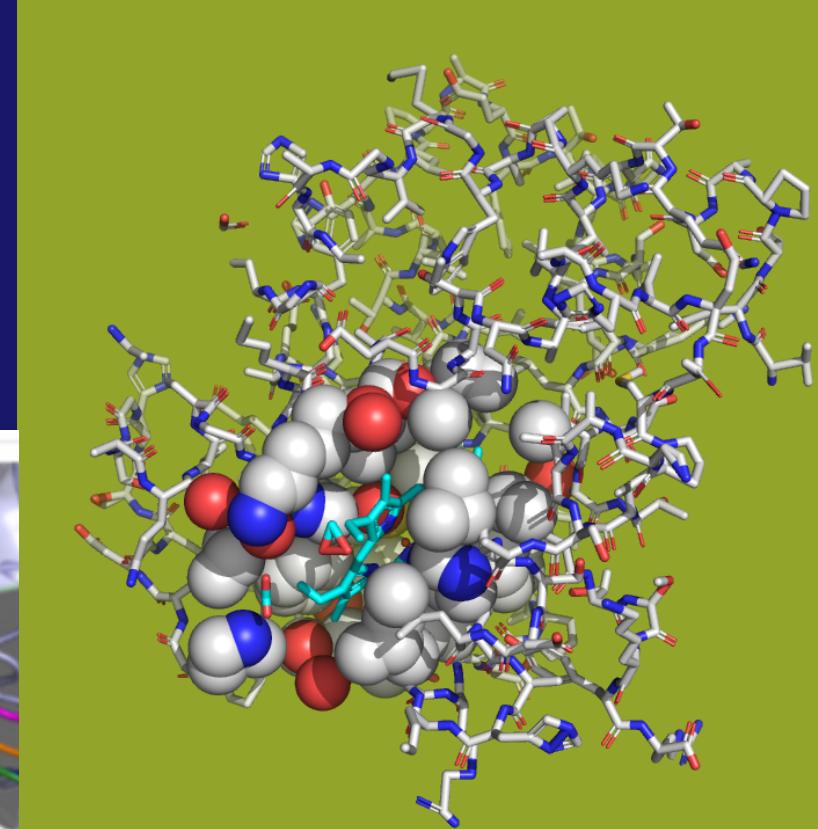
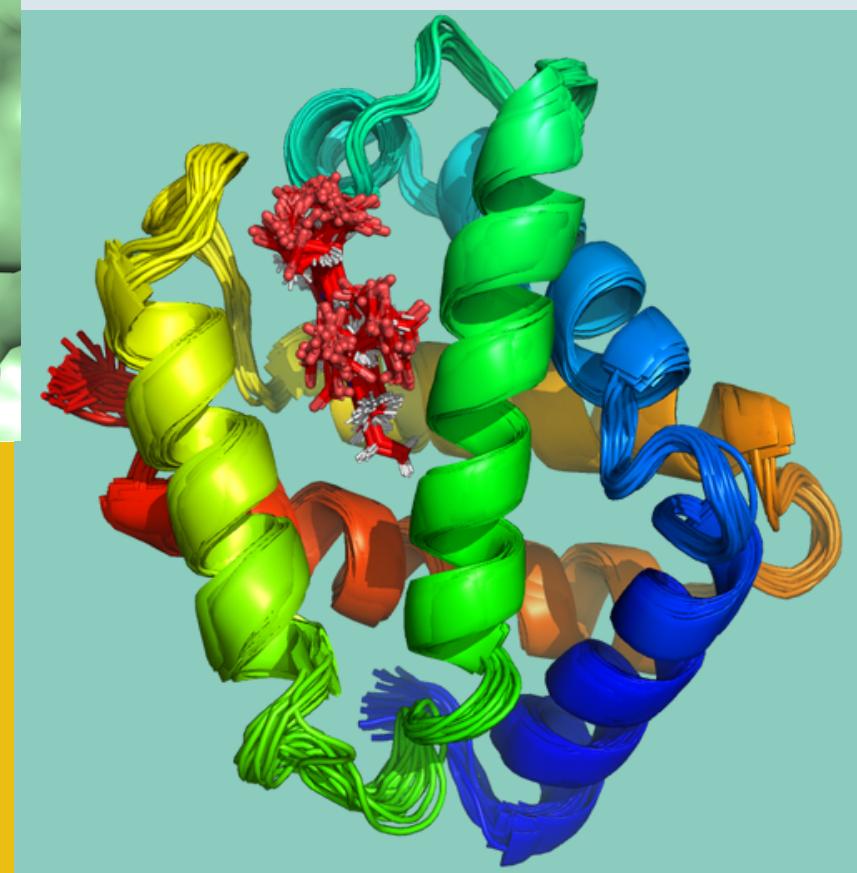
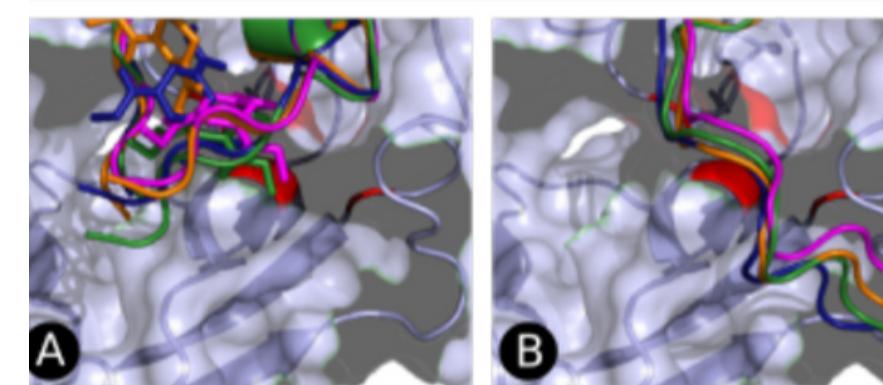
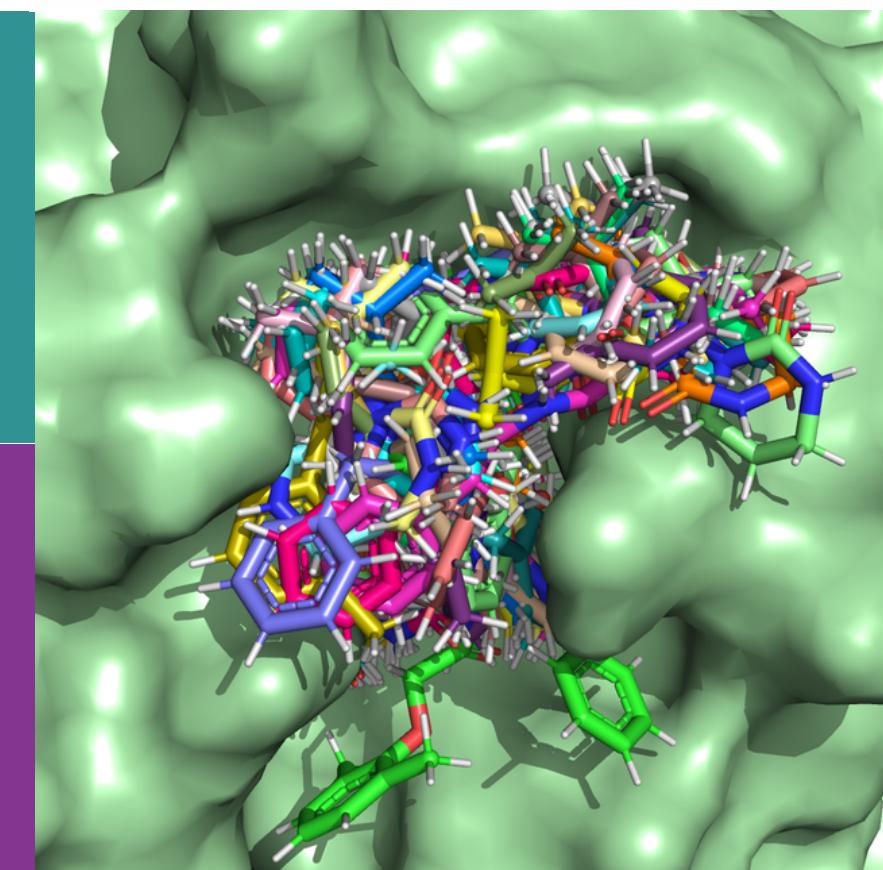
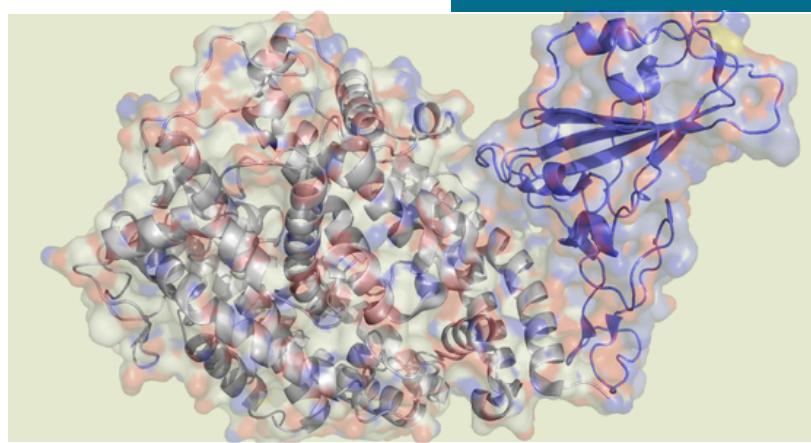


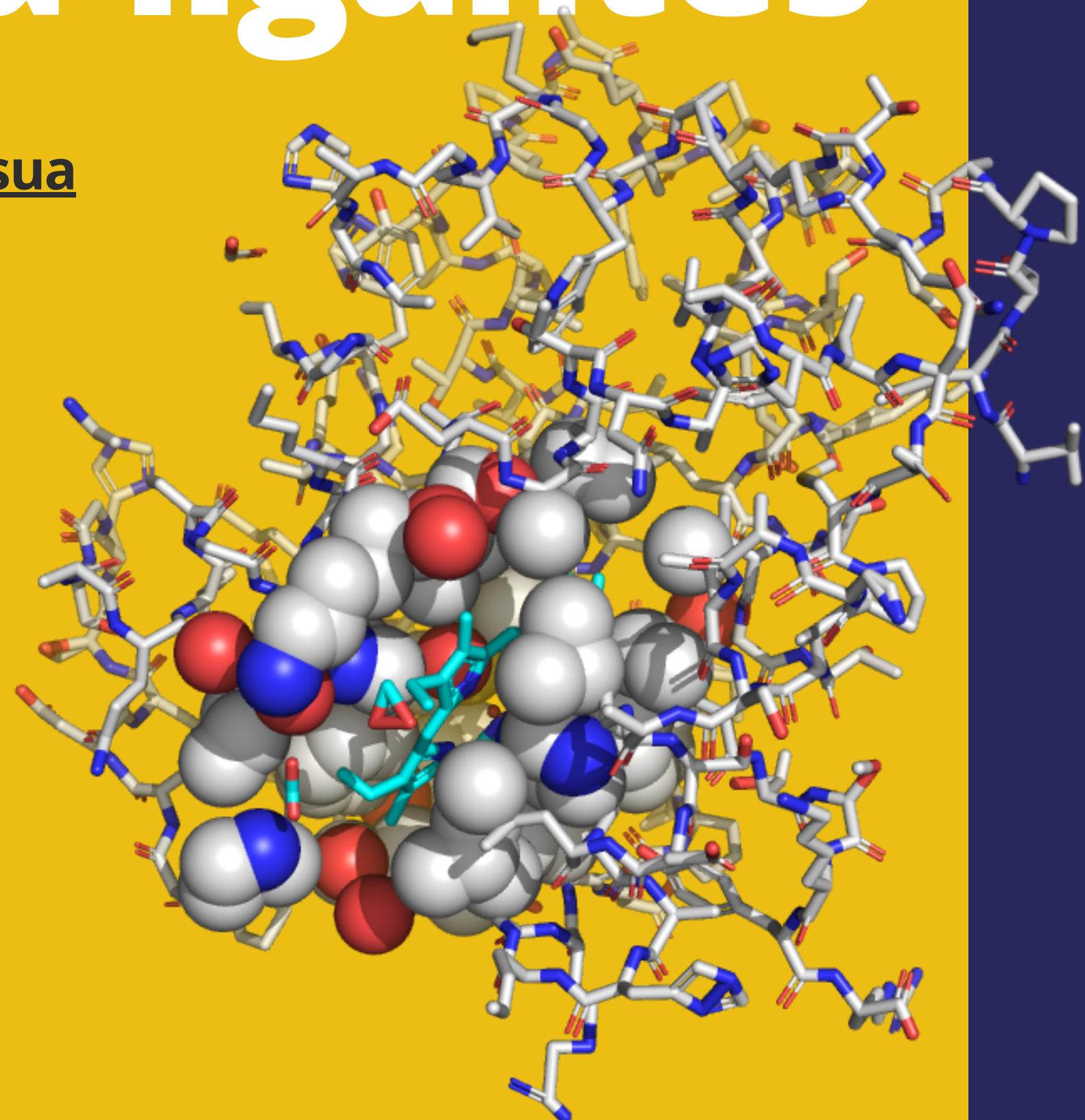
# Algoritmos de ancoragem molecular *(docking)*



# Interação proteína-ligantes

A função de proteínas se baseia na sua capacidade de interagir e se ligar reversivelmente com moléculas

- pequenas moléculas orgânicas
  - ~600 Da ou menos
    - 1 Da = 1,6605e-24 g
- Macromoléculas
- Íons



# Energia livre

- A termodinâmica nos permite usar a **caracterização macroscópica dos sistemas (conteúdo energético)** para prever a direção e a probabilidade de processos sem avaliar detalhes microscópicos
- Processos **espontâneos** envolvem o **decréscimo da energia livre do sistema** até o **equilíbrio**
  - $\Delta G = -RT \ln K_{eq}$
  - $K_{eq} = f[\text{produtos}]/f[\text{reagentes}]$ 
    - $R$  é a constante universal dos gases  $1,989 \text{ cal/mol K}$
    - $T$  é a temperatura em Kelvin (ambiente  $298 \text{ K}$ )
    - $K_{eq}$  é a constante de equilíbrio

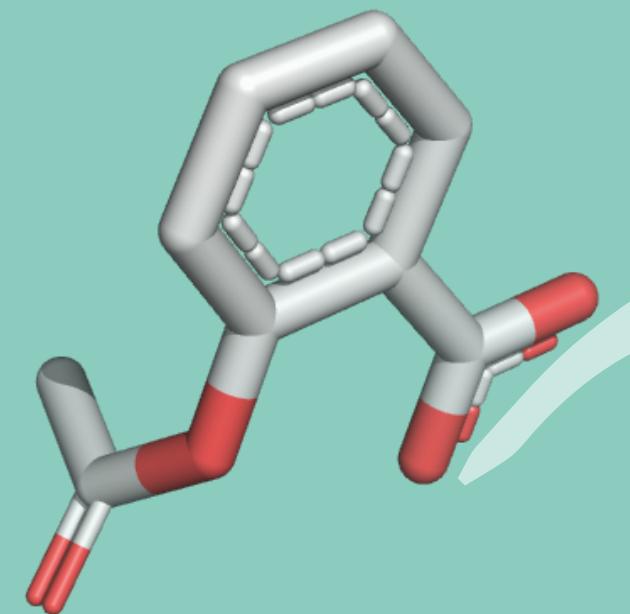
# Energia livre de ligação

- É a energia de ligação de proteínas com seus ligantes
  - $\Delta G_{lig} = GPL - (GP + GL)$ 
    - $GPL$  é a energia do complexo proteína-ligante (proteína na forma "**holo**")
    - $GP$  é a energia da proteína na forma "**apo**" (sem ligante)
    - $GL$ : é a energia do ligante
  - $\Delta G_{lig}$  tem:
    - pequena magnitude
    - forte dependência com a conformação
    - forte dependência do sistema (condições de temperatura, pressão, pH, etc)

# O problema da ancoragem molecular

## Encontrar

- o sítio de ligação da proteína
- a pose do ligante
- a afinidade
  - métrica de energia de ligação entre proteína-ligante



## Aspirina

pequena molécula

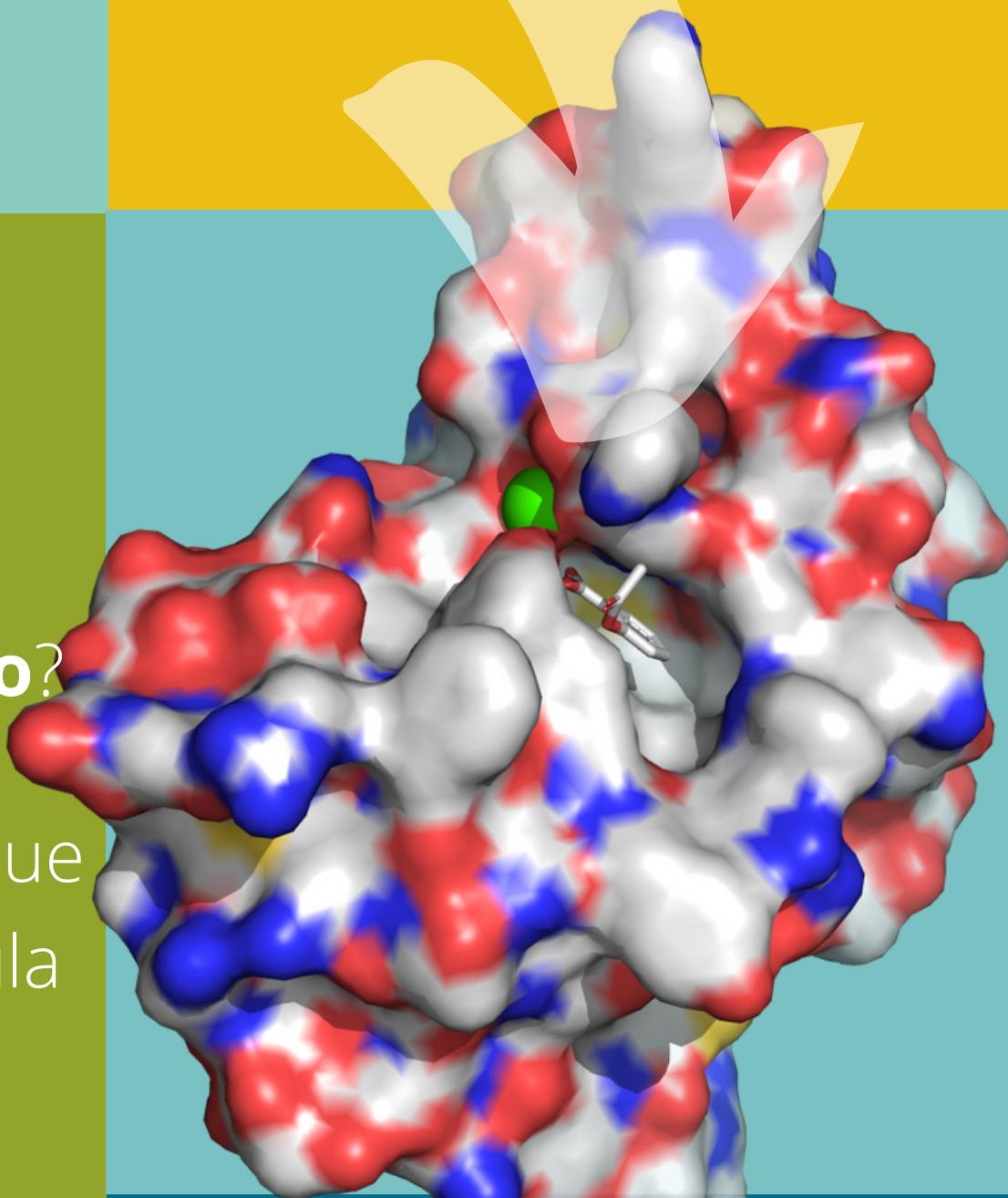
Qual a **pose**? Quais os ângulos de cada ligação rotacionável?

## Alvo

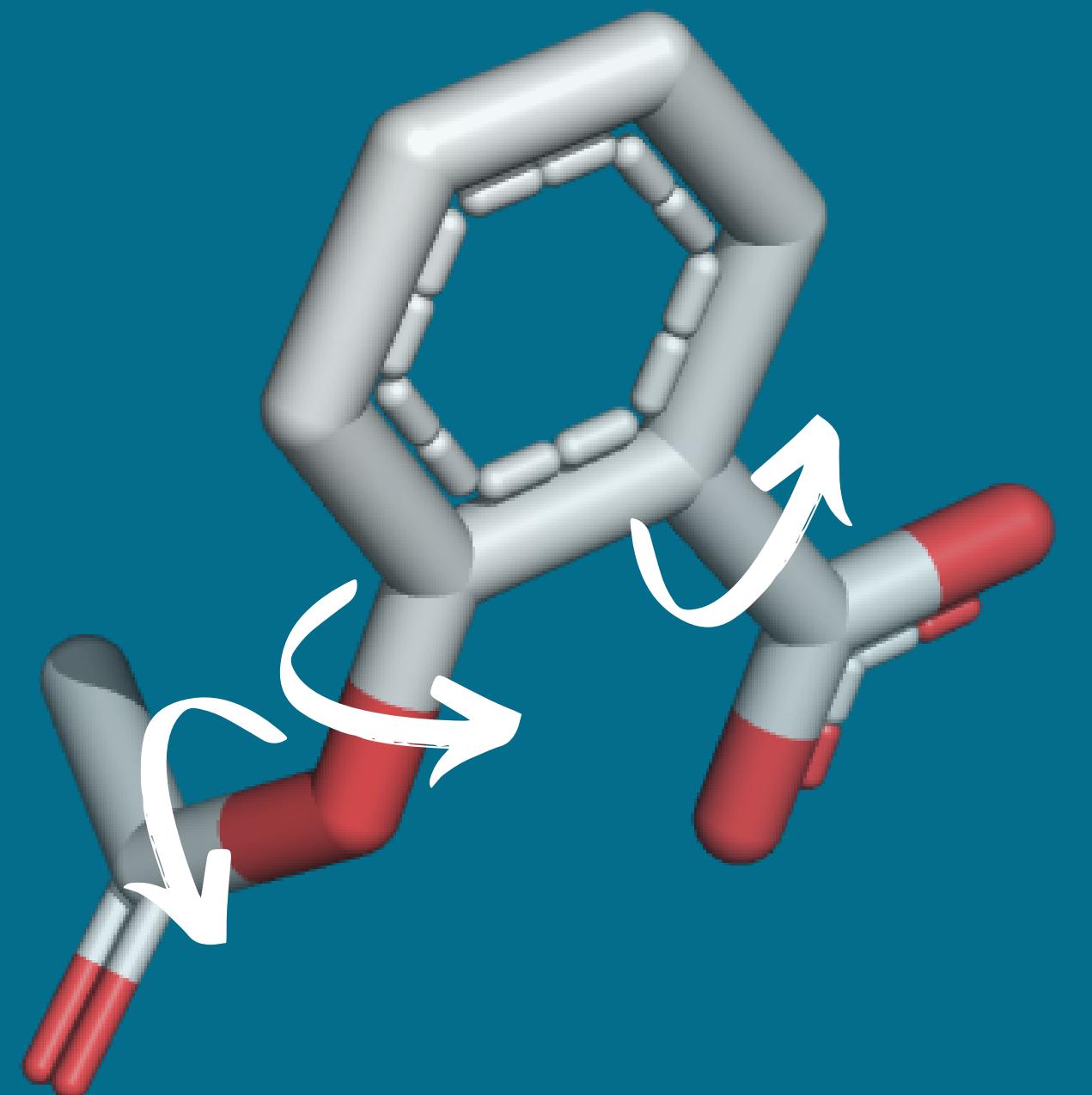
proteína

Qual o **sítio de ligação**?

Conjunto de resíduos que interage com a molécula



# Poses



Quantas ligações rotacionáveis tem a *aspirina*?

Quantos confórmeros tem a molécula?

À princípio, considerando apenas os ângulos inteiros,  
 $360^3$

# Poses



Quantas ligações rotacionáveis tem o *NAD*?

Quantos confórmeros tem a molécula?



# Algoritmo genético

- São um tipo de **algoritmo evolutivo**
  - **metaheurísticas** para **problemas de otimização**
- **Não** necessitam de **conhecimentos** sobre **como solucionar o problema**
  - apenas de uma forma de **avaliação** do resultado
- Baseiam-se na **codificação** do **conjunto das soluções possíveis**
  - e não nos parâmetros e nas regras da solução em si

# Componentes de um AG

## Função objetivo

expressão que resulta no valor que desejamos **otimizar**



## Indivíduo

representação do **espaço de busca**

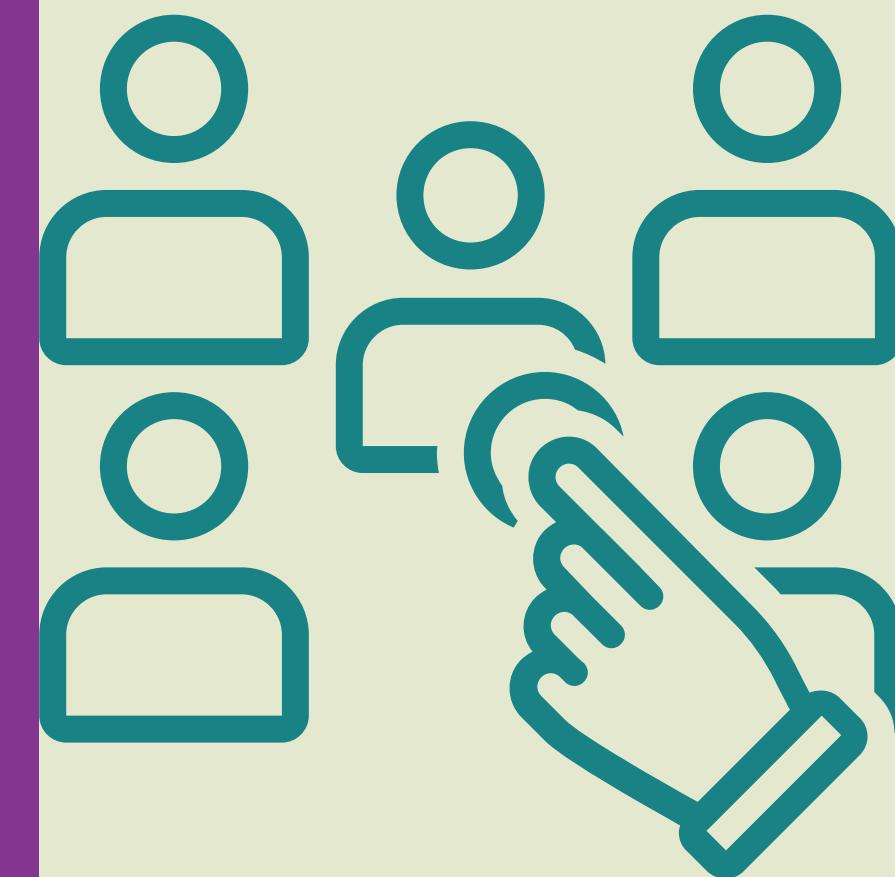
## Seleção

escolha dos indivíduos mais **bem adaptados** para sobrevivência

## Reprodução

Formada por:

- acasalamento
- recombinação
- mutação

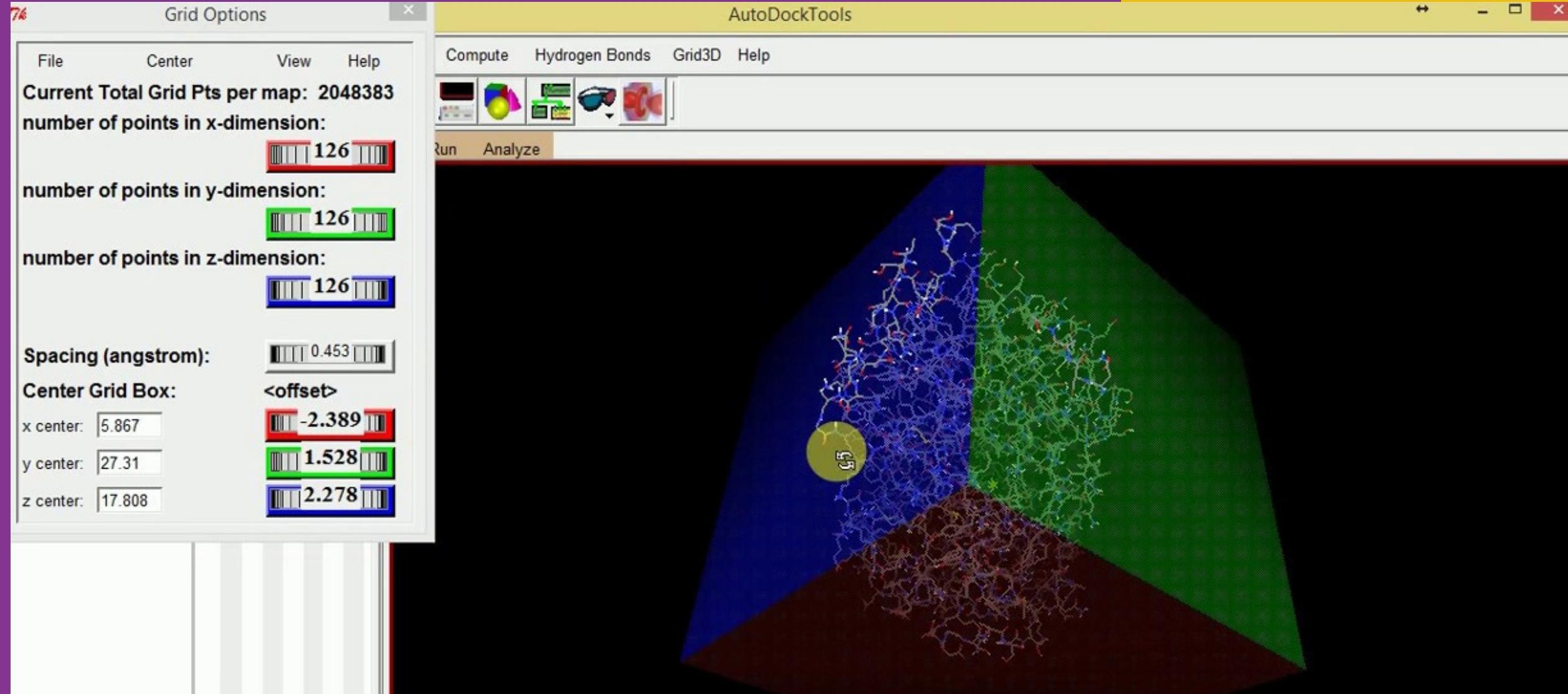




# Algoritmo genético

```
AlgoritmoGenético (população, funçãoObjetivo){  
    repita{  
        pais = seleção(população, funçãoObjetivo)  
        população = reprodução(pais)  
    }enquanto condição de parada não atingida  
    retorna o melhor indivíduo  
}
```

# Autodock



# Autodock

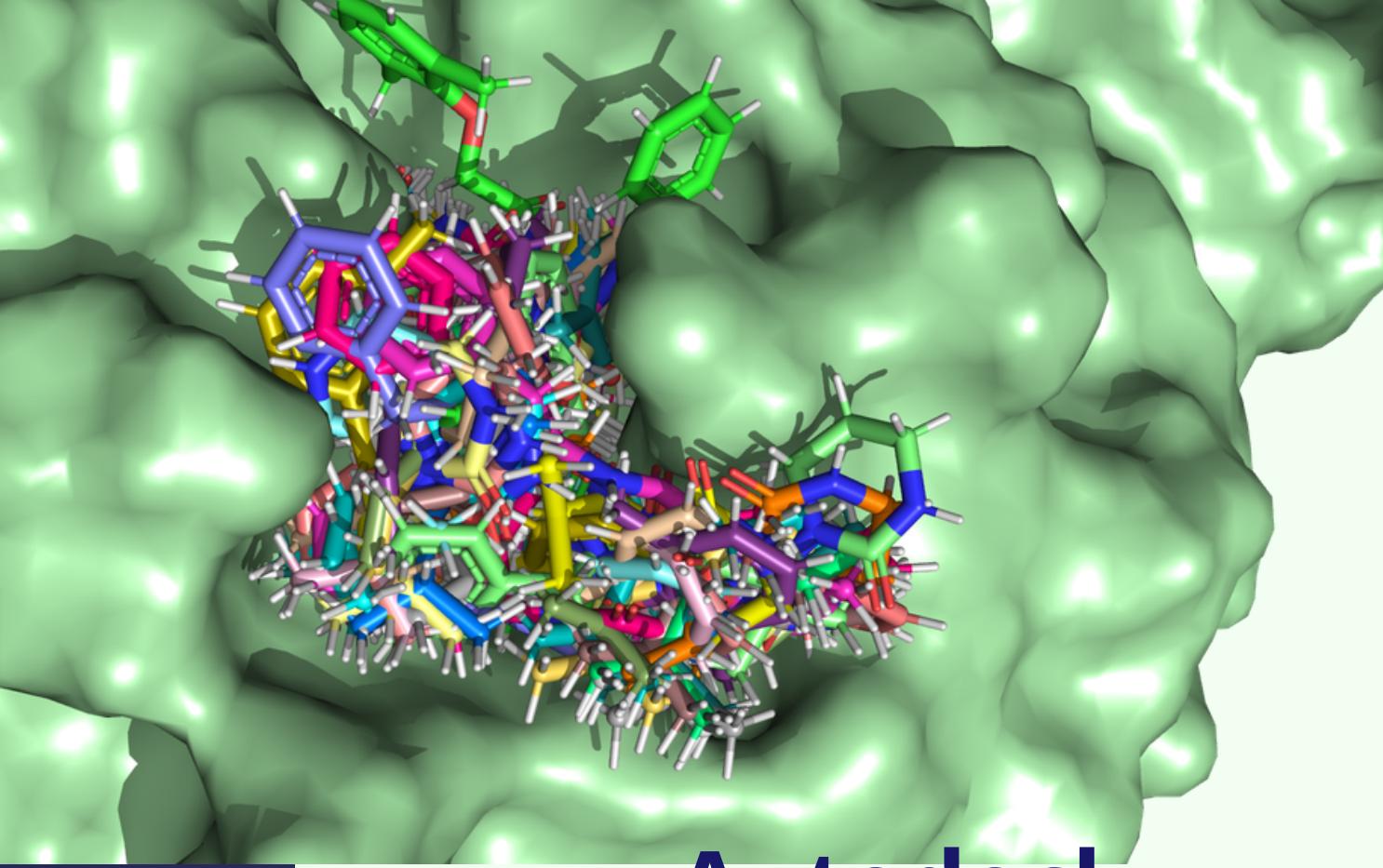
A ferramenta gratuita mais utilizada para *docking* de pequena moléculas

Usa **algoritmo genético** para otimizar a posição e a pose do ligante

Goodsell, David S., and Arthur J. Olson. "Automated docking of substrates to proteins by simulated annealing." *Proteins: Structure, Function, and Bioinformatics* 8.3 (1990): 195-202.

Morris, Garrett M., et al. "Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function." *Journal of computational chemistry* 19.14 (1998): 1639-1662.





**Autodock**

# 1 População inicial

Considerando o caso mais comum:

- proteína rígida
- molécula flexível

**O indivíduo é um confôrmero da molécula:**

- 3 componentes de **translação**
- 4 componentes de **rotação**: vetor unitário e 3 ângulos
- $n$  **ângulos** em  $[-180, +180]$  para as  $n$  **ligações rotacionáveis**

# 2 Avaliação

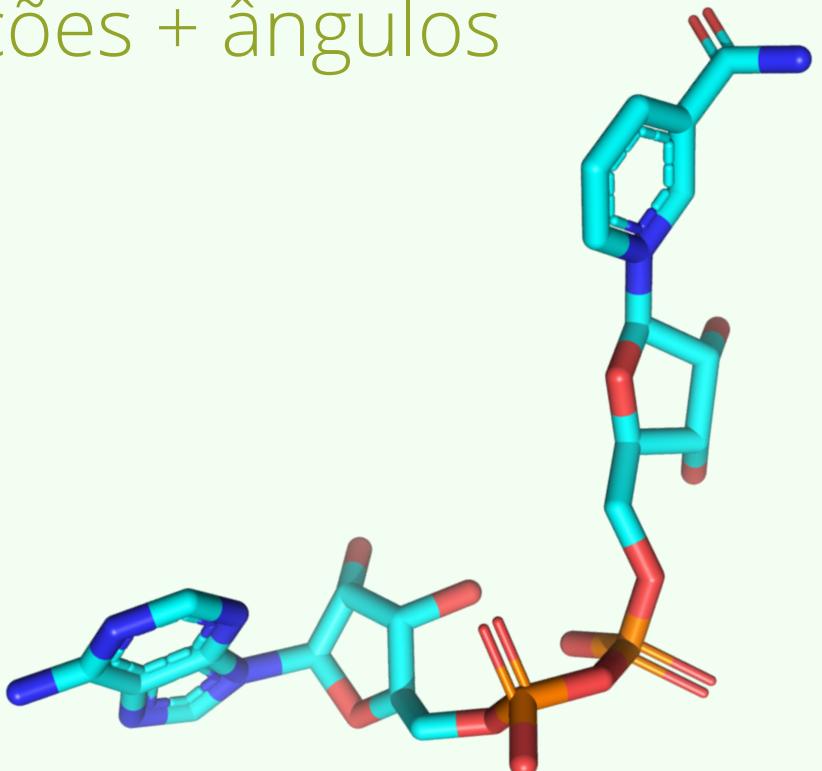
de acordo com a função objetivo

## Genótipo

Confórmeros

=

vetor de posições + ângulos



$$f(x)$$

## Fenótipo

Valor da função objetivo

=

função de energia das **interações  
intermoleculares**

$$\Delta G = \Delta G_{vdW} + \Delta G_{hbond} + \Delta G_{elec} + \Delta G_{sol} + \Delta G_{tor}$$

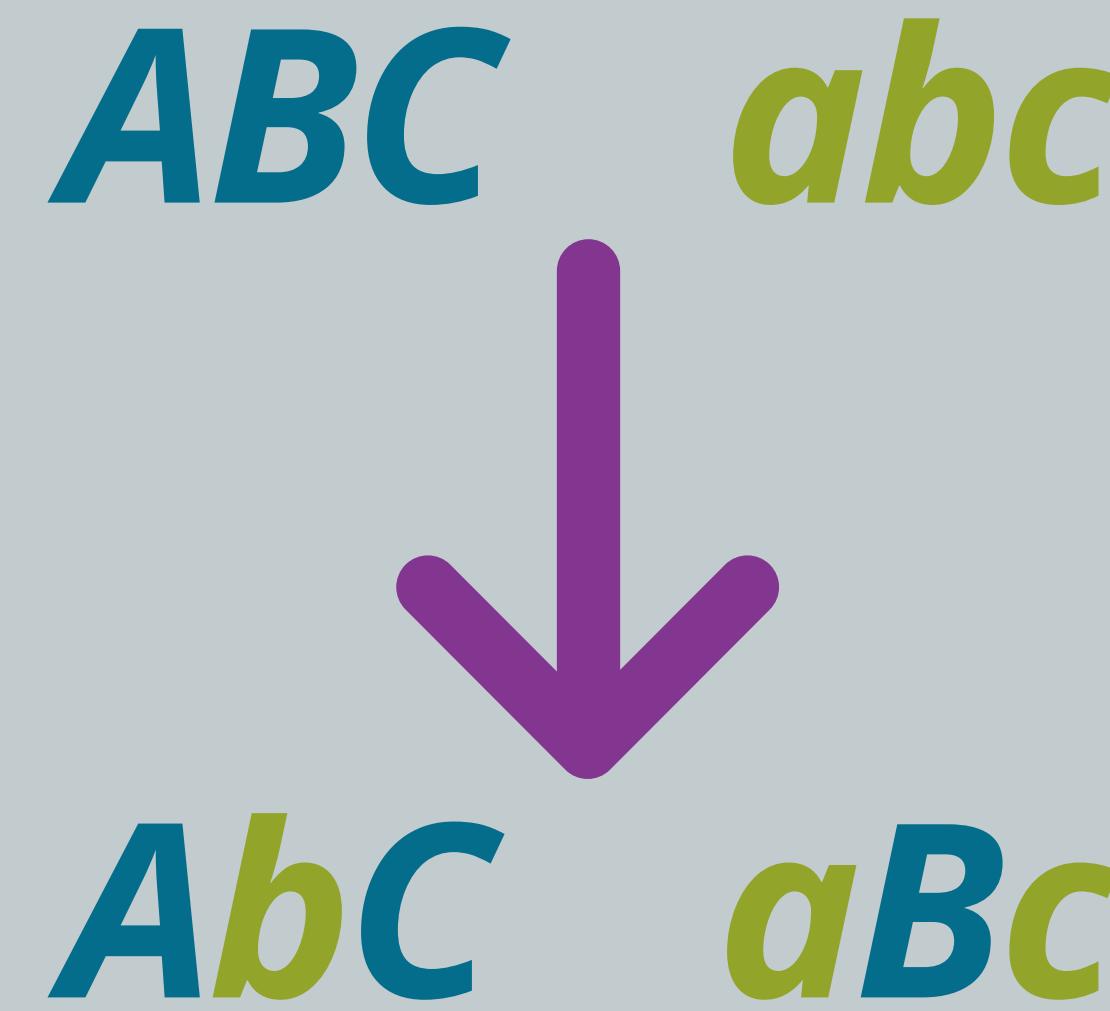
função definida e calibrada **empiricamente**



# 3 Seleção

**define os indivíduos que irão se reproduzir**

Indivíduos que possuem o valor da **função de aptidão acima da média** gerarão uma prole maior **proporcionalmente**



## 4 Cruzamento

- Cruzamento entre membros aleatórios da população
- Cruzamento
  - em **dois pontos**
  - **entre genes**
- Os **filhos substituem os pais**
  - **população de tamanho constante**

# 5 Mutação

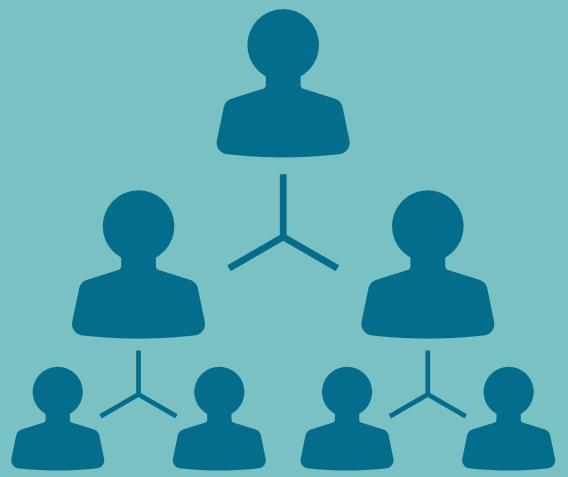


- As coordenadas de translação e rotação bem como os ângulos são **valores reais**
  - adição de um **valor aleatório real** dentro de uma distribuição pré-definida (Cauchy)

# 6 Elitismo

- Parâmetro inteiro que define quantos do indivíduos de melhor adaptados vão **automaticamente sobreviver**
- Pode ser 0 e, neste caso, todos estarão sujeitos à seleção





# Darwin

A grande maioria dos algoritmos genéticos imita as principais características da evolução darwiniana e aplicam a **genética Mendeliana**

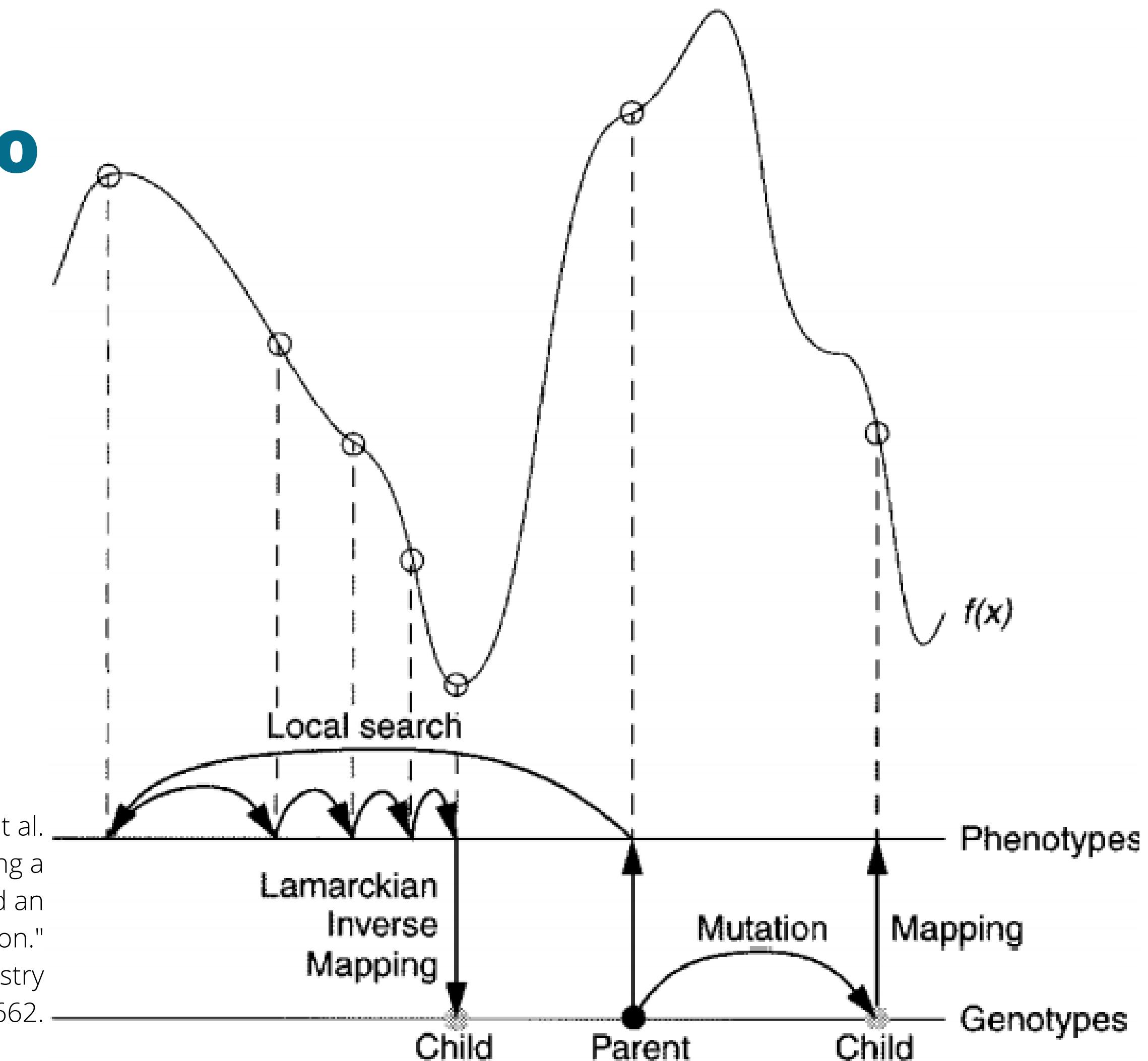


# Lamarck

No entanto, nos casos em que uma **função de mapeamento inversa** existe, por exemplo, uma que **produz um genótipo de um determinado fenótipo**, é possível fazer **buscas locais**, substituindo o indivíduo com o resultado da busca local

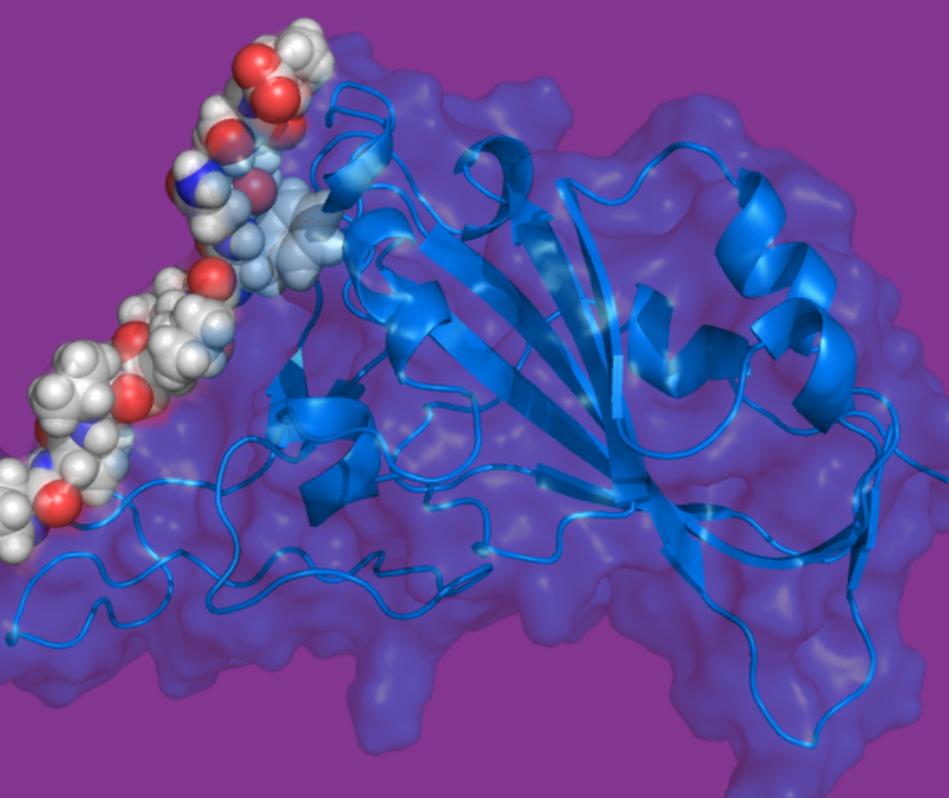
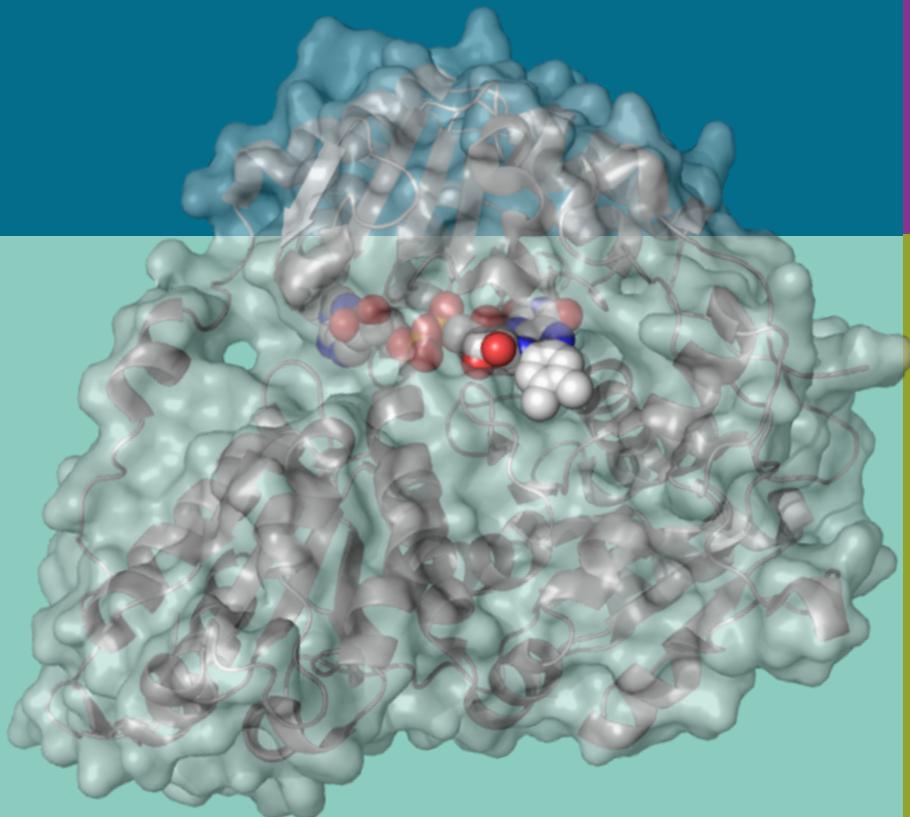
# Algoritmo genético Lamarckiano

melhora dos resultados através de buscas locais



## Exploração do espaço conformacional

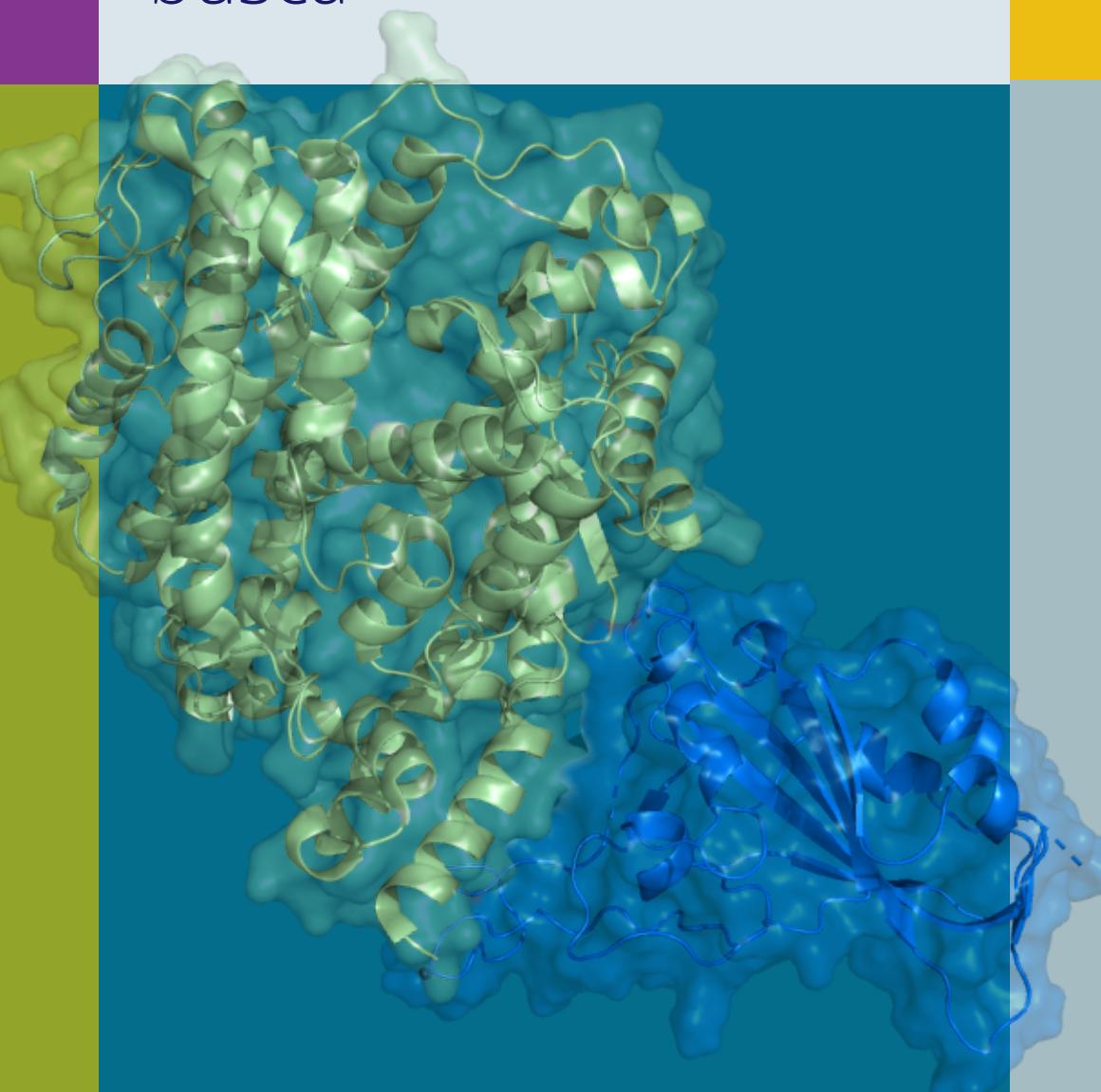
Enorme espaço de soluções



## Docking proteína-peptídeos

Peptídeos são curtas cadeias de aminoácidos

- 5 a 50 resíduos
- Tem MUITAS ligações rotacionáveis



## Docking proteína-proteína

Proteínas são macromoléculas e tratá-las como objetos flexíveis é proibitivo pelo tamanho do espaço de busca

## Desafios

### Função de aptidão (score)

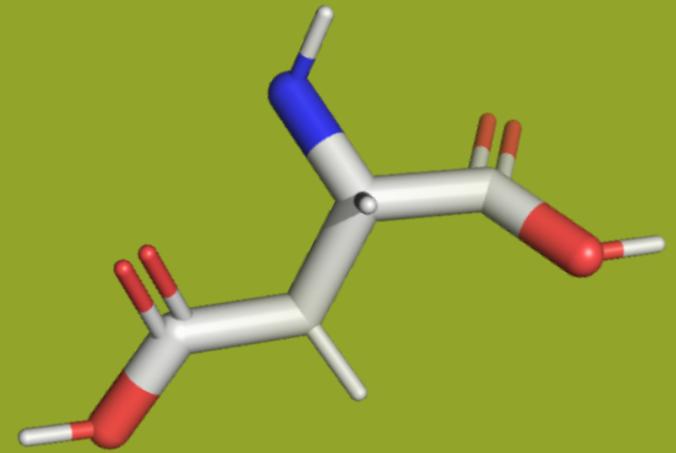
Há inúmeras funções que não são capazes de mimetizar o  $\Delta G$  até o dia de hoje

# Como avaliar?

Função de  
*score*

$\Delta G$

Avaliar a correlação entre a função de score e dados experimentais da energia de afinidade de ligação



Pose e interações

Principalmente por *redocking* que consiste em:

1. retirar o ligante experimental do sítio de ligação
2. ancorar o mesmo ligante usando o método
3. avaliar o **RMSD** do ligante experimental e o ancorado computacionalmente

**Escolha de um alvo terapêutico**

**Simulações de dinâmica molecular são realizadas por meio de computação de alto desempenho para geração de diversificadas conformações dos alvos**

**Triagem em larga-escala de bibliotecas de compostos por métodos baseados em ancoragem molecular**

Biblioteca com  
aproximadamente  
10.000 compostos  
repositionados

Biblioteca com  
aproximadamente  
1.000.000 de  
compostos

## Triagem em larga-escala de bibliotecas de compostos por métodos baseados em ancoragem molecular

Ordenação de compostos

Avaliação experimental com  
os compostos  
melhor classificados

Testes clínicos

