamento molecular



FONTE: Ferreira et al. Molecular Docking and Structure-Based Drug Design Strategies. *Molecules*. 20, 13384-13421, 2015. doi:10.3390/molecules200713384

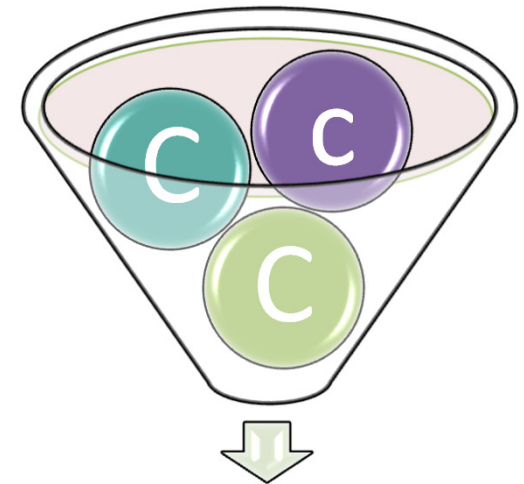
Métodos de busca

Método de busca	Função
Sistemática	Um conjunto de valores é estabelecido para cada grau de liberdade. O objetivo é explorar de forma combinatória todos os graus de liberdade da molécula durante a busca.
Determinística	Dado um mesmo estado inicial de entrada, é obtido sempre o mesmo resultado de saída. Métodos de simulação por dinâmica molecular e métodos clássicos de minimização de energia são exemplos.
Estocástica	O processo de otimização envolve movimentos aleatórios associados aos graus de liberdade. Este fato implica na possibilidade de se obter diferentes resultados como saída para um mesmo estado inicial de entrada.

FUNÇÕES DE AVALIAÇÃO

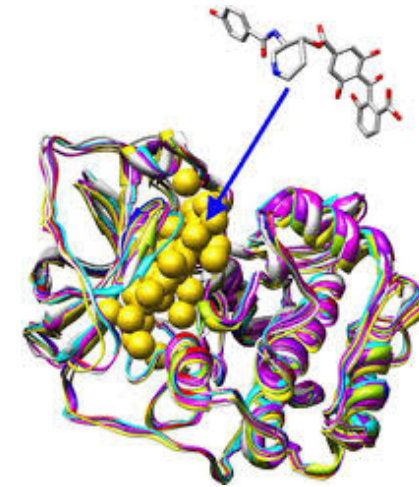
- Funções de busca → Várias conformações
- Avaliar a qualidade e verificar a afinidade com o receptor.
- Precisa ser capaz de ordenar ligantes:
 - Previsão do modo de ligação
 - Triagem virtual
 - Predição de afinidade

FUNÇÕES DE
AVALIAÇÃO



MELHOR
CONFORMAÇÃO

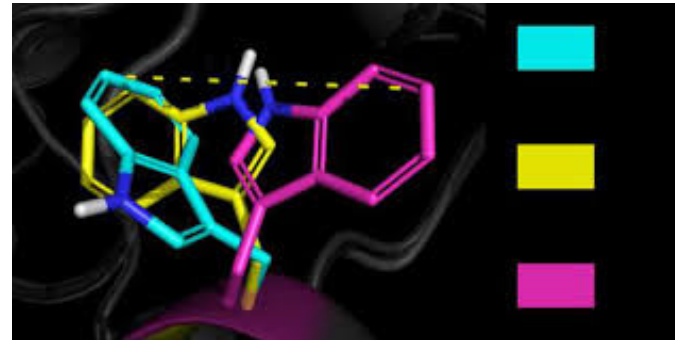
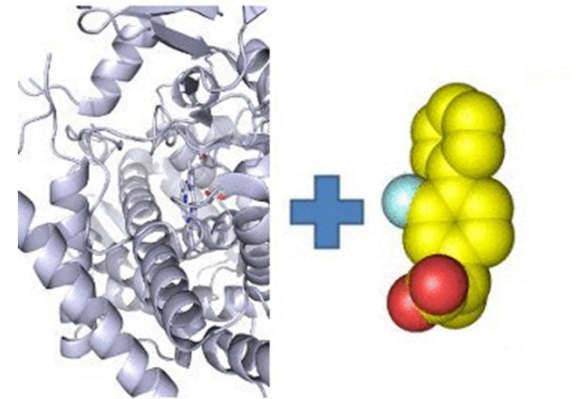
- Estas funções são modelos matemáticos
 - geralmente lineares,
 - propriedades físico-químicas
 - Relação de molécula ligante + receptor.
- Variam:
 - Número e tipo de termos
 - Complexidade matemática
 - Forma de parametrização
- Tipo de função:
 - Simples
 - Complexa



Custo
computacional

FUNÇÕES DE AVALIAÇÃO DE ATRACAMENTO MOLECULAR RECEPTOR- LIGANTE

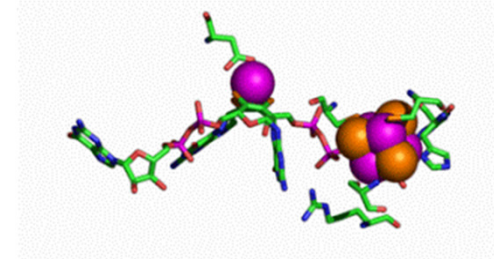
- Podem ser classificadas em três tipos:
 - Campo de força,
 - Empíricas
 - Conhecimento



- Utilizam modelos físicos
- Soma das energias de interação entre os ligantes e a proteína.
- Os termos de energia são divididos em termos:
 - ❖ Não-ligados → associados a interações de van der Waals, eletrostáticas e ligações de hidrogênio.
 - ❖ Ligados → representando normalmente a energia associada à torção de ligações químicas.

Campo de Força	Sistemas Moleculares
EFF	Alcanos
MM3	Geral
CVFF	Geral
Tripos	Geral
MMFF	Geral
Dreiding	Geral
Amber	Proteínas, ácidos nucleicos, carboidratos
CHARMM	Proteínas
GROMOS	Proteínas, ácidos nucleicos, carboidratos
MOMEK	Compostos de coordenação

- Complexos receptor-ligante com estruturas tridimensionais e afinidades conhecidas.
- Ajuste para maior acurácia.
- Diferenciação – Número de termos utilizados para modelar as equações

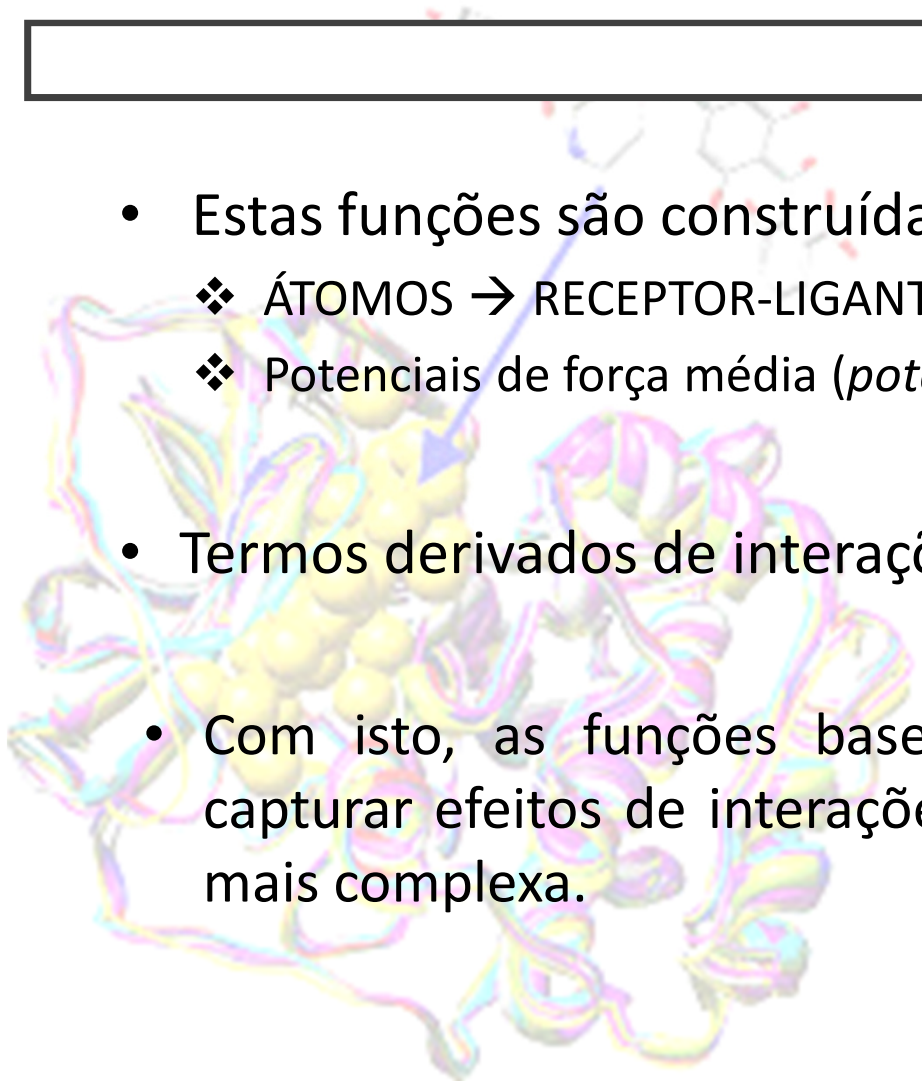


ChemScore

X-Score

GlideScore

- Estas funções são construídas a partir de análises estatísticas
 - ❖ ÁTOMOS → RECEPTOR-LIGANTE → EXPERIMENTAL
 - ❖ Potenciais de força média (*potentials of mean force*, PMF)
- Termos derivados de interações específicas pré-definidas.
- Com isto, as funções baseadas em conhecimento tendem a capturar efeitos de interações mais específicas e de modelagem mais complexa.



BASEADAS EM CONHECIMENTO

DESVANTAGEM

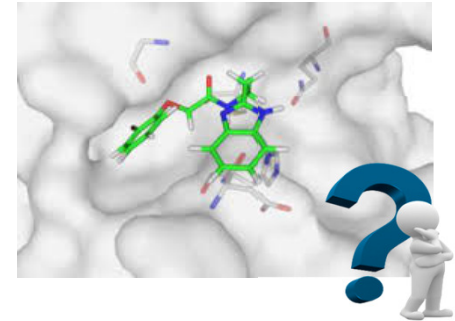
- Dependem de um conjunto de treinamento bastante amplo para a parametrização.
- Interações podem estar mal representadas
- Uso destas funções restrito.

VANTAGEM

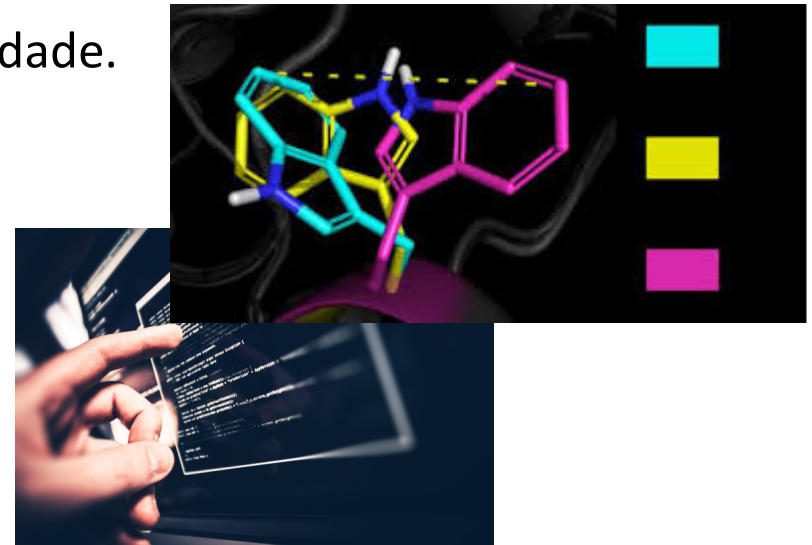
- Simplicidade de seus termos
- Rápidas comparadas as funções empíricas.

Ex.: DrugScore, RF-Score e PMF.

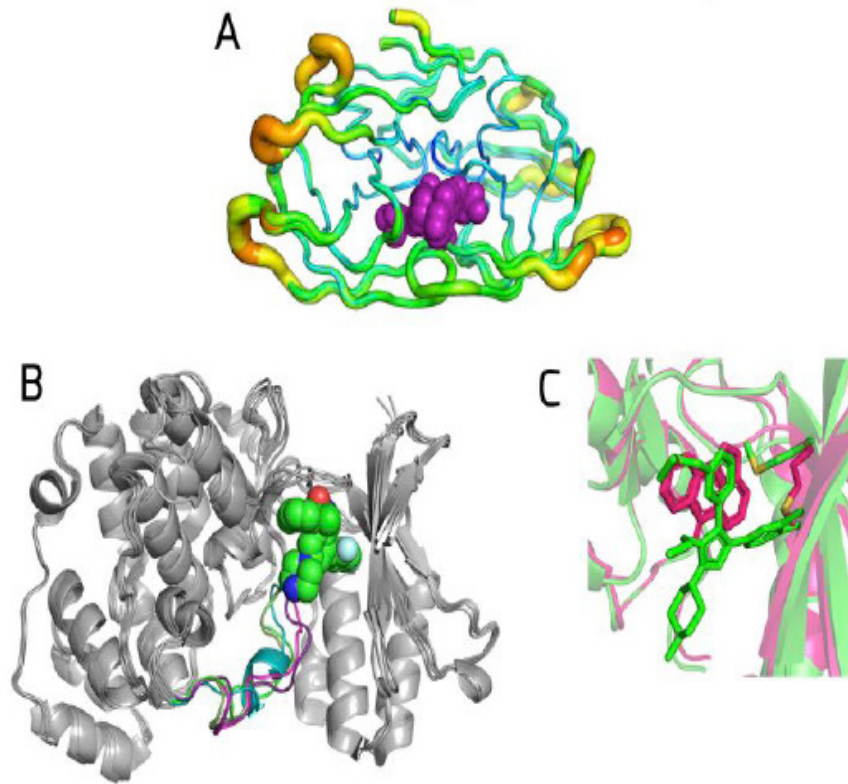
- Não existe um função de avaliação universal
- Para obter maior eficiência e confiabilidade:
 - Utilizar a função de avaliação que mais se adequa ao problema a ser pesquisado.
 - *Redocking*
- Análise qualitativas → Detecção de falsos-positivos



- Desafio → Atracamento molecular
- Impulsão de desenvolvimento de novas metodologias:
 - Importância no estudo de fármacos: Receptor +Ligante
 - Algoritmos computacionais: Maior possibilidade.



FLEXIBILIDADE DA PROTEÍNA



Movimentos locais → cadeias laterais de resíduos de aminoácidos

Movimentos de média escala → rearranjo de alças ou reposicionamento de hélices

Movimentos de grande escala → Domínios da proteína.

Figura 3-9: Graus de flexibilidade do receptor: (A) mobilidade do esqueleto peptídico da enzima protease do HIV-1, (B) diversas conformações de alça no sítio de ligação do ATP à enzima MAP cinase p38, e (C) mudança conformacional da cadeia lateral de resíduo na enzima cinase JNK3, influenciada por diferentes inibidores.

- Tipo de movimento determina a metodologia empregada.

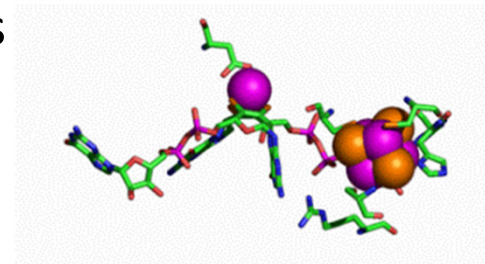
A

- i. Métodos associados ao mecanismo de encaixe induzido

- ii. Métodos associados ao mecanismo de conjunto de conformações (*ensemble docking*, em inglês)

B

- iii. Métodos híbridos, que levam os dois tipos de mecanismos e procuram considerar um amplo espectro de movimentos



MÉTODOS ASSOCIADOS AO MECANISMO DE ENCAIXE INDUZIDO

- Algoritmo de otimização local é utilizado para introduzir certa acomodação proteína-ligante no processo de atracamento

- ❖ Baseados na minimização do gradiente

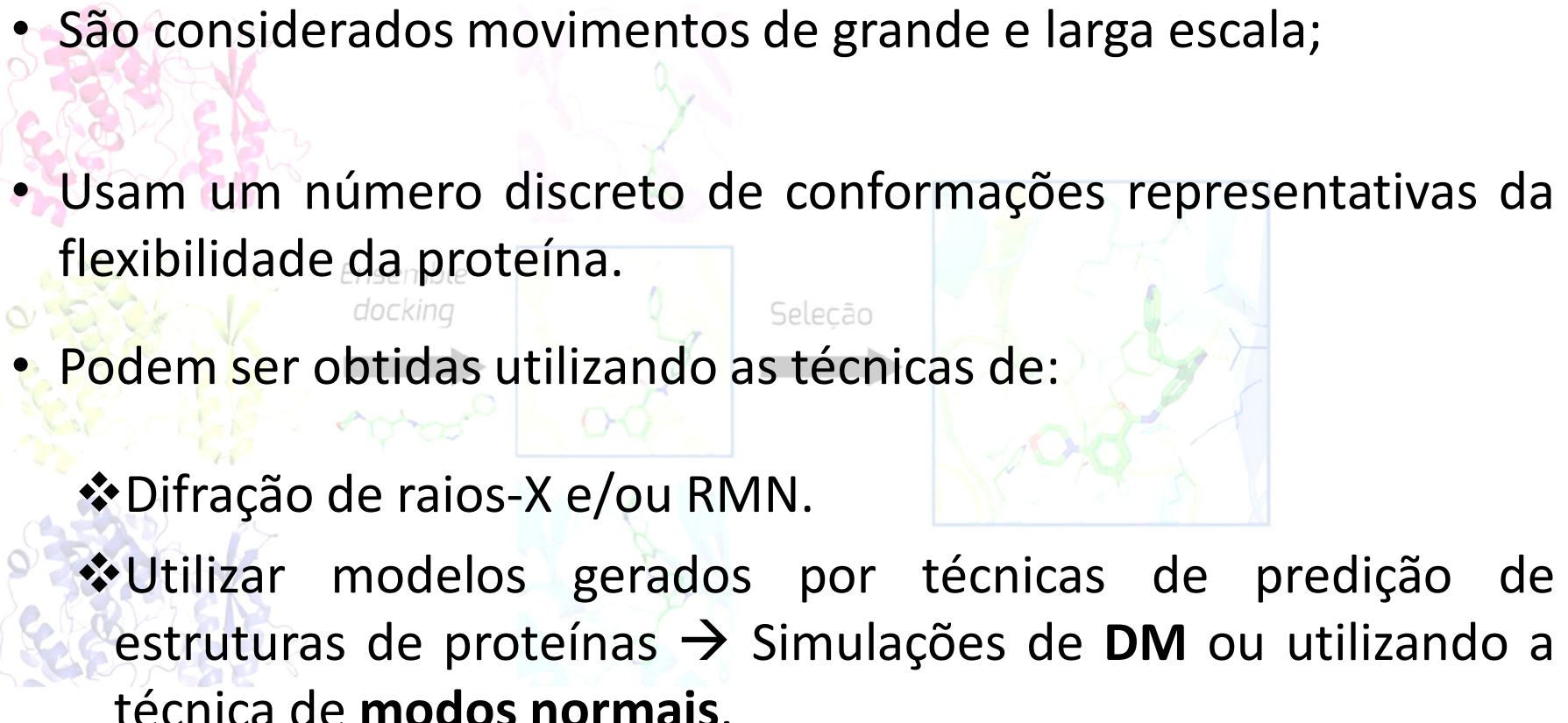
- ❖ Monte Carlo

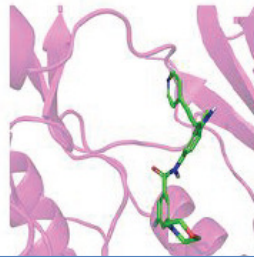
Reinvestigar as configurações ligante-proteína geradas durante o processo de busca.

MÉTODOS ASSOCIADOS AO MECANISMO DE ENCAIXE INDUZIDO

PROGRAMA	CARACTERÍSTICAS
Prodock	utiliza a minimização por gradiente durante o processo de busca para incorporar a flexibilidade em regiões da cadeia principal da proteína.
ICM/IFREDA	utiliza o método de Monte Carlo seguido de minimização de energia para otimizar cadeias laterais e/ou partes flexíveis do esqueleto peptídico.
AutoDock4 e GOLD	utilizam algoritmos genéticos para introduzir flexibilidade nas cadeias laterais de resíduos.
ROSETTALIGAND	utiliza o método de Monte Carlo para explorar simultaneamente os graus de liberdade associados ao ligante, às cadeias laterais dos resíduos e ao esqueleto peptídico da proteína.

MÉTODOS ASSOCIADOS AO MECANISMO DE CONJUNTO DE CONFORMAÇÕES

- São considerados movimentos de grande e larga escala;
 - Usam um número discreto de conformações representativas da flexibilidade da proteína.
 - Podem ser obtidas utilizando as técnicas de:
 - ❖ Difração de raios-X e/ou RMN.
 - ❖ Utilizar modelos gerados por técnicas de predição de estruturas de proteínas → Simulações de **DM** ou utilizando a técnica de **modos normais**.
- 



Há evidências significativas que associam o uso de **MÚLTIPLAS CONFORMAÇÕES** ao aumento significativo da probabilidade de obter **SUCESSO** em estudos de **ATRACAMENTO MOLECULAR**.

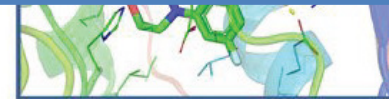
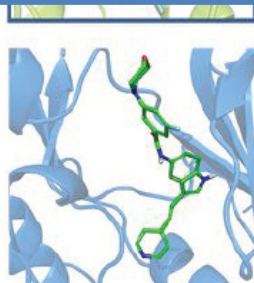
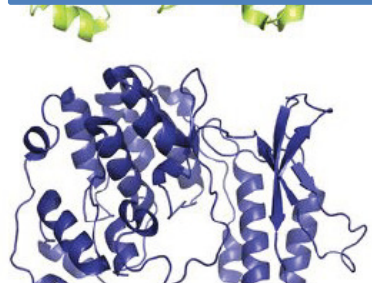


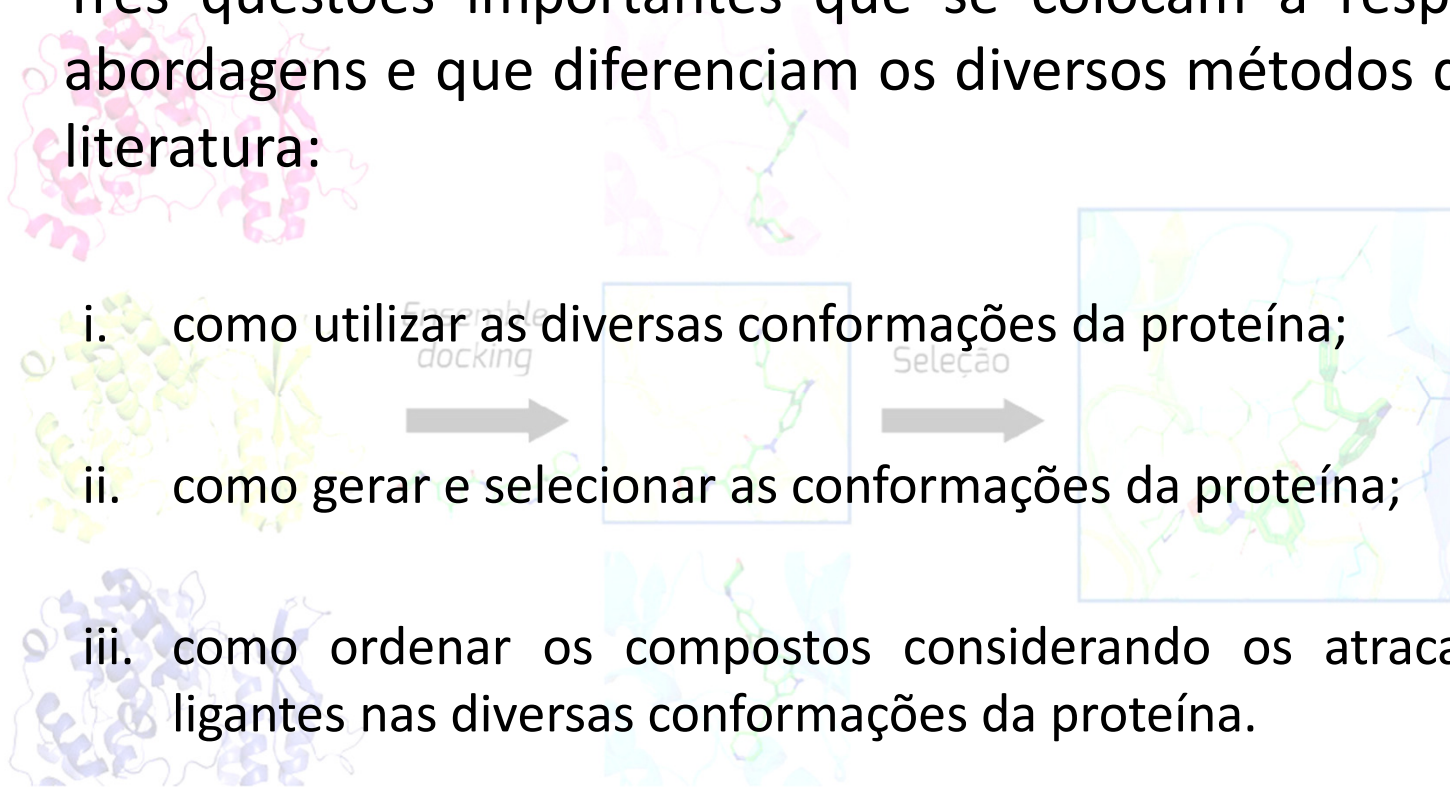
Figura 7-9: Atracamento molecular utilizando conjunto de conformações (adaptado de Guedes e colaboradores, 2013).

- Três questões importantes que se colocam a respeito destas abordagens e que diferenciam os diversos métodos descritos na literatura:

i. como utilizar as diversas conformações da proteína;

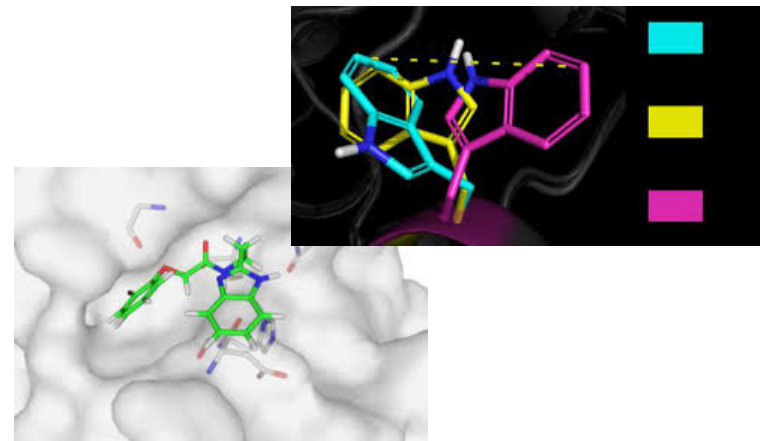
ii. como gerar e selecionar as conformações da proteína;

iii. como ordenar os compostos considerando os atracamentos dos ligantes nas diversas conformações da proteína.



i. UTILIZAÇÃO DAS CONFORMAÇÕES

- Considerar cada conformação da proteína como rígida
- Agrupamento por semelhança ou *construção de clusters*:
- Grades de energia combinadas.



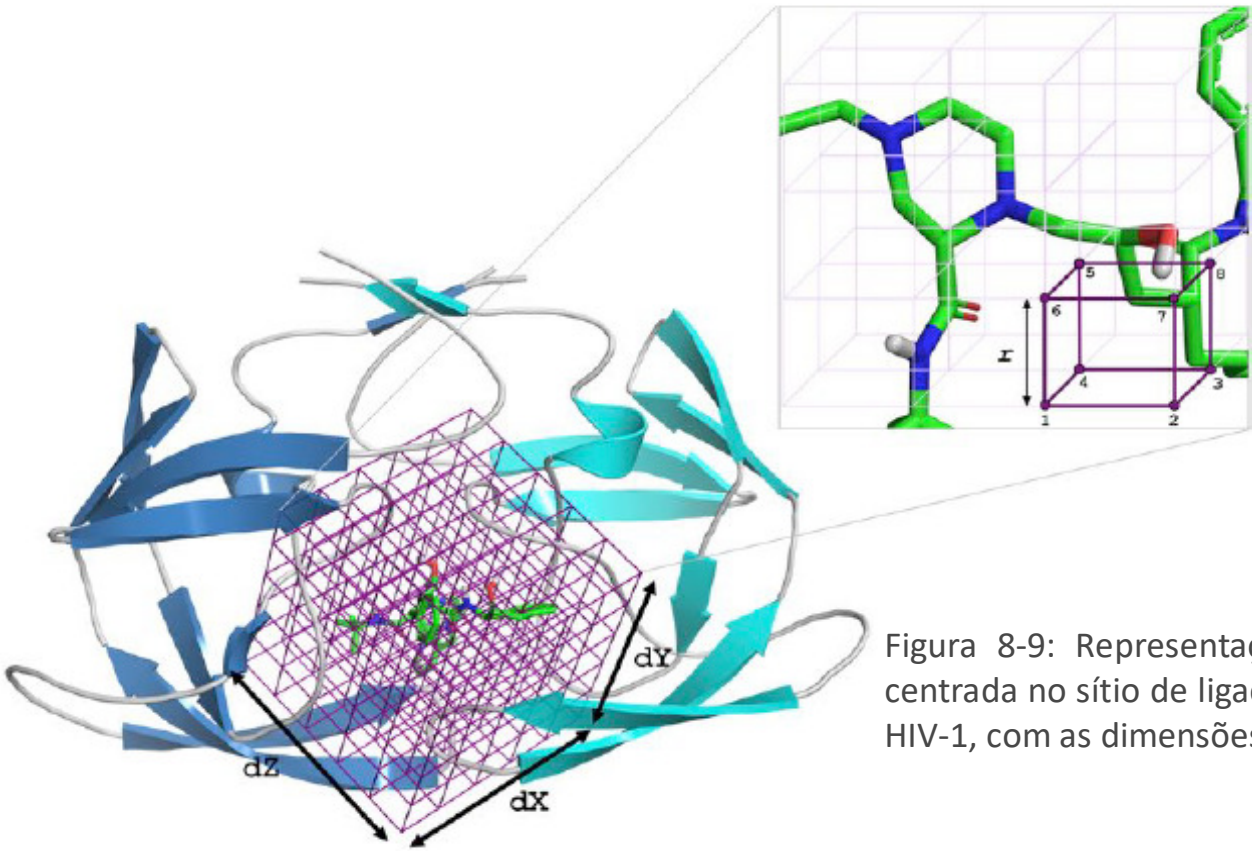


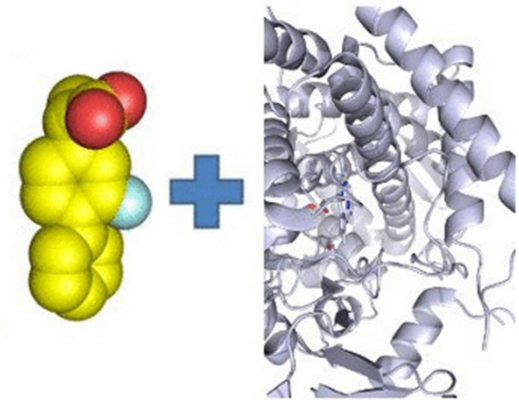
Figura 8-9: Representação de uma grade de energia cúbica centrada no sítio de ligação do inibidor indinavir da protease do HIV-1, com as dimensões de cada eixo (dX , dY e dZ).

PROGRAMA	OBSERVAÇÕES
DOCK	Primeiro a implementar conjuntos de grades de energia para a inclusão da flexibilidade da molécula receptora.
FlexE	Cada ligante é atracado em cada uma das conformações representativas da flexibilidade da proteína.
OUTROS PROGRAMAS QUE USAM GRADE DE ENERGIA	
GOLD, Glide, AutoDock Vina e DockThor	



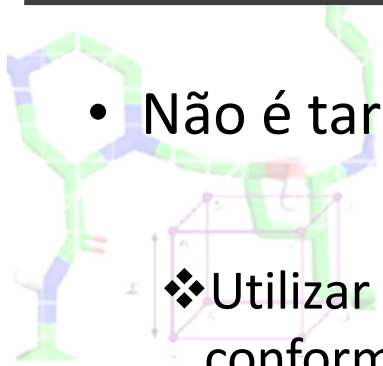
ii. GERAÇÃO DE CONFORMAÇÕES DA PROTEÍNA:

- Flexibilidade local de um receptor
- Movimentos de mais larga escala
- Modos de ligação podem ser obtidos → resultados experimentais ou de simulações de atracamento molecular.



iii. ORDENAÇÃO DOS COMPOSTOS

- Não é tarefa fácil – Grande desafio

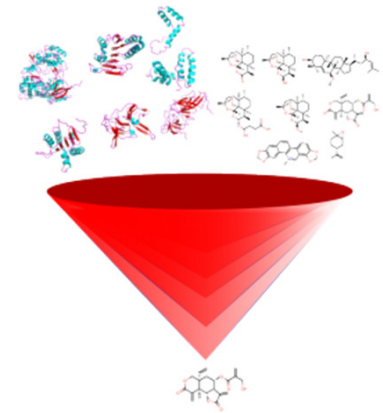


❖ Utilizar a média das energias dos ligantes com relação às múltiplas conformações da proteína.

❖ Considerar a melhor/menor energia obtida por um ligante ao interagir com determinada conformação → Possibilidade de novos fármacos.

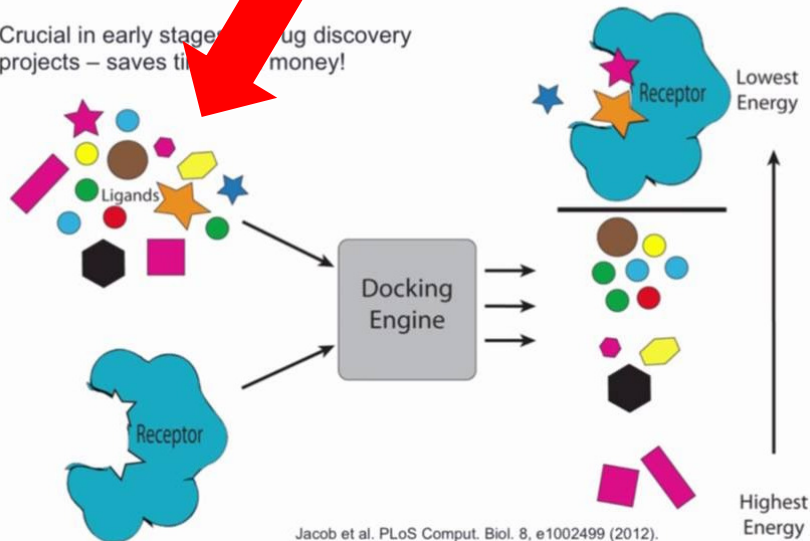
TRIAGEM VIRTUAL (*virtual screening*)

- Metodologia de atracamento molecular em larga escala, através da qual muitos **ligantes** são avaliados no sítio de ligação de um **receptor**;
- Busca de moléculas candidatas a novos fármacos;
 - rápidas, eficazes e de baixo custo.
- Complementar os resultados obtidos pela triagem experimental (*high-throughput screening*);
- Bancos: ZINC, BindingDB, PubChem, SuperNatural e ChEMBL.



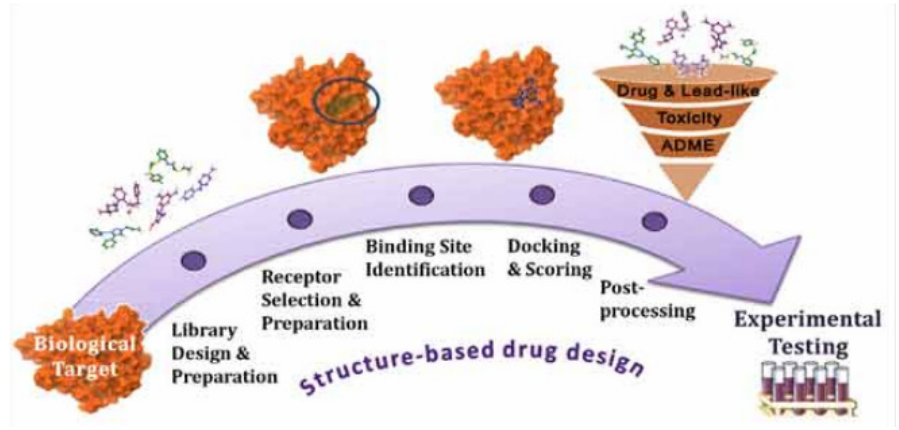
Virtual Screening

Crucial in early stages of drug discovery projects – saves time and money!

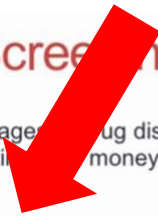


Jacob et al. PLoS Comput. Biol. 8, e1002499 (2012).

Fonte: <https://www.youtube.com/watch?v=ST-v5PbyVPY>



Fonte: <https://www.nature.com/articles/aps2009159>



- Propriedades físico-químicas definidas pelo usuário;
 - Ex.: número de ligações rotacionáveis e logP ou, em alguns deles, desenhar o fragmento desejável na estrutura dos ligantes;
 - objetivo de reduzir o número de compostos a serem analisados pela triagem virtual, especificando o perfil desejado para estes ligantes.
- Geralmente o banco fornece uma tabela com as principais propriedades químicas dos compostos;
 - Construção de uma biblioteca de ligantes própria - FAF-Drugs.

Baseadas na estrutura do Receptor (*structure-based*)

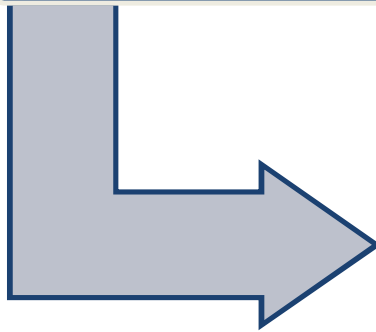
- Com base na **estrutura tridimensional do ligante**;
- Mais utilizado quando a estrutura tridimensional da molécula receptora está disponível com boa qualidade;

Baseadas na estrutura do Ligante (*ligand-based*)

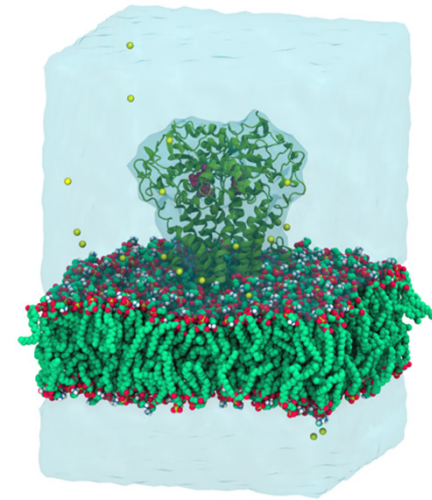
- Quando não é possível obter a **estrutura tridimensional do receptor**;
- Esta abordagem consiste na análise de similaridade de propriedades estruturais e físico-químicas de compostos ativos e inativos.
- estudo da relação estrutura-atividade (SAR, *structure–activity relationship* ou QSAR, *quantitative structure–activity relationship*) e a modelagem farmacofórica.

Validação do Protocolo

Avaliar se o método de busca e a função de avaliação escolhidos são capazes de reproduzir o modo de ligação experimental de compostos originalmente complexados com o receptor alvo.



Análise que deve ser feita é a capacidade de o protocolo diferenciar as moléculas ativas das inativas, conhecidas como casos falso-positivos





structure-based virtual screening powered by fast and flexible ligand docking

idock is also available as a [standalone tool](#) for users to run locally on their computers with their own screening library

Submitted jobs

Description	Compounds	Submitted	Status	Progress	Result
T0R_330~330	982,442	2018/03/04 23:18:15	Queued for execution	0.00000%	🕒
T0R_331~350	906,079	2018/03/04 23:19:09	Queued for execution	0.00000%	🕒
T0R_351~371	886,100	2018/03/04 23:19:54	Queued for execution	0.00000%	🕒

- A descoberta e planejamento de novos fármacos é um processo muito caro e muito demorado.
 - novo fármaco ao mercado são necessários de 10 a 20 anos e o custo estimado é de cerca de 800 milhões de dólares.
- Abordagens *in silico* que possam reduzir estes custos e acelerar o processo de descoberta e planejamento de novos fármacos;
 - moléculas que foram descobertas/otimizadas utilizando técnicas computacionais e que estão na fase de ensaios clínicos ou que já foram aprovadas para uso terapêutico.

