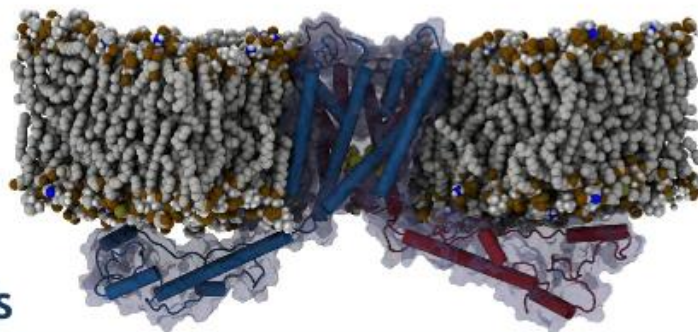




V Escola de Modelagem Molecular em Sistemas Biológicos

23-27 de Agosto - LNCC - Petrópolis



# Minicurso: Estudos de Docking Proteína-Ligante


Camila S. de Magalhães – UFRRJ, [camilasm@ufrrj.br](mailto:camilasm@ufrrj.br)

Diogo Marinho – LNCC, [diogoma@lncc.br](mailto:diogoma@lncc.br)

Isabella Alvim – LNCC, [isabella@lncc.br](mailto:isabella@lncc.br)

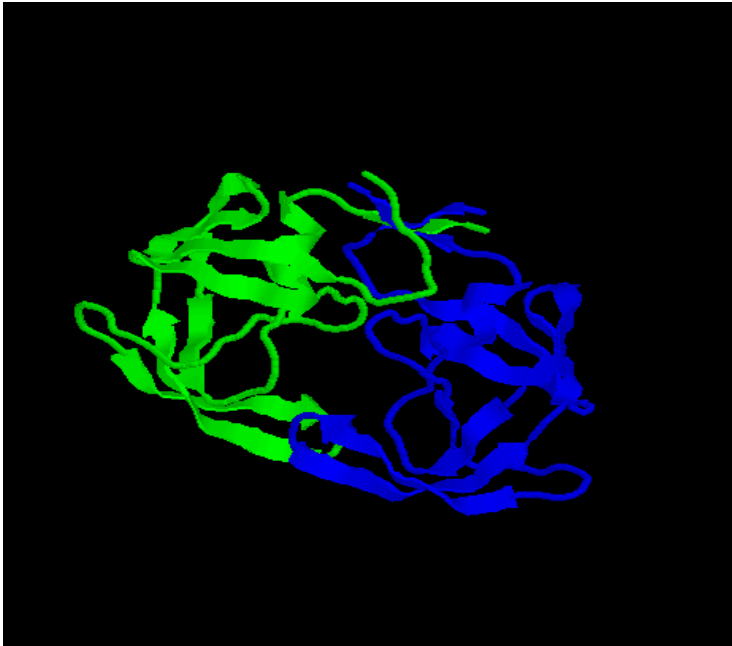
Laurent E. Dardenne – LNCC, [dardenne@lncc.br](mailto:dardenne@lncc.br)

# Programa do Curso

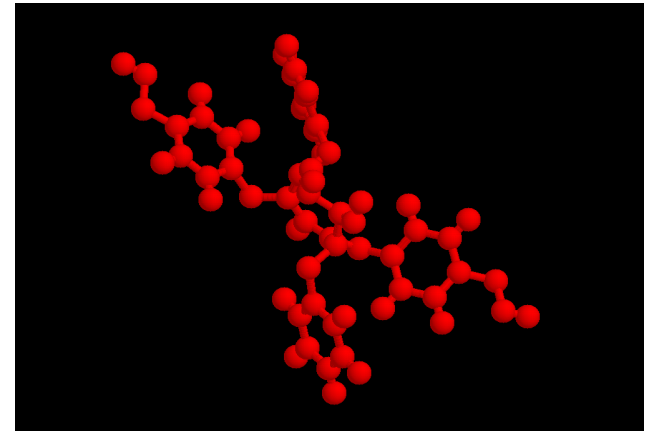
- ▶ **1ºDia:**
    - Introdução: Docking Molecular
    - Preparação dos arquivos da proteína e do ligante
  - ▶ **2ºDia:**
    - Docking com AutoDock 4
  - ▶ **3ºDia:**
    - Virtual Screening com AutoDock
  - ▶ **4ºDia:**
    - Docking com diferentes protocolos e outros complexos
- 

# O Problema de Docking

- ▶ Dada uma estrutura de **proteína**, como uma pequena molécula **ligante** interage com ela?



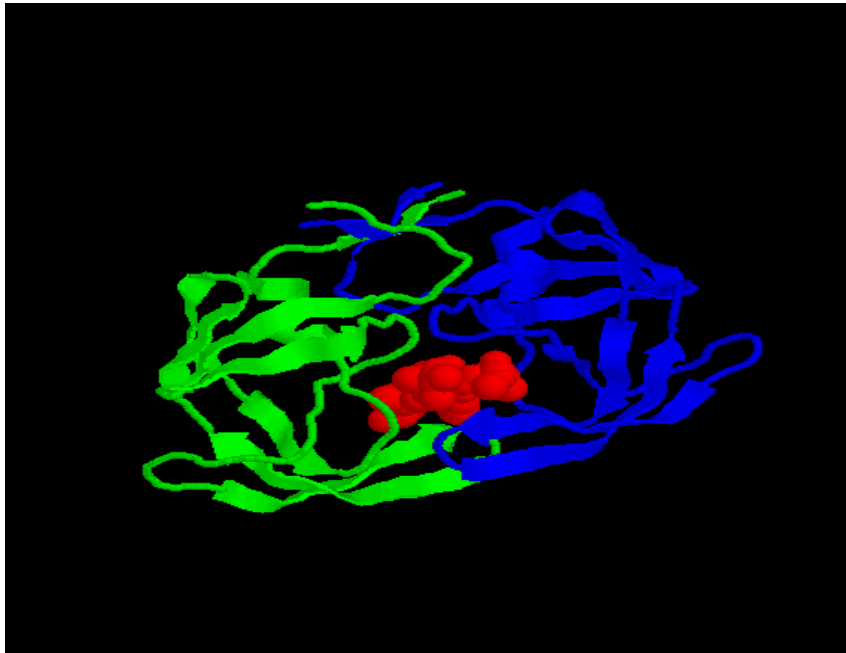
Proteína



Ligante

# O Problema de Docking

- ▶ Complexo Proteína–ligante:



# O Problema de Docking

- ▶ O que é Docking?
  - Predição do modo de ligação de uma pequena molécula ligante (**composto/fármaco**) no sítio ativo (região de ligação) de uma proteína (**alvo molecular/receptor**)
  - Obtenção de uma estimativa da afinidade dessa ligação (“Scoring”)

# Por que Métodos de Docking são Importantes?

- ▶ Desenho Racional de Fármacos Baseado em Estrutura
- ▶ Possibilidade de redução do tempo e dos custos associados ao desenvolvimento de um novo medicamento
- ▶ Área de pesquisa altamente ativa e desafiante, que tem atraído o interesse de universidades, indústrias farmacêuticas e empresas de biotecnologia

# Métodos de Docking Proteína–Ligante

- ▶ Métodos de docking se diferenciam principalmente em relação:
  - ▶ Grau de Flexibilidade Molecular
  - ▶ Função de Avaliação (Scoring)
  - ▶ Método de Busca

# Métodos de Docking Proteína–Ligante

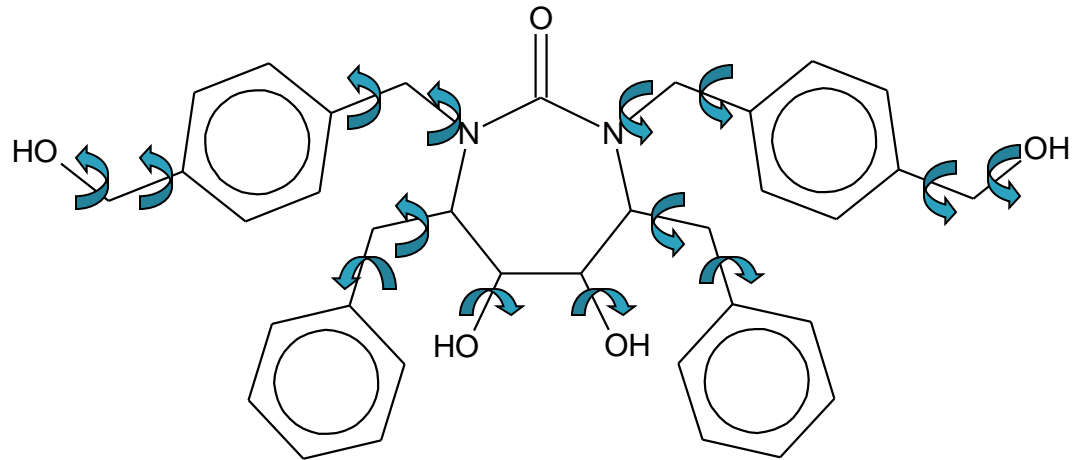
- ▶ Classificação dos Métodos de Docking:
  - Docking Rígido
    - Proteína e Ligante Rígidos
  - **Docking Ligante Flexível**
    - **Proteína Rígida e Ligante Flexível**
  - Docking Totalmente Flexível
    - Proteína e Ligante Flexíveis

# Métodos de Docking Proteína–Ligante

- Flexibilidade Molecular: Ligante

Tipos de Graus de Liberdade do Ligante:


- Translacionais
- Rotacionais
- Conformacionais



# Métodos de Docking Proteína–Ligante

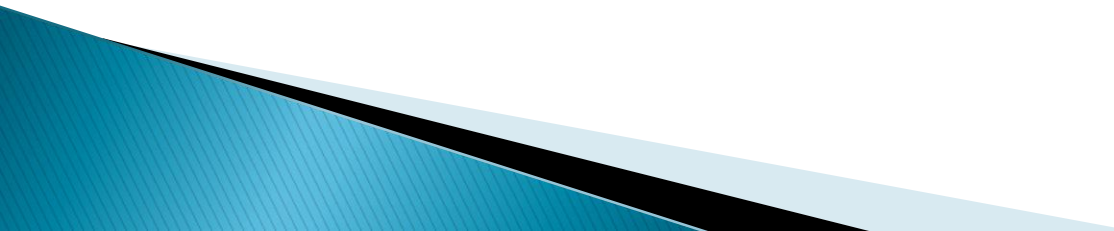
- ▶ Programa **Autodock** (<http://autodock.scripps.edu/>)
  - The Scripps Research Institute
  - Docking Ligante Flexível
  - Método de Busca: **LGA (Lamarckian Genetic Algorithm)**

# Algoritmo Genético (GA)

- ▶ Inspiração: Teoria da Evolução por seleção natural proposta por Darwin
  - ▶ Principais características:
    - Trabalham sobre uma representação da solução
    - Operam sobre uma população de soluções (indivíduos)
    - Utilizam regras de transição probabilísticas
- 

# Algoritmo Genético (GA)

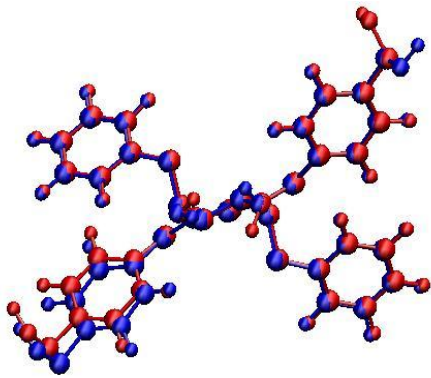
## ▶ AG para Docking:

- Representação da Solução (Conformação do Ligante)
  - Operadores Genéticos
  - Algoritmo Genético
  - Função Scoring
- 

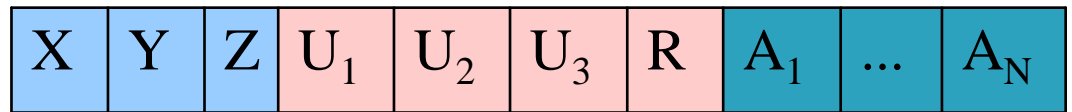
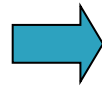
# Algoritmo Genético (GA)

- Representação do Ligante

Estrutura 3D



Cromossomo (DOF's)



Translacional

Rotacional

Conformacional

Docking Rígido

Docking Flexível

# Operadores Genéticos

- Crossover de 2-pontos

CROMOSSOMO 1

$V_1 V_2 V_3 V_4 V_5 | V_6 V_7 | V_8$

CROMOSSOMO 2

$V_1 V_2 V_3 V_4 V_5 | V_6 V_7 | V_8$

FILHO 1  $V_1 V_2 V_3 V_4 V_5 V_6 V_7 V_8$

FILHO 2  $V_1 V_2 V_3 V_4 V_5 V_6 V_7 V_8$

# Operadores Genéticos

- **Mutação**

CROMOSSOMO

$V_1 V_2 V_3 V_4 V_5 V_6 V_7 V_8$

Mutação:

$$\mathbf{V_3} = V_3 + \Delta$$

Novo CROMOSSOMO

$V_1 V_2 \mathbf{V_3} V_4 V_5 V_6 V_7 V_8$

# Algoritmo Genético (GA)

- Gera População inicial
- Seleciona Operador Genético
- Seleciona Indivíduos (pais)
- Aplica Operador Genético
- Avalia Novo Indivíduo
- Insere Novo Indivíduo (filho)

## Condição de parada:

- Número Máximo de Avaliações

# Programa Autodock

- ▶ **Algoritmo Genético Lamarckiano (LGA)**
  - Inspirado na (desacreditada!) teoria de Lamarck:
    - Características adquiridas pelo indivíduo podem ser herdadas (codificadas nos genes)
  - AG Híbrido = AG + BL
  - Substituição do resultado da busca local no cromossomo do indivíduo

# Função Scoring – Programa Autodock

## ► Função Empírica: Autodock 4

$$V = W_{vdw} \sum_{i,j} \left( \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) + W_{hbond} \sum_{i,j} E(t) \left( \frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) + W_{elec} \sum_{i,j} \frac{q_i q_j}{\epsilon(r_{ij}) r_{ij}} + W_{sol} \sum_{i,j} (S_i V_j + S_j V_i) e^{(-r_{ij}^2 / 2\sigma^2)}$$

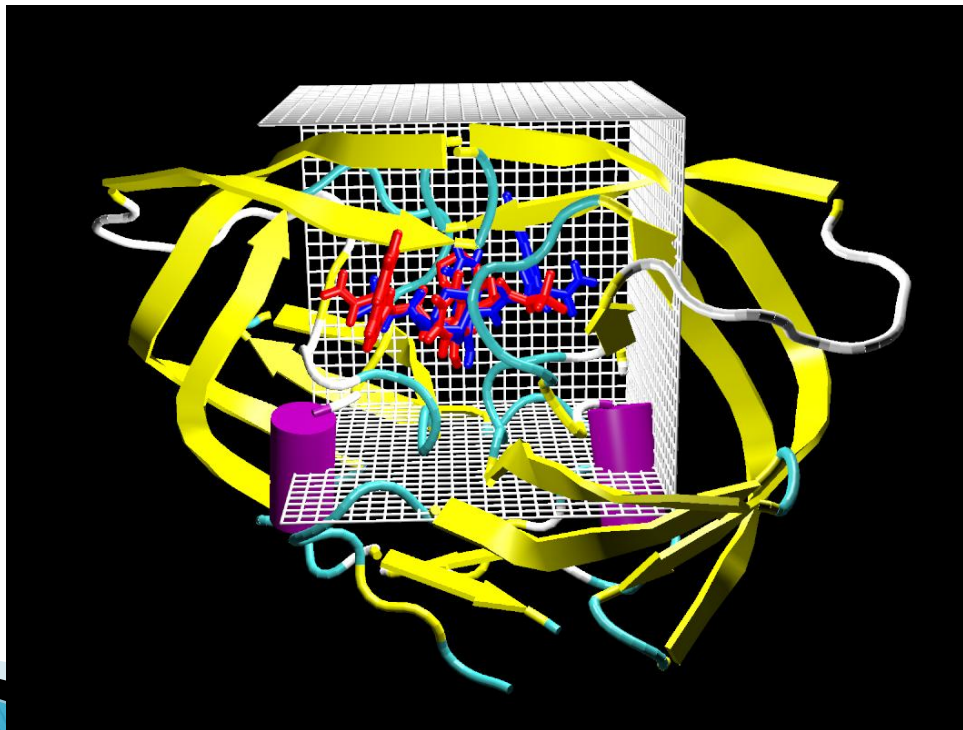
- $W_{vdw}$ ,  $W_{bond}$ ,  $W_{elec}$ ,  $W_{sol}$ , parâmetros otimizados com 188 complexos proteína–ligante com dados de afinidade experimental conhecidos ( $K_i$ )

- Para todos os pares de átomos  $i,j$ 
  - Receptor–Ligante } **alto** custo computacional!!!!
  - Ligante–Ligante

# Metodologia Baseada em Grade

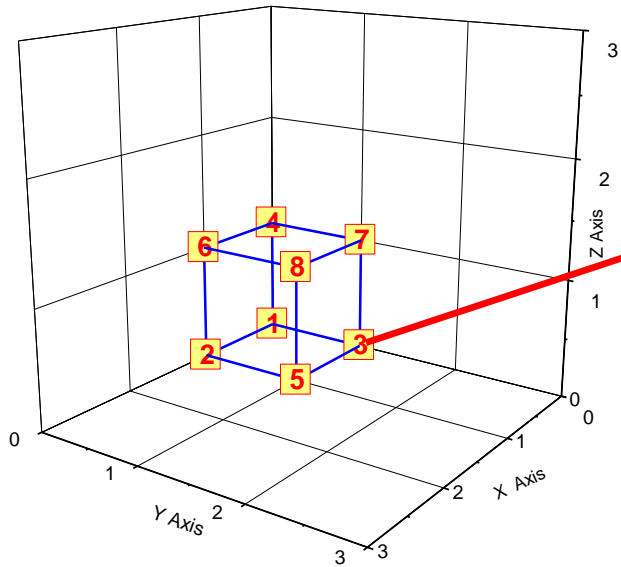
## Cálculo da função : (Grid Maps)

- ▶ Docking na Grade: A energia de interação proteína-ligante para cada conformação investigada é calculada através de uma interpolação trilinear dos valores (pré-calculados) em oito pontos localizados nos vértices de uma célula da malha construída centralizada no sítio ativo do receptor.



Receptor Rígido

# Metodologia Baseada em Grade



$$\left\{ \sum_j^{N_{\text{point}}} \sum_{i=1}^{N_{\text{Protein}}} \left\{ 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma'_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i^p}{D(r_{ij}) r_{ij}} \right\} \right\}$$



$$\sum_{j=1}^{N_{\text{Ligand}}} \left\{ V(LJ) + q_j^l V(EP) \right\} \rightarrow$$

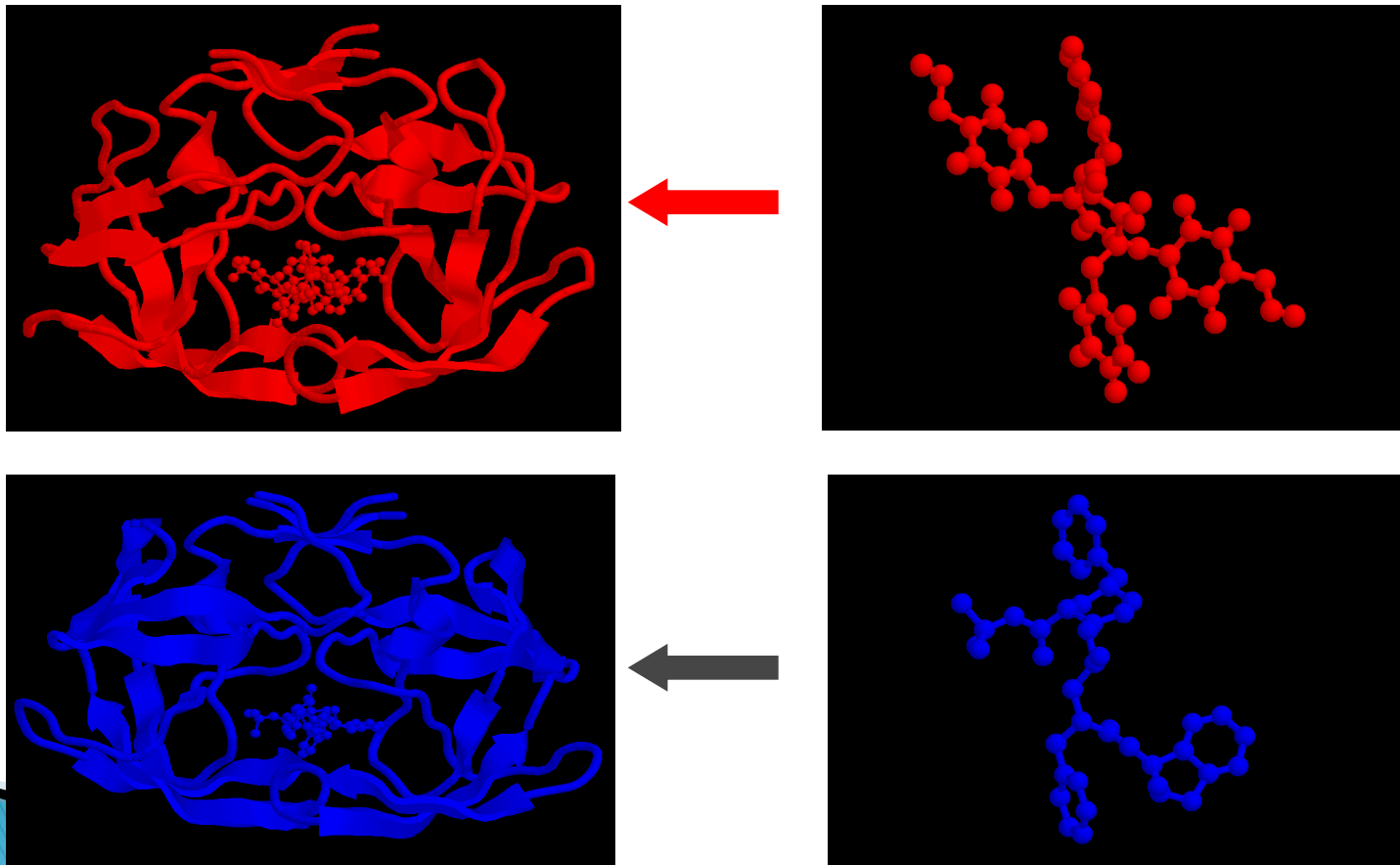
Potencial de Lennard-Jones e potencial eletrostático pré-calculado em cada ponto de cada célula da grade

# Programa AutoDock


- ▶ Grid Maps Autodock:
  - Tipo de átomo do ligante
  - Eletrostática
  - Desolvatação

# Métodos de Docking

- ▶ Avaliação do Método: Predição do Modo de Ligação
- ▶ Reproduzir estruturas experimentais (Redocking)



# Avaliação dos Métodos de Docking

- ▶ Valor do RMSD (Root Mean Square Deviation) entre coordenadas dos átomos da estrutura encontrada pelo algoritmo e da estrutura experimental:
- ▶ Se menor do que 2.0 Å  sucesso.  
(embora até 2.5 Å seja considerado um bom resultado)

$$\text{RMSD} = \sqrt{\frac{\sum_i^N (R_i - R_{ref_i})^2}{N}}$$

# Docking com o programa AutoDock

- ▶ Redocking: Indinavir (1HSG.pdb)
  - Primeiro passo:
    - Preparação dos arquivos da Proteína e do Ligante
      - Seguir o Tutorial disponível em `/home/vemmsb/1HSG/`
        - Login: vemmsb
        - Senha: emmsb

# Docking com o programa AutoDock

## ► Roteiro para 1º Dia:

1. Download dos arquivos da proteína e do Ligante
  - Visualizar a estrutura com pymol e kate
  - Separar os arquivos da proteína e do ligante
  - Adicionar átomos de hidrogênio no ligante com Babel

# Docking com o programa AutoDock

## ► Roteiro para 1º Dia:

### 2. Preparação dos arquivos da proteína

- Abrir pdb da proteína com **adt** (AutoDock Tools 4)
  - Retirar as moléculas de água
  - Adicionar hidrogênios
  - Calcular Cargas (Gasteigner)
  - Salvar arquivo .pdbqt da proteína (entrada para Autodock)

# Docking com o programa AutoDock

## ► Roteiro para 1º Dia:

### 3. Preparação dos arquivos do ligante

- Abrir pdb do ligante no **adt** (AutoDock Tools 4)
  - Detectar root
  - Definir torções (rígido: torções=0)
  - Salvar arquivo .pdbqt do ligante (entrada para Autodock)

# Docking com o programa AutoDock

## ► Roteiro para 1º Dia:

- Repetir o processo anterior para gerar diferentes arquivos .pdbqt do ligante para:
  - 6 torções mais externas (*fewest*)
  - 6 torções mais internas (*most*)
  - 14 torções (ligante totalmente flexível)

# Docking com o programa AutoDock

- ▶ Roteiro para 2º Dia: Docking Indinavir: (1HSG.pdb)
  - ▶ Menu Grid: Preparar arquivo de parâmetros para gerar a grid (.gpf)
    - Abrir o arquivo .pdbqt da proteína
    - Abrir o .pdbqt do ligante
    - Escolher dimensão da grid
    - Gerar a grid: Executar Programa Autogrid 4

# Docking com o programa AutoDock

- ▶ Roteiro para 2º Dia: Docking Indinavir: (1HSG.pdb)
  - ▶ Menu Docking: Preparar arquivo de parâmetros para Docking (.dpf)
    - Selecionar o arquivo .pdbqt da proteína
    - Escolher o ligante
    - Definir parâmetros do Algoritmo Genético
    - Salvar arquivos de parâmetros .dpf
    - Fazer o Docking: Executar Programa AutoDock 4
    - Analisar o resultado

# Docking com o programa AutoDock

- ▶ Roteiro para 3º Dia: Virtual Screening
  - ▶ Introdução a Virtual Screening
    - ZINC
    - Separar os arquivos dos ligantes
    - Preparar os ligantes
    - Determinar os tipos de átomo
    - Preparar o receptor
    - Preparar arquivos de parâmetros da grid
    - Executar Autogrid

# Docking com o programa AutoDock

- ▶ Roteiro para 3º Dia
  - ▶ Virtual Screening:
    - Validar o protocolo com um controle positivo
    - Preparar os diretórios do docking e os arquivos de parâmetros de cada ligante da biblioteca
    - Executar uma lista de dockings
    - Analisar os resultados

# Docking com o programa AutoDock

- ▶ Roteiro para 4º Dia:
  - ▶ Análise dos resultados de Docking com diversos protocolos
  - ▶ Docking com outros complexos

# Avaliações da Função Energia

## ► Recomendações AutoDock

# Graus de Liberdade Conformacionais do Ligante	ga_num_evals
0	25 000 a 250 000
1-10	250 000 a 25 000 000
> 10	> 25 000 000 (!!)

# Programa Autodock

## ▶ Principais Tipos de átomos Autodock:

- Átomo unido
- Aceptores de ligações de hidrogênio:
  - NA, AO e AS
- Hidrogênios doadores: HD
- Carbonos aromáticos: A