Des. e Prod. de medicamento - 2023

PROF. FERNANDO BARROSO
FCFRP/USP

APLICAÇÃO DE BIOINFORMÁTICA EM CIÊNCIAS FARMACÊUTICAS 1

05.10.2023

Estrutura dessa aula

- Contexto
- Motivação

1.1. Introdução a Bioinformática e seus aspectos estruturais;

1.2. Interações fundamentais e os campos de forças para simulações moleculares

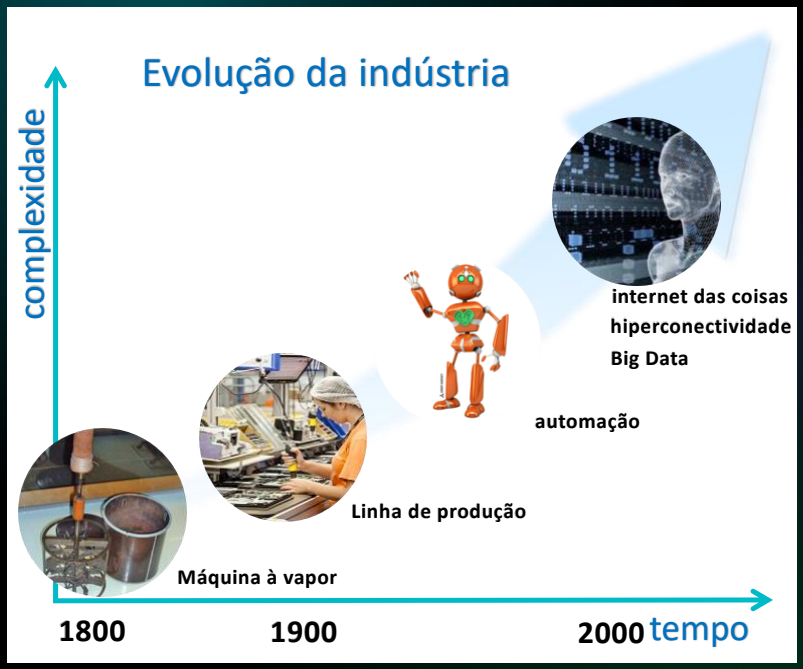
1.3. Conceitos básicos de simulações moleculares

a) Dinâmica molecular

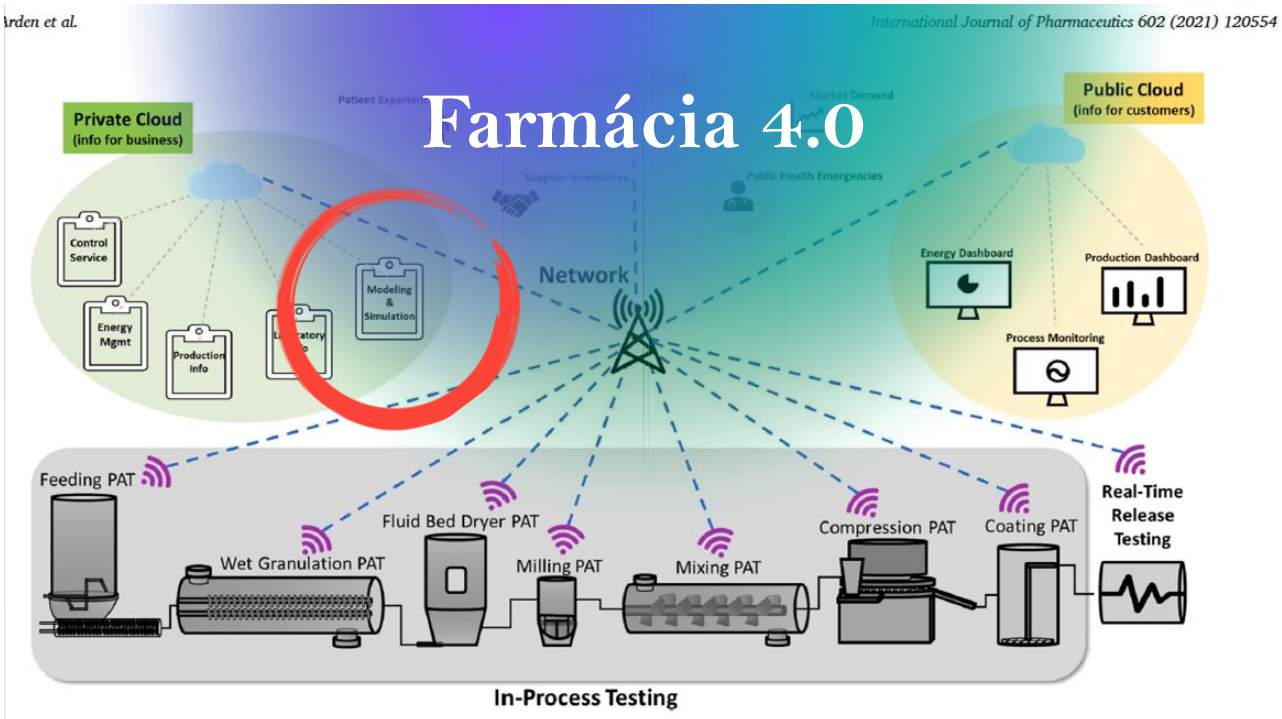
b) "Docking" molecular

1.4. Modelagem molecular e o desenvolvimento racional de fármacos

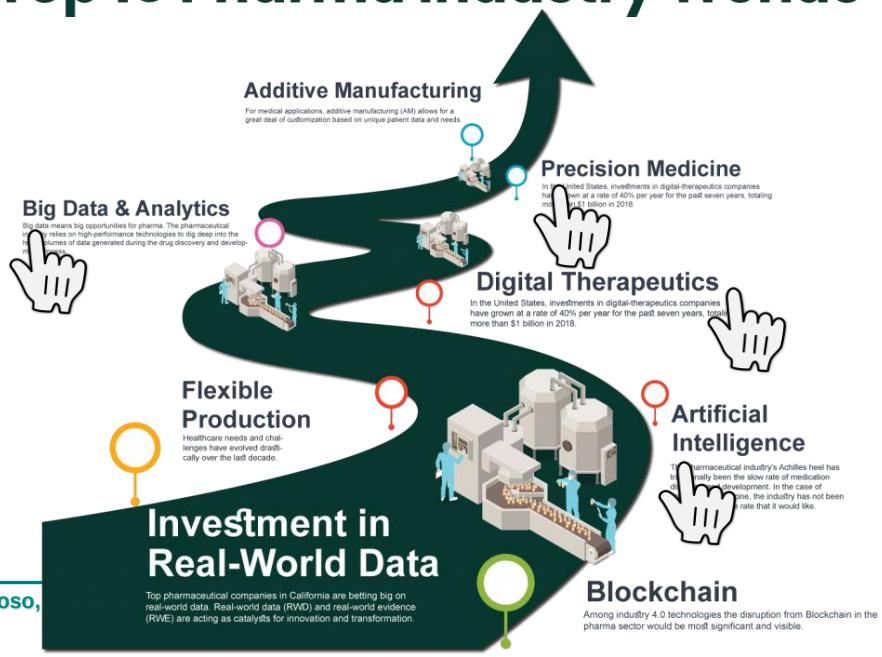
Farmácia 4.0



Farmácia 4.0



Top 10 Pharma Industry Trends



© Fernando Barroso,
flbarroso@usp.br

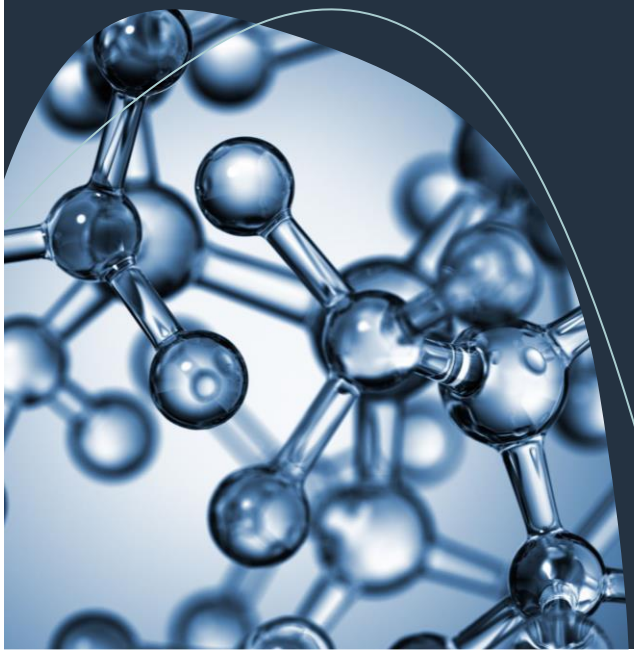
5

Global Molecular Modelling Market is Expected to Account for USD 7,244.69 Million by 2028

Global Molecular Modelling Market, By Regions, 2021 to 2028

DATA BRIDGE MARKET RESEARCH

DATA BRIDGE MARKET RESEARCH



Motivação

1. Qual o sítio de ligação?

8

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

2. Como otimizar um ligante?

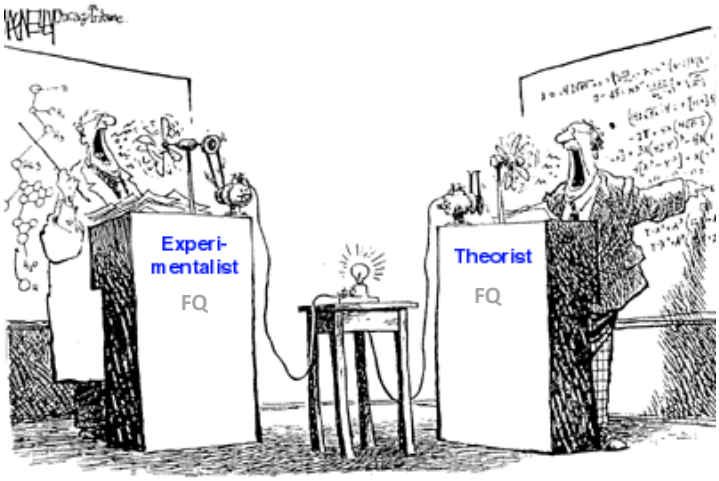
© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

3. Qual o melhor ligante?

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

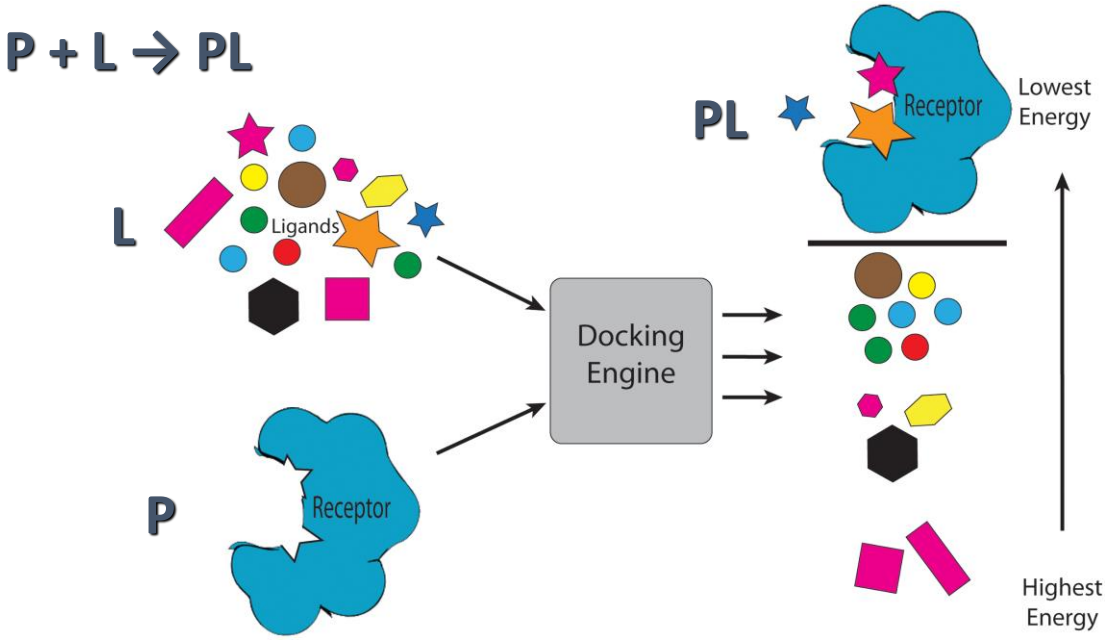
Físico-química

Físico-Química



11

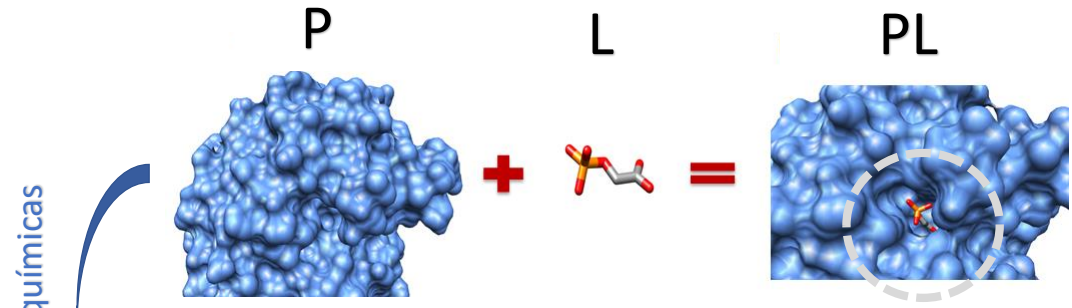
Fernando Barroso, FCFRP/USP
flbarroso@usp.br



12

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

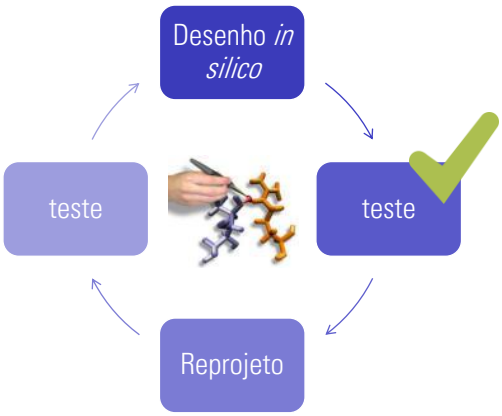
Como prever isso?

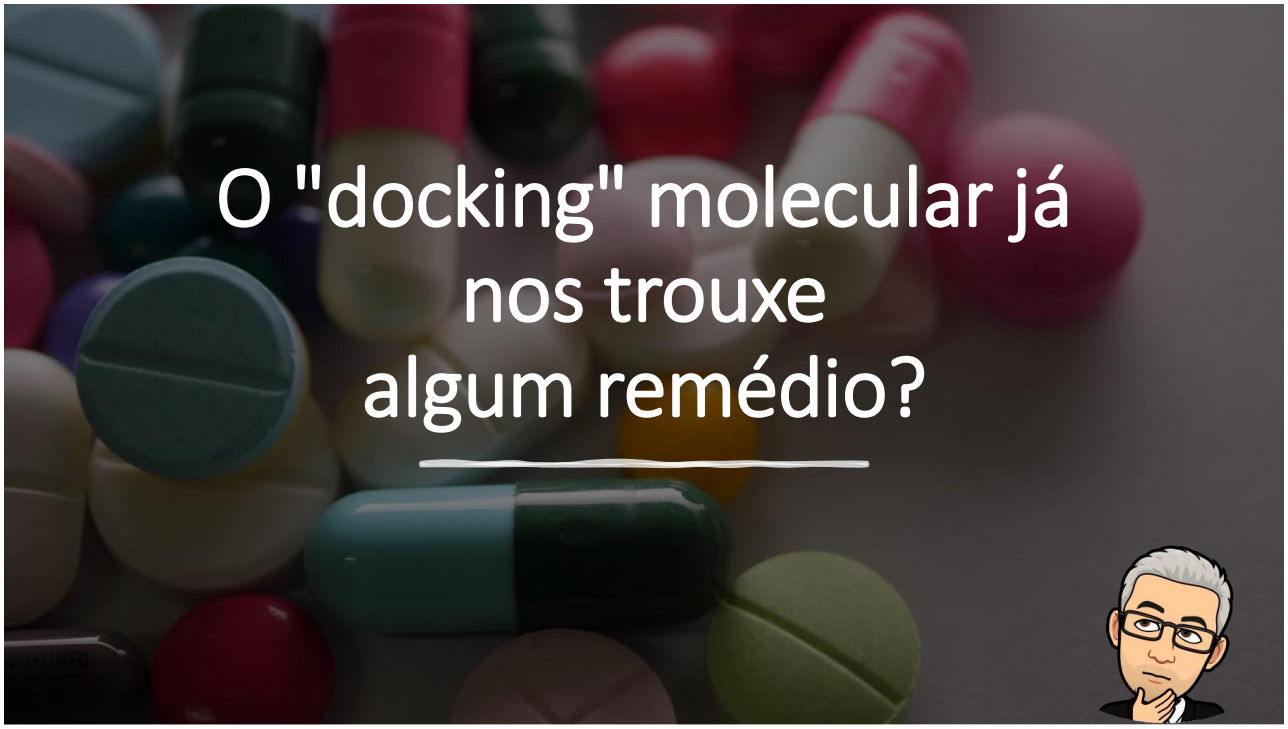


$$P + L \rightleftharpoons PL \quad \Delta G$$

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

PROTEIN/DRUG DESIGN





Has Molecular Docking Ever Brought us a Medicine?

Mark Andrew Phillips, Marisa A. Stewart, Darby L. Woodling and Zhong-Ru Xie

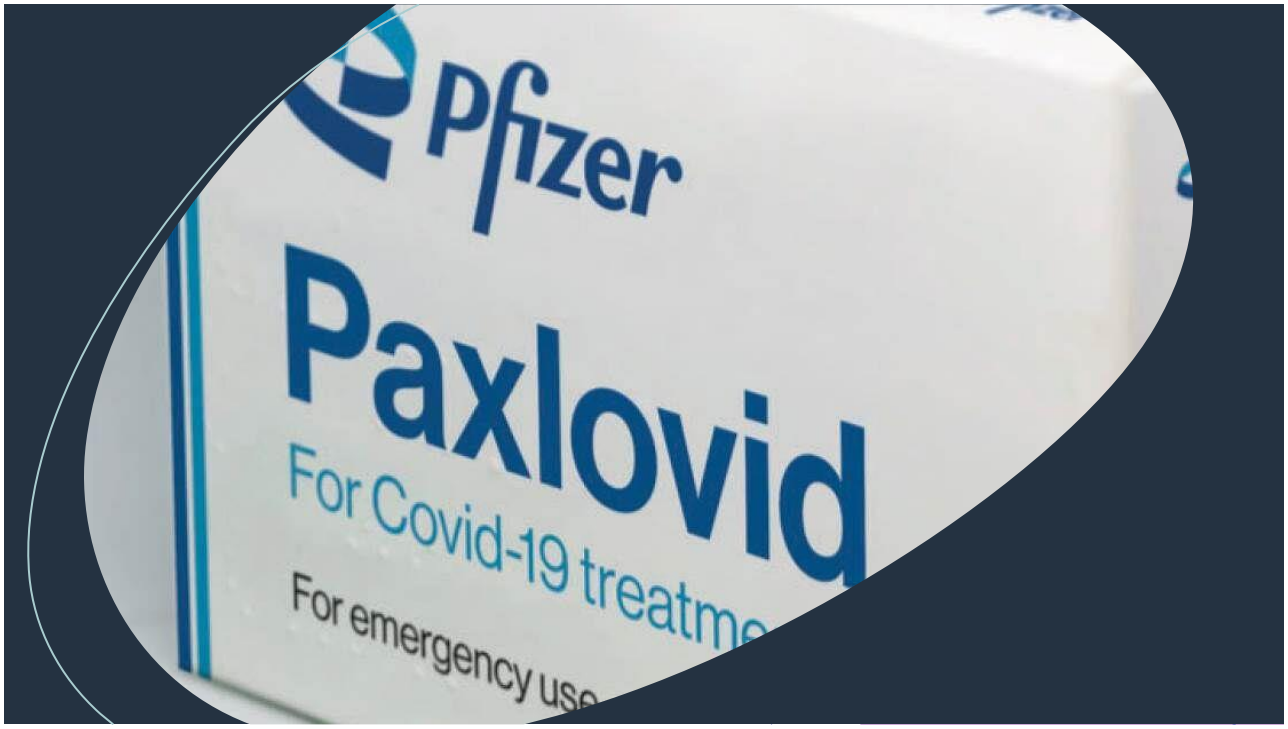
Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.72898

- ✓ **Zanamivir** (influenza)
- ✓ **Raltegravir, Nelfinavir & Saquinavar** (HIV)
- ✓ **Dorzolamide** (Glaucoma)

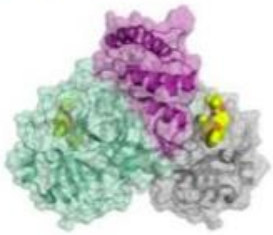
16

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br



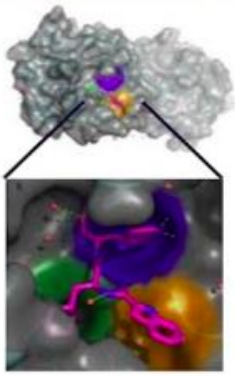
Main protease (Mpro) as a SARS-CoV-2 drug target

- Essential enzyme for processing the polyproteins translated from viral RNA
- Highly conserved across coronavirus family



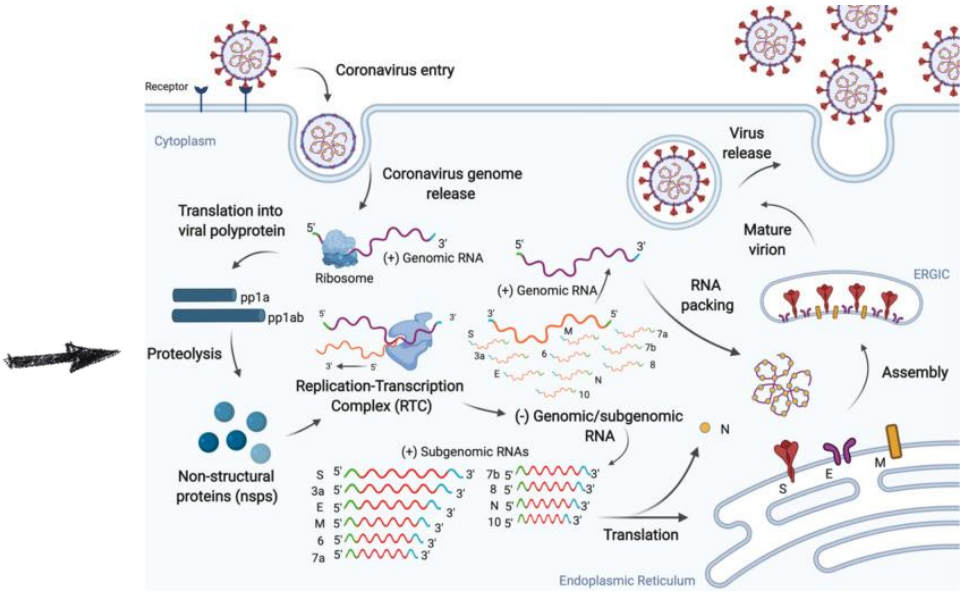
Structure of SARS-CoV-2 Mpro
The active form of Mpro is a homo-dimer with orthogonally positioned catalytic sites

- Highly structurally enabled
- >400 public PDB structures as of November 11, 2021, many with ligands bound



18

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

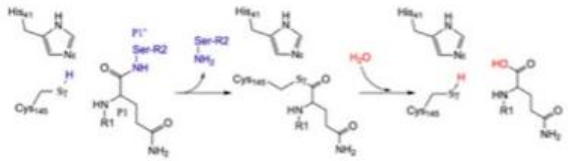


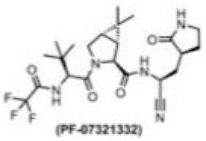
© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

10.1101/2020.12.05.409821

Main protease (Mpro) as a SARS-CoV-2 drug target

Mpro is a protease that uses a nucleophilic cysteine to cleave peptide bonds





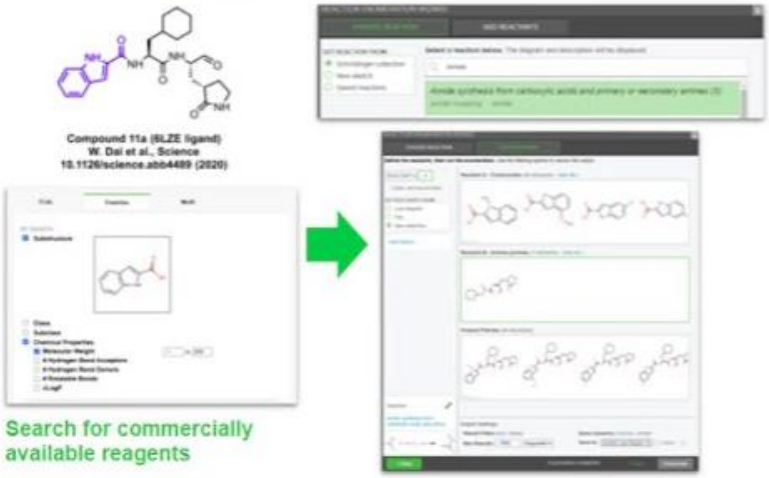
- First oral Mpro inhibitor to enter clinical development
- Peptidomimetic with cysteine modifying warhead

Schrödinger

MedRxiv 2021 DOI: 10.1101/2021.07.28.21261232

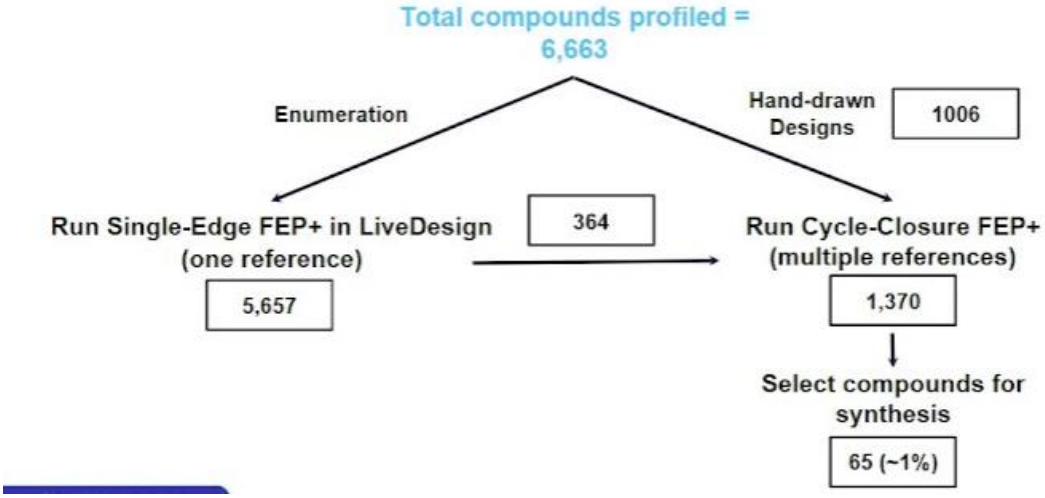
© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

Building virtual libraries using reaction enumeration in LiveDesign



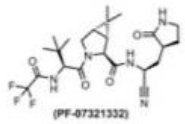
© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

FEP+: Workflow and throughput



© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

Lead compounds possess pan-coronavirus activity and similar potency to Pfizer oral Mpro inhibitor



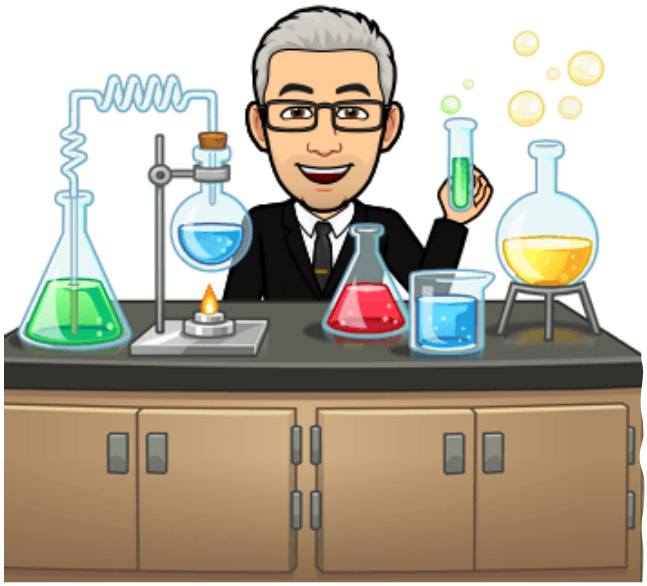
Prioritized for synthesis based on FEP+ after reaction enumeration

Designed based on WaterMap

Str Name	(PF-07321332)	Lead 1	Lead 2	Lead 3
Enzyme IC_{50} (nM)				
Mpro_SARS-CoV-2	12.6	16.2	21.6	12.1
Mpro_229E	67.8	48.4	26.4	21.8
Mpro_HKU1	17.2	23.8	16	7.74
Mpro_MERS	146	276	208	88.7
Mpro_NL63	218	180	40	37.6
Mpro_OC43	32.3	59.8	35.1	18.9
Mpro_SARS-CoV-1	21.3	27.8	30	11.9
CathepsinB(h)	>50000	38000	>50000	>50000
CPE EC_{50} (μM) / Parent Cell line CC_{50} (μM)				
SARS-CoV-2 (HeLa-ACE2)	0.059 / >40	0.11 / >40	0.12 / >40	0.097 / >40
OC43 (Huh7)	0.11 / >50	0.28 / >50	0.29 / >50	0.12 / >50
229E (MRC5)	0.10 / >50	0.11 / >50	0.06 / >50	0.07 / >50

Schrödinger

52



Como fazemos isso?

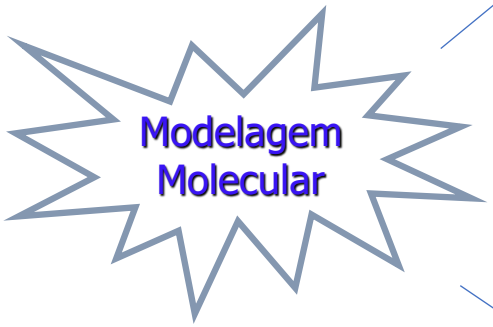
Algumas definições e conceitos

Formalismos para usarmos o docking molecular

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

Nomenclatura

docking

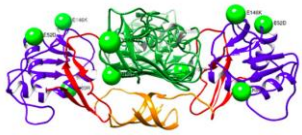


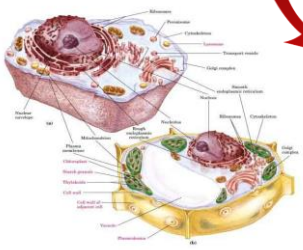
- Química Teórica
- Química Computacional
- Biologia Computacional
- Bioquímica Computacional
- Biofísica Computacional
- Bioinformática
- Biocomputação
- Simulação Computacional
- Simulação Molecular
- Experimento em computador
- Experimentos "*in silico*"
- Computer-aid.....etc.

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

Barroso Research lab

Bioinformática Estrutural





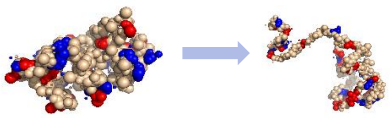
... está relacionada com o entendimento da biologia em termos de suas moléculas (no sentido da físico-química) e aplicando técnicas de "informática" para entender e organizar as informações associadas com estas moléculas, em larga-escala". (M. Gerstein)

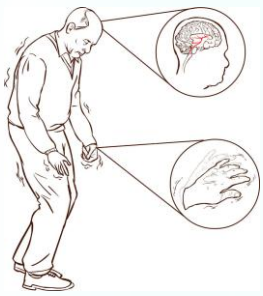
Fernando Barroso, FCFRP/USP
flbarroso@usp.br

Barroso Research lab

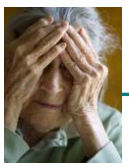
Bioinformática Estrutural

desenovelamento de proteínas





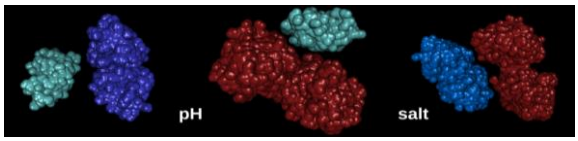
- ✓ Base molecular para o entendimento das doenças



Fernando Barroso, FCFRP/USP
flbarroso@usp.br

Barroso Research lab

Bioinformática Estrutural

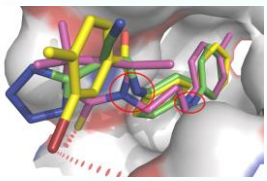


- ✓ Base molecular para o entendimento das doenças

Fernando Barroso, FCFRP/USP
flbarroso@usp.br

Barroso Research lab

Bioinformática Estrutural



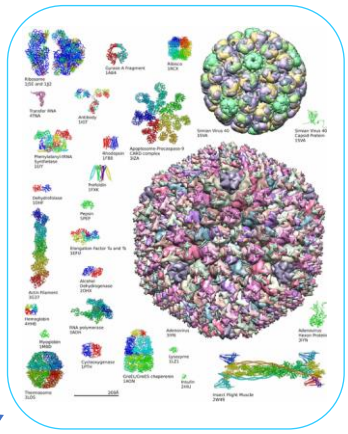
- ✓ Base molecular para o entendimento das doenças
- ✓ Engenharia de novas moléculas

Fernando Barroso, FCFRP/USP
flbarroso@usp.br

patients

Next Generation Sequencers

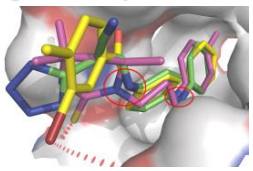
whole genome sequencing



Bioinformatics

- ✓ Rationalize processes
- ✓ Molecular design
- ✓ To deeply understand protein function
- ✓ To identify new targets
- ✓ To develop methods and algorithms to determine protein characteristics
- ✓ To analyze drug binding

in silico drug discovery



biophysical characterization

Algumas vantagens...

- Não tem risco de envenenamento, mau-cheiro ou explosão.
- Não consome reagentes (caros)
- Os experimentos podem ser feitos sob condições drásticas (alta pressão, elevada temperatura, etc...

32

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

Algumas vantagens...

- É fácil trocar a(s) molécula(s)
  .... o meio ambiente
  ..... temperatura, pH, concentrações,...

..... BASTA TROCAR ALGUNS PARÂMETROS NO *INPUT*!

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

Algumas vantagens...

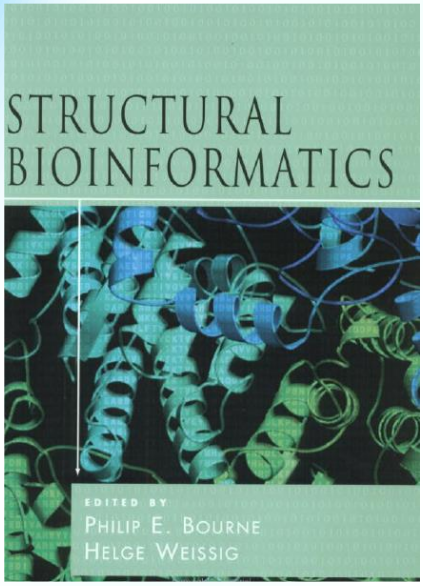
- Informações diretas no nível molecular
- Experimento livre de contaminação e impurezas

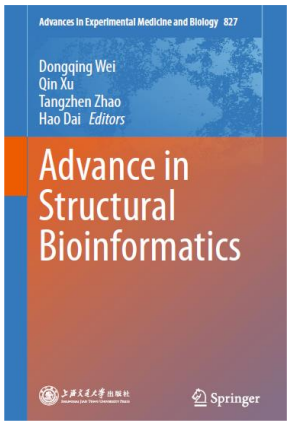
"As simulações são ficções que aspiram a imitar a realidade. Bonitas imagens e até mesmo alguns bons números não garantem uma boa ciência" (Peter Steinbach, Center of Molecular Modeling, NIH, EUA)

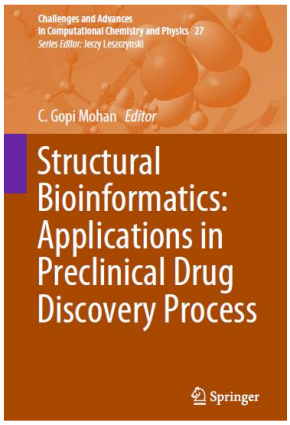
© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

Barroso Research lab

Bons livros





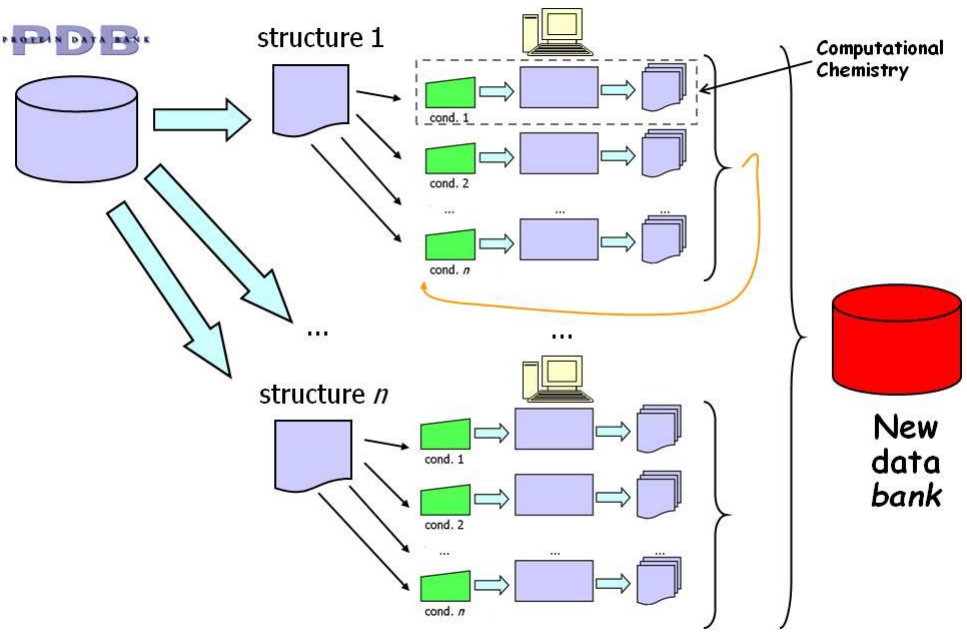


Fernando Barroso, FCFRP/USP
flbarroso@usp.br

Sequencing

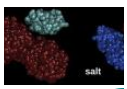
proteô
data m
alinhar
análise
genom

Est



cular
tes

ática,

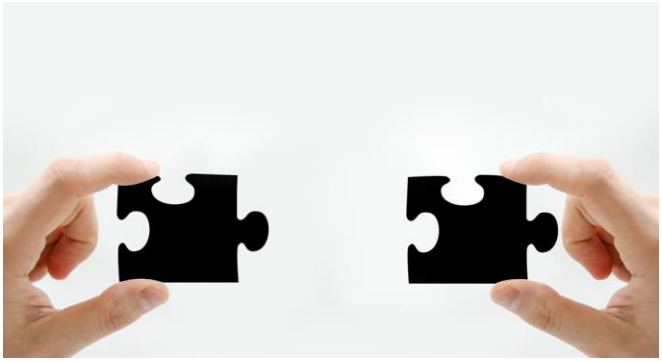


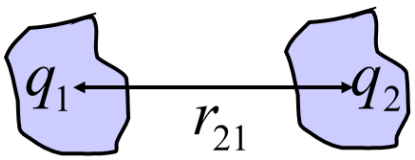
36

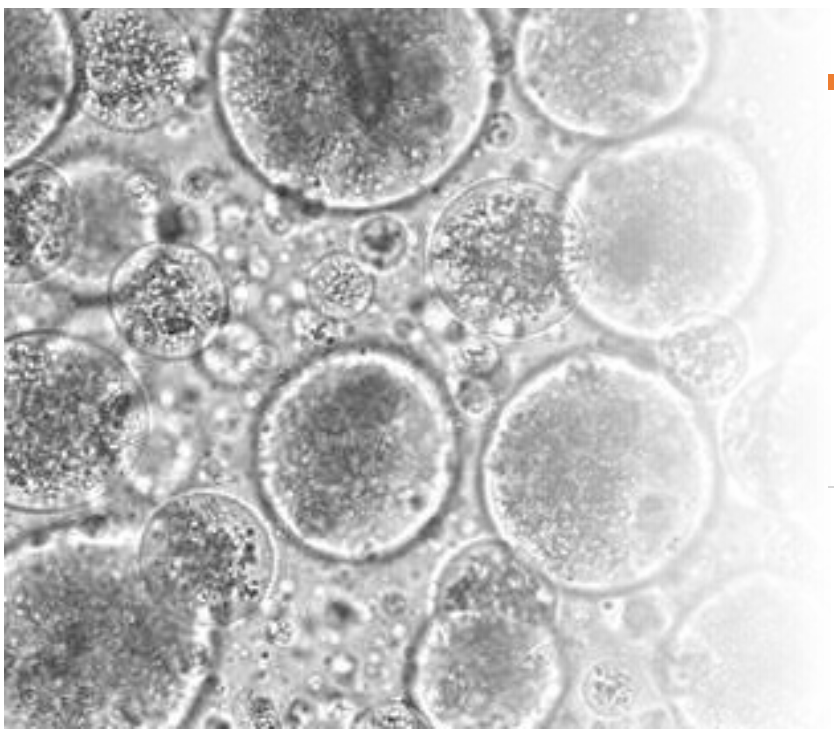
© Fernand
flbarroso

Docking

O que controla a distância entre essas moléculas?







soluto

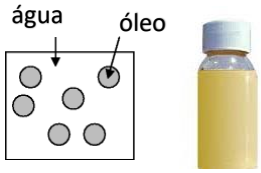
1 nm

1 µm

Coloides

são misturas heterogêneas

Sistema de dispersão farmacêutico



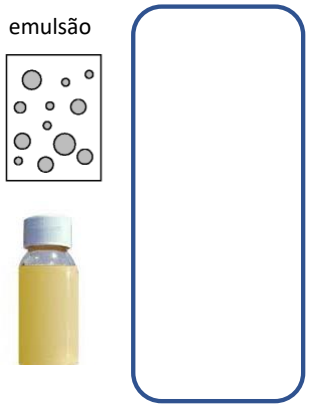
Emulsão

Uma emulsão é uma **dispersão** de pelo menos **dois líquidos imiscíveis**, um dos quais está disperso como gotículas no outro líquido, e **estabilizado por um agente emulsionante**.

É termodinamicamente instável!

Estabilidade?

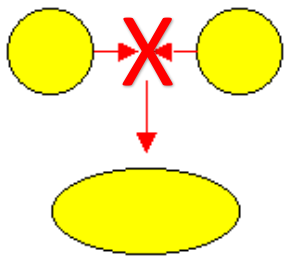
Como evitar a separação de fases?



É irreversível!
Precisa ser evitada!

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

Sistema de dispersão farmacêutico



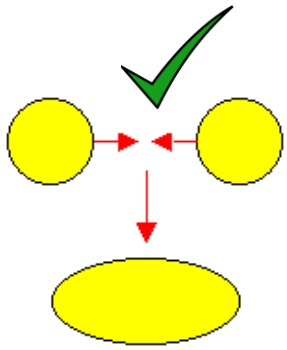
O que controla a distância entre as (macro)partículas?

- ✓ Forças intermoleculares
- ✓ Como interpretá-las?
- ✓ Como quantifica-las?

Evitar a associação!

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

Docking molecular



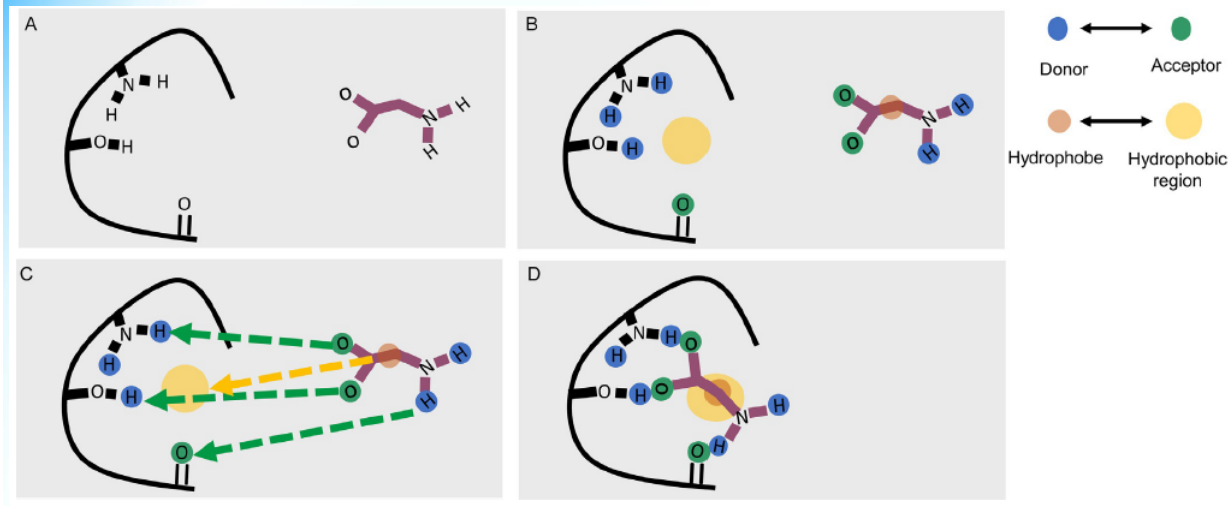
O que controla a distância entre as (macro)partículas?

- ✓ Forças intermoleculares
- ✓ Como interpretá-las?
- ✓ Como quantifica-las?

41

Garantir a associação!

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br



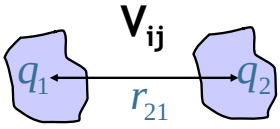
42

Fig. 1 Rigid docking approach: (A) protein and ligand. (B) The initial pose generation through pharmacophore point matching. Donor–acceptor pharmacophore point and hydrophobic pharmacophore points are added to the protein and the ligand. (C) The searching algorithm tries to match the protein and the ligand fitting points by matching donor with acceptors and hydrophobic atoms with hydrophobic cavities. (D) Different solutions are found.

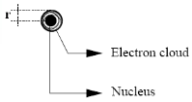
https://doi.org/10.1016/bs.pmch.2021.01.004

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

Interações intermoleculares



V_{ij} =



repulsão eletrônica

+

transferência de carga

+

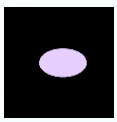
interação elétrica multipolo-multipolo

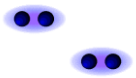
+

interação elétrica multipolo-multipolo induzido

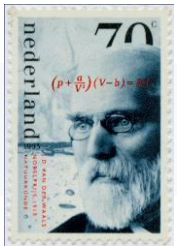
+

dispersão





Forças de van der Walls

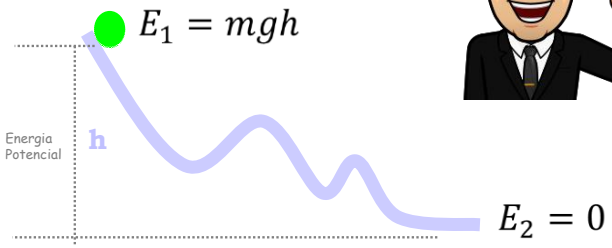


sempre atrativas

Não podem ser alteradas!

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

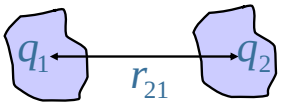
- ✓ Formalismo **Newtoniano**
- ✓ Formalismo **Hamiltoniano**



Força (partícula)

Energia (campo)

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br



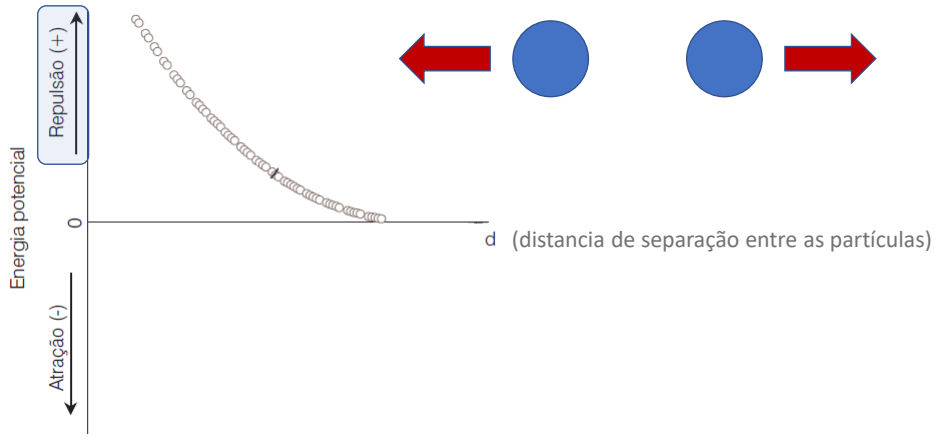
Propriedades de *campo* vs de *partícula*

Saiba as diferenciar!

	Propriedades de partícula	Inter-relações	Propriedades de campo
Grandeza vetorial	*Força* $\mathbf{F} = \frac{1}{4\pi\epsilon_0}\frac{q_1 q_2}{\lvert r\rvert^2}\hat{\mathbf{r}}$	$\mathbf{F} = q\mathbf{E}$	*Campo elétrico* $\mathbf{E} = \frac{1}{4\pi\epsilon_0}\frac{q}{\lvert r\rvert^2}\hat{\mathbf{r}}$
Inter-relações	$\mathbf{F} = -\nabla U$		$\mathbf{E} = -\nabla V$
Grandeza escalar	*Energia Potencial* $U = \frac{1}{4\pi\epsilon_0}\frac{q_1 q_2}{\lvert r\rvert}$	$U = qV$	*Potencial* $V = \frac{1}{4\pi\epsilon_0}\frac{q}{\lvert r\rvert}$

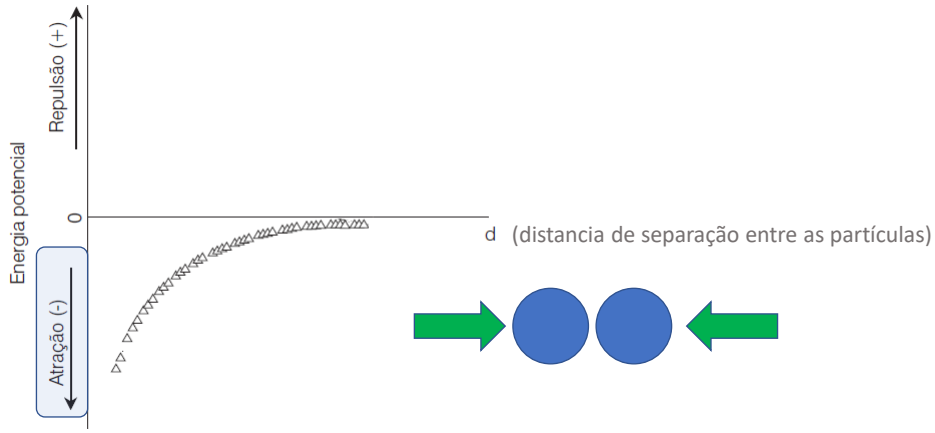
© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

Interpretação: energia potencial de interação



© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

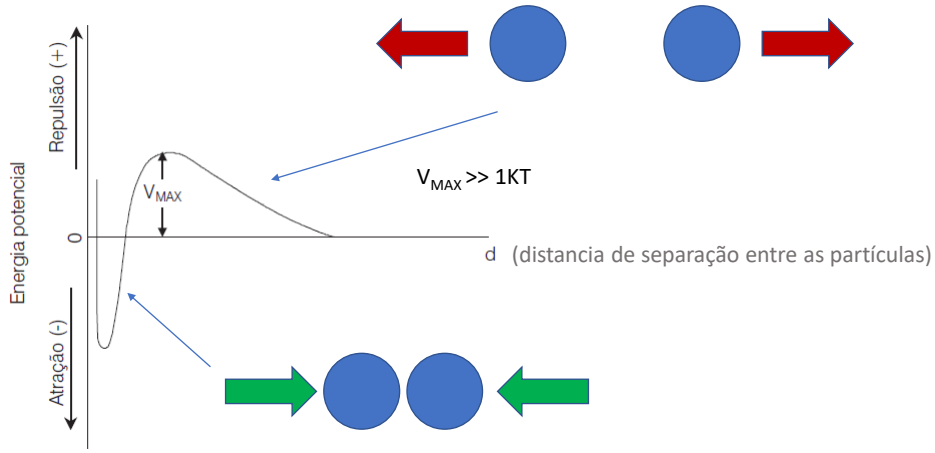
Interpretação: energia potencial de interação



47

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

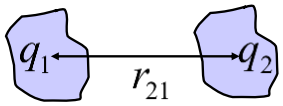
Interpretação: energia potencial de interação



48

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

Potencial *vs* Potencial de interação



$$U_t(r_{21}) = E_1 + E_2 + \frac{1}{2}\left(q_1\phi_2(r_{21}) + q_2\phi_1(r_{21})\right)$$

- ✓ Energia total de interação
- ✓ Energia configuracional
- ✓ Energia interna do sistema

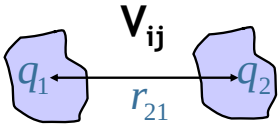
$$U_t = \frac{1}{2}\left(q_1\phi_2(r_{21}) + q_2\phi_1(r_{21})\right)$$

$$U_t = \frac{1}{2}\sum_{i=1}^{N}\sum_{i=1}^{N} \underbrace{q_i\overbrace{\phi_j(r_{ji})}^{\text{potencial}}}_{\text{"potencial" de interação}}$$

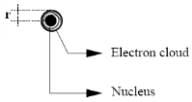
ou
energia potencial
ou
energia de interação

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

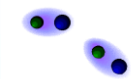
Interações intermoleculares

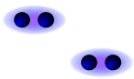


V_{ij} =

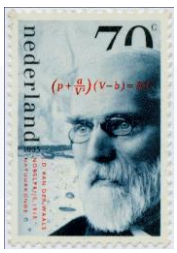


repulsão eletrônica
+
transferência de carga
+
interação elétrica multipolo-multipolo
+
interação elétrica multipolo-multipolo induzido
+
dispersão





Forças de van der Walls



sempre atrativas

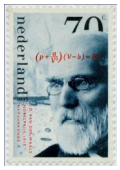
Não podem ser alteradas!

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

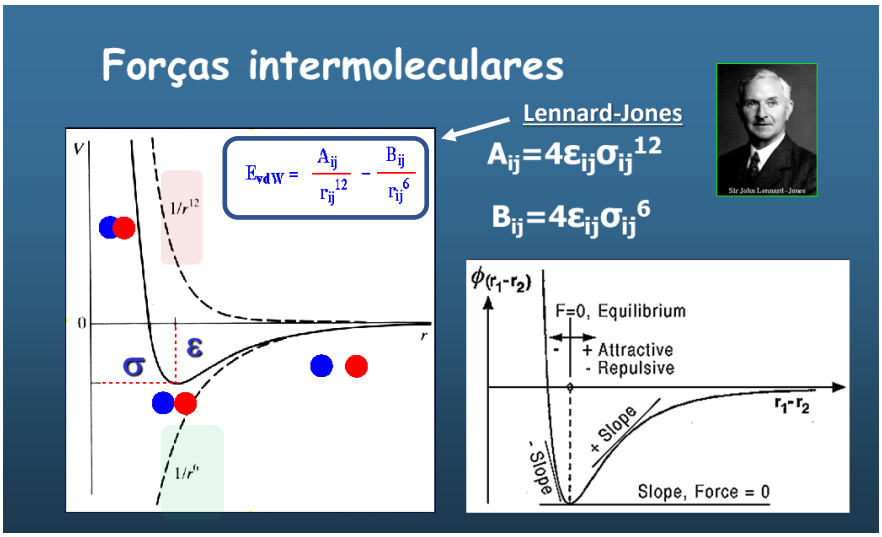
Como quantifica-las?

$$u_{vdW}(r) = \frac{A}{r_{ij}^{12}} - \frac{B}{r_{ij}^{6}}$$

Para átomos

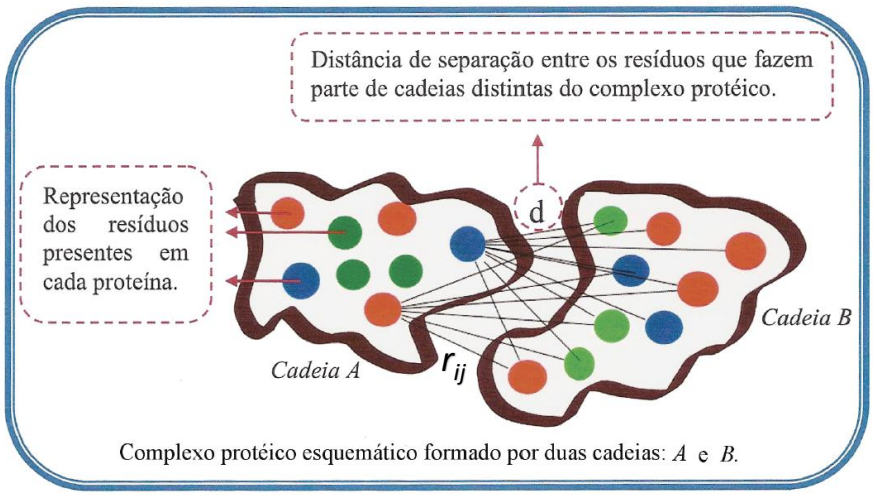


Forças de van der Walls



© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

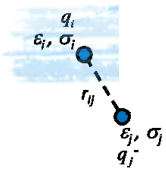
$$u_{vdW}(r) = \frac{A}{r_{ij}^{12}} - \frac{B}{r_{ij}^{6}}$$



Complexo protéico esquemático formado por duas cadeias: *A* e *B*.

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

Campo de força

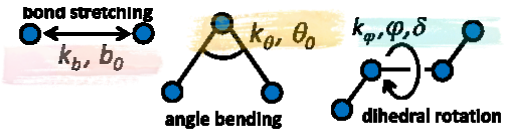


THE JOURNAL OF CHEMICAL PHYSICS VOLUME 49, NUMBER 11 1 DECEMBER 1968

Consistent Force Field for Calculations of Conformations, Vibrational Spectra, and Enthalpies of Cycloalkane and *n*-Alkane Molecules

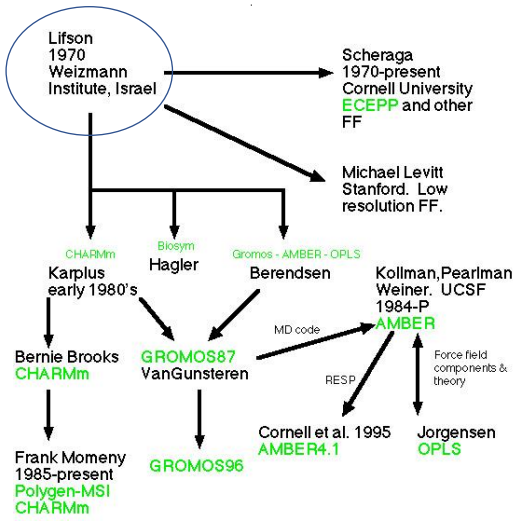
S. LIFSON AND A. WARSHEL
Department of Chemical Physics, Weizmann Institute of Science, Rehovot, Israel

(Received 13 May 1968)



DOI:10.1016/j.colsurfb.2014.06.050

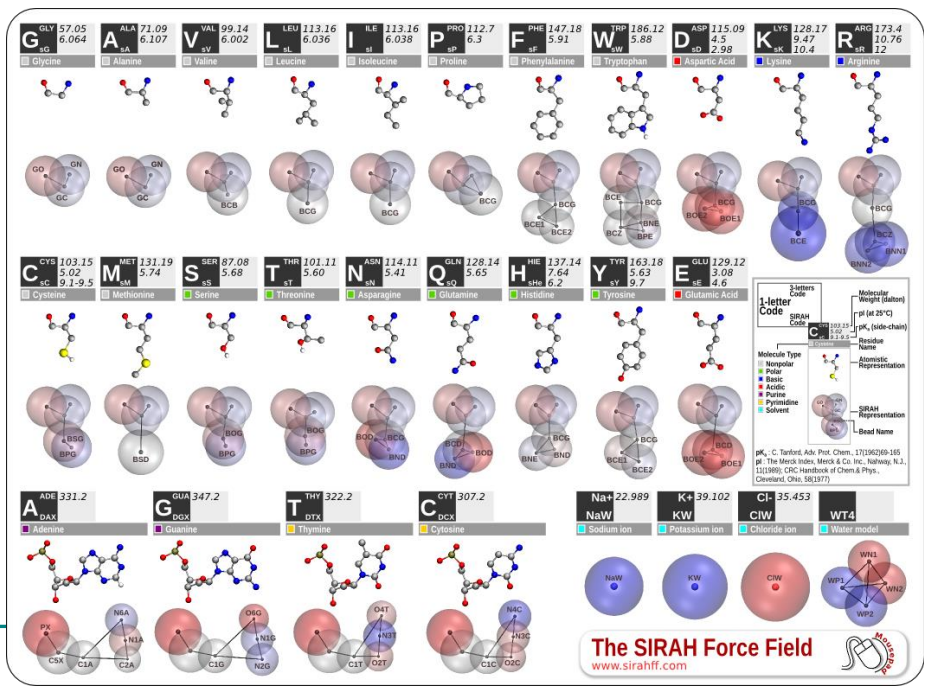
$$U = \frac{1}{2}\sum_{\text{bonds}} k_b(b-b_0)^2 + \frac{1}{2}\sum_{\text{angles}} k_\theta(\theta-\theta_0)^2 + \frac{1}{2}\sum_{\text{dihedrals}} k_\varphi(1+\cos(n\varphi-\delta)) + \sum_{\substack{\text{non-bonded}\\ \text{pairs}}} \left[\frac{A}{r^{12}} - \frac{B}{r^6}\right] + \frac{q_1 q_2}{\varepsilon\varepsilon_0 r}$$



© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

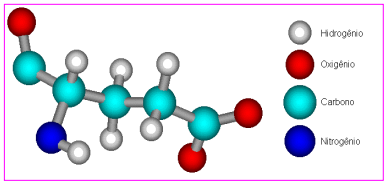
Institut Pasteur

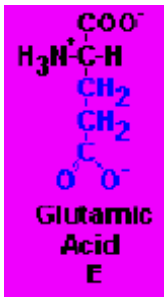
SIRAH
Southamerican Initiative for a Rapid and Accurate Hamiltonian



© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

Campos de força

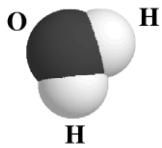




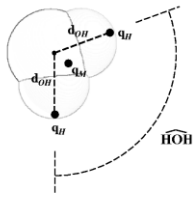
Atom	Gromos	Parse	Charmm	Reduced
N	-0.328	-0.400	-0.470	0,000
CA	0.050	0.000	0.070	0,000
C	0.355	0.550	0.510	0,000
O	-0.380	-0.550	-0.510	0,000
CB	0.010	0.000	-0.180	0,000
CG	-0.155	0.000	-0.280	0,000
CD	0.365	0.100	0.620	0,000
OE1	-0.575	-0.550	-0.760	-0.500
OE2	-0.575	-0.550	-0.760	-0.500
H	0.233	0.400	0.310	0,000
HÁ	0.000	0.000	0.090	0,000
HB1	0.000	0.000	0.090	0,000
HB2	0.000	0.000	0.090	0,000
HG1	0.000	0.000	0.090	0,000
HG2	0.000	0.000	0.090	0,000

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

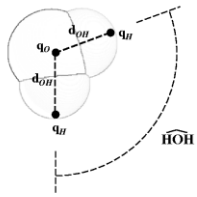
Campos de força



TIP4P



SPC



Model	q_O e	q_M e	q_H e	d_{OH} Å	$\widehat{HOH}$ deg	C_6(O,O) $(kJmol^{-1}nm^6)$	C_{12}(O,O) $(kJmol^{-1}nm^{12})$
TIP3P	-0.834		0.417	0.9572	104.52	2.4889×10^{-3}	2.4352×10^{-6}
TIP4P		-1.04*	0.52	0.9572	104.52	2.5543×10^{-3}	2.5145×10^{-6}
SPC	-0.82		0.41	1.0	109.47	2.6171×10^{-3}	2.6331×10^{-6}
SPC/E	-0.8476		0.4238	1.0	109.47	2.6171×10^{-3}	2.6331×10^{-6}

Table 5.1: Popular water model parameters.
* This single negative charge is located along the $\widehat{HOH}$ bisector with $d_{OM} = 0.15$ Å.

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

Campos de força

TABLE 2: Water Model Properties at 298 K, 1 atm[a]

	r_c (nm)	LRC	μ (D)	ϵ_r	ρ (kg m^{-3})	α_P (10^{-4} K^{-1})	$\partial\alpha_P/\partial T$ (10^{-6} K^{-2})	κ_T (GPa^{-1})	$k_BT^2\alpha_P$ (kJ mol^{-1})
SPC	1.40	no	2.27	66	973.8	7.4	6.2	0.53	0.55
SPC/E	1.40	no	2.35	71	995.5	5.0	7.5	0.47	0.37
SPC/E	0.85	yes	2.35	71	998.9	5.0	6.8	0.46	0.37
TIP3P	0.85	yes	2.35	98	986.0	8.9	5.3	0.57	0.66
TIP4P	0.85	yes	2.18	52	994.1	5.6	8.9	0.51	0.41
TIP4P-Ew	0.85	yes	2.32	65	995.7	3.2	10.5	0.47	0.23
experiment				78	997.2	2.56	9.7	0.468	0.19

[a] The dipole moment μ, dielectric constant ϵ_r, density ρ, thermal expansion coefficient α_P and its temperature derivative, isothermal compressibility κ_T, and the liberation correction in the hydration enthalpy are given. r_c is the LJ cutoff radius. LRC is the long-range dispersion correction. All quantities, except for κ_T, were determined from 800 ns *NPT* Simulations: ϵ_r from dipole fluctuations,[49] temperature derivatives at 298 K from finite differences using the temperatures 278, 298, and 323 K. κ_T was determined from a finite difference using 10 ns simulations with ±1% density differences. The experimental data were taken from Wagner et al.[47]

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

Eur Biophys J (2005) 34: 273–284
DOI 10.1007/s00249-004-0448-6

ARTICLE

Chris Oostenbrink · Thereza A. Soares
Nico F. A. van der Vegt · Wilfred F. van Gunsteren

Validation of the 53A6 GROMOS force field

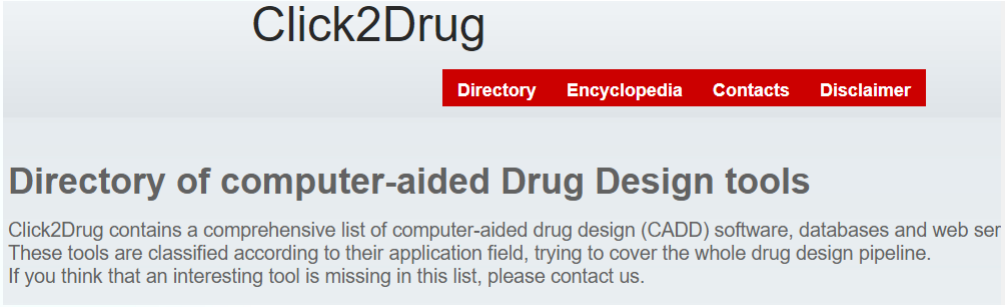
Table 3 Comparison of structural parameters obtained from experimental and simulation structures of the DNA dodecamer

	45A4	53A6	NMR	X-ray
Local base-pair parameters				
Shear (nm)	−0.002	0.007	0.0	−0.018
Stretch (nm)	−0.017	−0.017	−0.039	−0.015
Stagger (nm)	−0.035	−0.033	−0.022	−0.002
Buckle (°)	0.0	−0.6	−0.1	2.5
Propeller (°)	−7.1	−7.1	−12.3	−19.0
Opening (°)	−2.0	−2.2	2.1	3.0
Sugar-ring puckering				
C3′-endo (%)	5	4	0	4
C4′-exo (%)	4	4	0	0
O4′-endo (%)	20	19	0	12
C1′-exo (%)	44	44	63	42
C2′-endo (%)	23	25	37	42
C3′-exo (%)	2	2	0	0
C2′-exo (%)	1	1	0	0
Pseudorotation				
Phase (°)	118.8	121.1	132.1	133.7
Amplitude (°)	42.5	42.4	29.3	41.3
Local base-pair helical parameters				
X-displacement (nm)	−0.349	−0.321	−0.128	−0.052
Y-displacement (nm)	0.004	−0.009	0.001	−0.022
Rise (nm)	0.278	0.300	0.329	0.329
Inclination (°)	15.5	14.2	4.6	5.0
Tip (°)	−0.2	−0.1	0.0	1.3
Helical twist (°)	33.4	33.3	35.1	37.8
Local base-pair step parameters				
Shift (nm)	−0.002	0.003	0.0	0.002
Slide (nm)	−0.109	−0.092	−0.046	0.004
Rise (nm)	0.342	0.357	0.338	0.341
Tilt (°)	0.1	0.1	0.0	−0.9
Roll (°)	8.4	7.8	2.5	3.4
Twist (°)	30.4	30.5	34.6	36.1

Averages are over all bases, sugars, base pairs and inter-base-pair parameters observed in 4 ns of simulation (parameter sets 45A3 and 53A6), five NMR model structures (Tjandra et al. 2000) or the X-ray structure (Shui et al. 1998). Local base-pair parameters, helical parameters and step parameters were calculated using the three DNA definitions (Lu and Olson 2003; Olson et al. 2001). Sugar puckering, pseudorotation phase and pucker amplitudes are according to Altona et al. (1968) and Altona and Sundaralingam (1972)

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

Ferramentas disponíveis



https://www.click2drug.org/index.php

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

Mais populares

AutoDock

https://autodock.scripps.edu/

AutoDock Vina

https://vina.scripps.edu/

https://dock.compbio.ucsf.edu/DOCK_6/index.htm

GOLD

https://www.ccdc.cam.ac.uk/solutions/csd-discovery/components/gold/

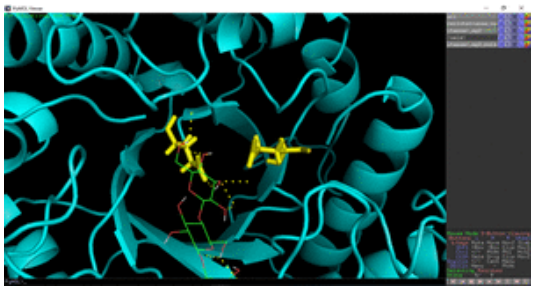
https://www.rosettacommons.org/software/servers

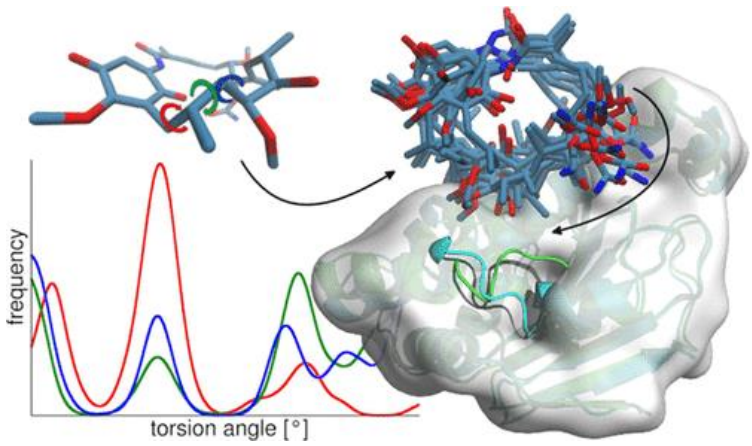
© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

Exemplo de resultados

AutoDock Vina

mode	affinity (kcal/mol)	dist from best mode rmsd l.b.	rmsd u.b.
1	-7.5	0.000	0.000
2	-7.3	1.764	3.961
3	-7.2	3.178	10.371
4	-6.9	1.668	2.928
5	-6.7	3.877	7.199
6	-6.6	1.571	9.263
7	-6.6	3.307	10.436
8	-6.5	2.670	5.733
9	-6.4	1.168	2.462





Fernando Barroso, FCFRP/USP

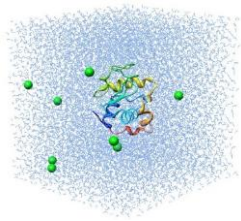
Diferentes opções

Laboratory of Computational Biophysical Chemistry

1) Modelo (rígido ou flexível)

2) "Scoring function"

$$f(x)$$



Modelagem → **Solução do modelo**

Qual campo de força?

simplificação
redução # sítios

precisão
relevância prática

Dinâmica molecular?
Dinâmica de Langevin?
Algoritmo genético?
Monte Carlo?
Poisson-Boltzmann?

Table 1. Molecular docking software.

Software	Posing	Scoring	Availability	Reference
Vina	Iterated Local Search + BFGS Local Optimiser	Empirical/ Knowledge-Based	Free (Apache License)	Trott, 2010 [3]
AutoDock4	Lamarckian Genetic Algorithm, Genetic Algorithm or Simulated Annealing	Semiempirical	Free (GNU License)	Morris, 2009; Huey, 2007 [31,32]
Molegro/MolDock	Differential Evolution (Alternatively Simplex Evolution and Iterated Simplex)	Semiempirical	Commercial	Thomsen, 2006 [9]
Smina	Monte Carlo stochastic sampling + local optimisation	Empirical (customisable)	Free (GNU License)	Koes, 2013 [33]
Plants	Ant Colony Optimisation	Empirical	Academic License	Korb, 2007; Korb, 2009 [34,35]
ICM	Biased Probability Monte Carlo + Local Optimisation	Physics-Based	Commercial	Abagyan, 1993; Abagyan, 1994 [36,37]
Glide	Systematic search + Optimisation (XP mode also uses anchor-and-grow)	Empirical	Commercial	Friesner, 2004 [38]
Surflex	Fragmentation and alignment to idealised molecule (Protomol) + BFGS optimisation	Empirical	Commercial	Jain, 2003; Jain 2007 [39,40]
		Physics-based (GoldScore), Empirical		Jones, 1997;

Laboratory of Computational ysical Chemistry

https://doi.org/10.3390/ijms20184574

Barroso, FCFRP/USP

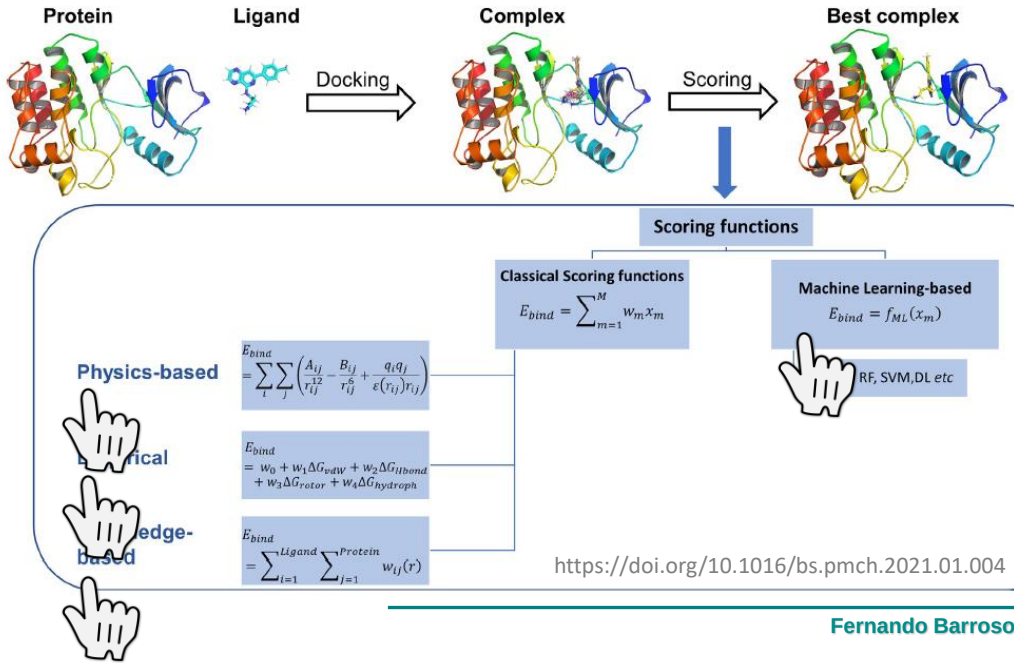
Docking

Como escolher o melhor método?

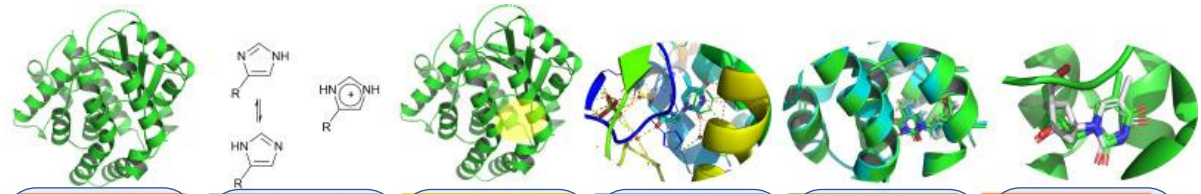
		Correlation		Entire set		Small		Large		Hydrophilic		Hydrophobic		Proteins	
		Pearson	Spearman	RMSD	% Pairs	RMSD	% Pairs	RMSD	% Pairs	RMSD	% Pairs	RMSD	% Pairs	RMSD	% Pairs
Top score															
Best docking	Program	eHits	eHits	GOLD	GOLD	GOLD	GOLD	GOLD	GOLD	GOLD	GOLD	eHits	eHits	GOLD	GOLD
program	Results	0.38	0.29	2.68	58.45	1.96	67.11	3.50	48.45	2.30	65.67	2.91	48.33	4.03	46.50
Second best	Program	Surflex	Surflex	eHits	eHits	eHits	eHits	eHits	eHits	eHits	eHits	GOLD	GOLD	AutoDock	AutoDock
program	Results	0.33	0.22	2.76	54.02	1.96	64.80	3.59	37.71	2.62	50.96	3.06	50.28	4.06	46.86
Averaged results*	Results	0.23	0.16	3.63	45.13	2.50	56.69	4.53	34.84	3.26	49.19	4.02	42.10	5.45	36.80
MetaPose	Results	0.42	0.43	2.57	62.67	1.75	70.60	3.11	52.25	2.08	67.31	2.63	57.90	3.59	50.00
MetaScore	Results	0.48	0.47	3.43	52.70	2.17	62.47	3.70	42.81	2.67	56.65	3.19	48.46	4.36	28.78
VoteDock	Results	0.49	0.50	2.20	68.67	1.58	76.18	2.82	59.04	1.90	72.49	2.28	62.70	3.26	57.00
Best pose															
Best docking	Program	eHits	eHits	GOLD	GOLD	eHits	eHits	GOLD	GOLD	GOLD	GOLD	eHits	eHits	GOLD	GOLD
program	Results	0.29	0.38	1.66	73.83	1.20	83.30	2.01	66.38	1.43	79.05	1.74	72.42	2.63	61.77
Second best	Program	Surflex	Surflex	eHits	eHits	GOLD	GOLD	Surflex	Surflex	eHits	eHits	Surflex	Surflex	LigandFit	LigandFit
program	Results	0.22	0.30	1.69	73.80	1.31	81.07	2.42	64.82	1.64	72.30	2.04	69.04	2.71	49.87
Averaged results*	Results	0.16	0.19	2.22	65.63	1.54	76.92	2.99	53.54	2.02	68.46	2.52	62.44	3.64	48.60
MetaPose	Results	0.30	0.34	1.54	82.53	0.97	87.83	1.86	74.62	1.29	84.64	1.59	77.89	1.86	74.62
MetaScore	Results	0.39	0.44	1.29	82.12	0.93	89.91	1.65	72.35	1.17	84.52	1.41	77.73	1.94	65.84
VoteDock	Results	0.40	0.45	1.64	78.91	1.35	83.75	2.09	70.19	1.46	80.01	1.83	74.05	2.34	65.17

Consenso!

10.1002/jcc.21642



https://doi.org/10.1016/bs.pmch.2021.01.004

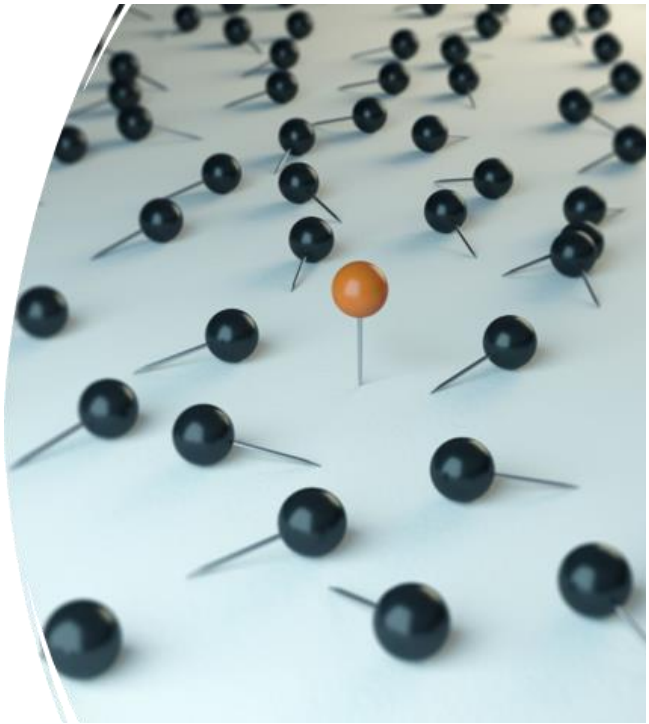


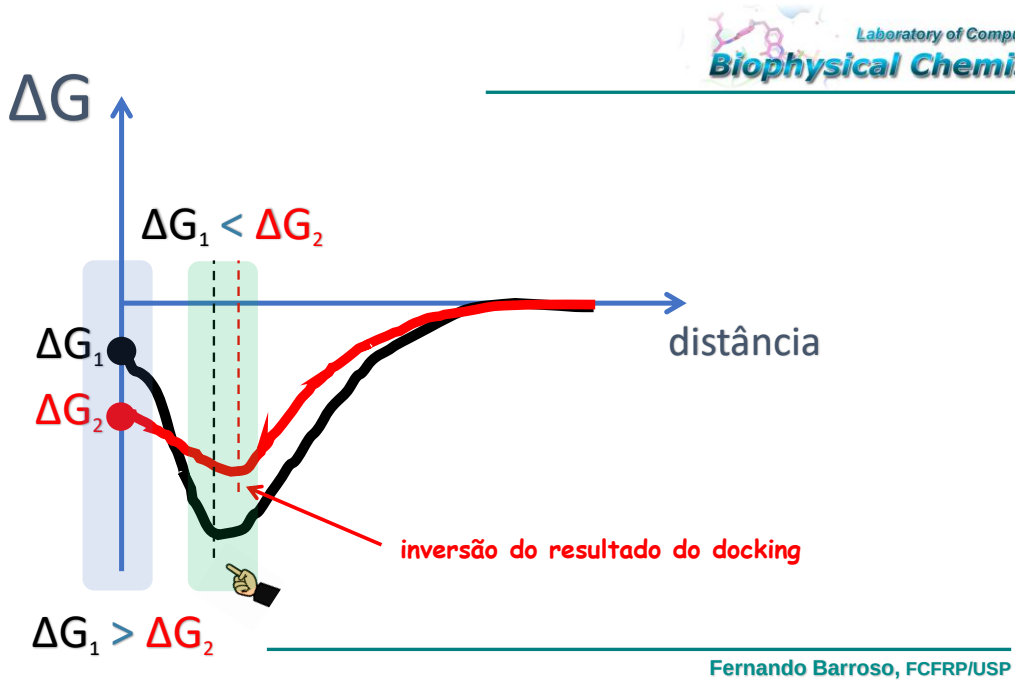
Protein and ligand selection	Protein ligand preparation	Binding site definition	Structural water	Molecular docking	Evaluation
• Structure of the protein target (3D coordinates) • 3D coordinates of ligand(s)	• Add hydrogens • Check protonation state(s) • Select proper tautomer/protomer /stereoisomer	• Experimentally (3D structure of protein-ligand complex is available) • Cavity detection	• Waters that coordinate H-bonds between protein and ligand	• Exploring the conformational space (searching algorithm) • Ranking candidate solutions (scoring function)	• Are all the H-bond donors and acceptors in the ligand satisfied? • If the complex is known, is the binding mode of a ligand to protein reproduced?

https://doi.org/10.1016/bs.pmch.2021.01.004

Docking

Outro importante detalhe...



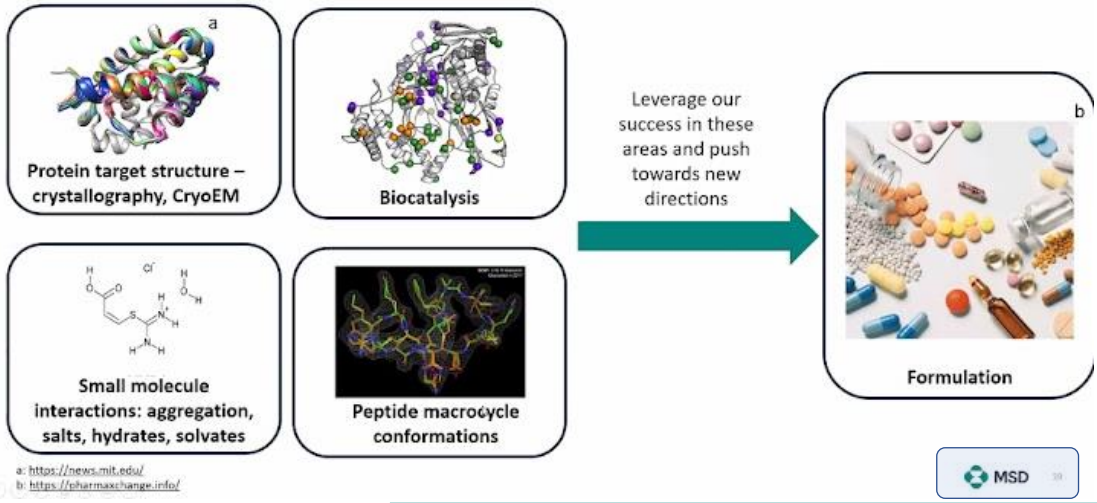


Dinâmica molecular

Resolução numérica de equações de movimento

Fernando Barroso, FCFRP/USP

Molecular dynamics across the discovery/development pipeline



Fernando Barroso, FCFRP/USP

THE JOURNAL OF CHEMICAL PHYSICS VOLUME 31, NUMBER 2 AUGUST, 1959

Studies in Molecular Dynamics. I. General Method*

B. J. Alder and T. E. Wainwright

Lawrence Radiation Laboratory, University of California, Livermore, California

(Received February 19, 1959)

A method is outlined by which it is possible to calculate exactly the behavior of several hundred interacting classical particles. The study of this many-body problem is carried out by an electronic computer which solves numerically the simultaneous equations of motion. The limitations of this numerical scheme are enumerated and the important steps in making the program efficient on the computers are indicated. The applicability of this method to the solution of many problems in both equilibrium and nonequilibrium statistical mechanics is discussed.

PHYSICAL REVIEW VOLUME 159, NUMBER 1 5 JULY 1967

Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules*

Loup Verlet†

Belfer Graduate School of Science, Yeshiva University, New York, New York

(Received 30 January 1967)

The equation of motion of a system of 864 particles interacting through a Lennard-Jones potential has been integrated for various values of the temperature and density, relative, generally, to a fluid state. The equilibrium properties have been calculated and are shown to agree very well with the corresponding properties of argon. It is concluded that, to a good approximation, the equilibrium state of argon can be described through a two-body potential.

SP

Table 9: History and extrapolated future of computer simulations of molecular dynamics. The future is deduced from extrapolation based on an observed increase of computing speed of a factor 10 every 5 years over the past decades (see Figure 31).

Year	Molecular system (type, size)	Length of the simulation [s]
1957	first molecular dynamics simulation (hard discs)	
1964	atomic liquid (argon)	10^{-11}
1971	molecular liquid (water)	5×10^{-12}
1977	protein in a vacuum	2×10^{-11}
1983	protein in water	2×10^{-11}
1989	protein–DNA complex in water	10^{-10}
1997	polypeptide folding in solvent	10^{-7}
2001	micelle formation	10^{-7}
200x	folding of a small protein	10^{-3}
And the future ...		
2001	biomolecules in water (ca. 10^4 atoms)	10^{-8}
2029	biomolecules in water (folding sooner?)	10^{-3}
2034	*E. coli* bacteria (ca. 10^{11} atoms)	10^{-9}
2056	mammalian cell (ca. 10^{15} atoms)	10^{-9}
2080	biomolecules in water (as fast as nature)	10^{6}
2172	human body (ca. 10^{27} atoms)	1


articles

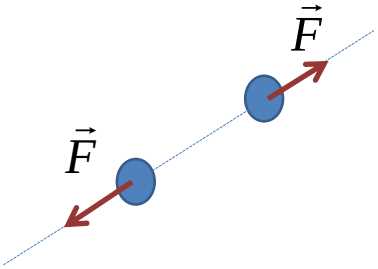
Dynamics of folded proteins

J. Andrew McCammon, Bruce R. Gelin & Martin Karplus

Department of Chemistry, Harvard University, Cambridge, Massachusetts 02138

Fernando Barroso, FCFRP/USP

Angew. Chem. Int. Ed. 2006, 45, 4064–4092



$$\vec{F} = m\vec{a} \qquad \vec{F} = -\nabla V(\vec{r})$$

$$F_i = m_i a_i \qquad F_i = -\nabla_i V \qquad -\frac{dV}{dr_i} = m_i \frac{d^2 r_i}{dt^2} \qquad F_i = m_i a_i$$

Posição $\mathbf{r_i(t)}$

Energia Potencial

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

$$\vec{F} = m\frac{dv}{dt} \xrightarrow{\text{Forças conservativas}} v(t) = \frac{F}{m}t + v_0$$

Equação de movimento

$$v(t) = \frac{dr}{dt} = \frac{F}{m}t + v_0$$

$$r(t) = \frac{F}{2m}t^2 + v_0 t + r_0$$

Solução (F = cte)

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

Numericamente:

$$r(t+\delta t) = r(t) + v(t)\delta t + \frac{1}{2}a(t)\delta t^2 + \ldots$$

$$v(t+\delta t) = v(t) + a(t)\delta t + \frac{1}{2}b(t)\delta t^2 + \ldots$$

$$a(t+\delta t) = a(t) + b(t)\delta t + \ldots$$

77

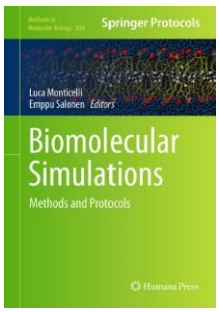
- É fundamental a estabilidade numérica!
- Conservação de energia total e momento para grande δt

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

2.2. The Verlet Algorithm

The Verlet algorithm (10) is obtained by adding together the Taylor expansions of $\vec{r}(t+\Delta t)$ and $\vec{r}(t-\Delta t)$ about time t. It reads as follows:

$$\vec{r}_i(t+\Delta t) = 2\vec{r}_i(t) - \vec{r}_i(t-\Delta t) + \frac{1}{m}\vec{F}_i(t)(\Delta t)^2. \qquad (37)$$



Não envolve velocidades, más,....

$$\vec{v}_i(t) = \frac{\vec{r}_i(t+\Delta t) - \vec{r}_i(t-\Delta t)}{2\Delta t}.$$

78

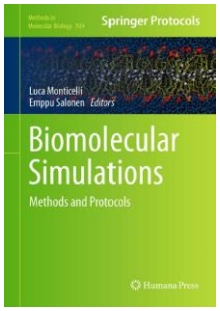
© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

2.3. The Leapfrog Algorithm

The leapfrog algorithm (12) is a modification of the Verlet algorithm, given by

$$\vec{v}_i\left(t+\frac{\Delta t}{2}\right)=\vec{v}_i\left(t-\frac{\Delta t}{2}\right)+\frac{1}{m_i}\vec{F}_i(t)\Delta t \qquad (39)$$

$$\vec{r}_i(t+\Delta t)=\vec{r}_i(t)+\vec{v}_i\left(t+\frac{\Delta t}{2}\right)\Delta t. \qquad (40)$$

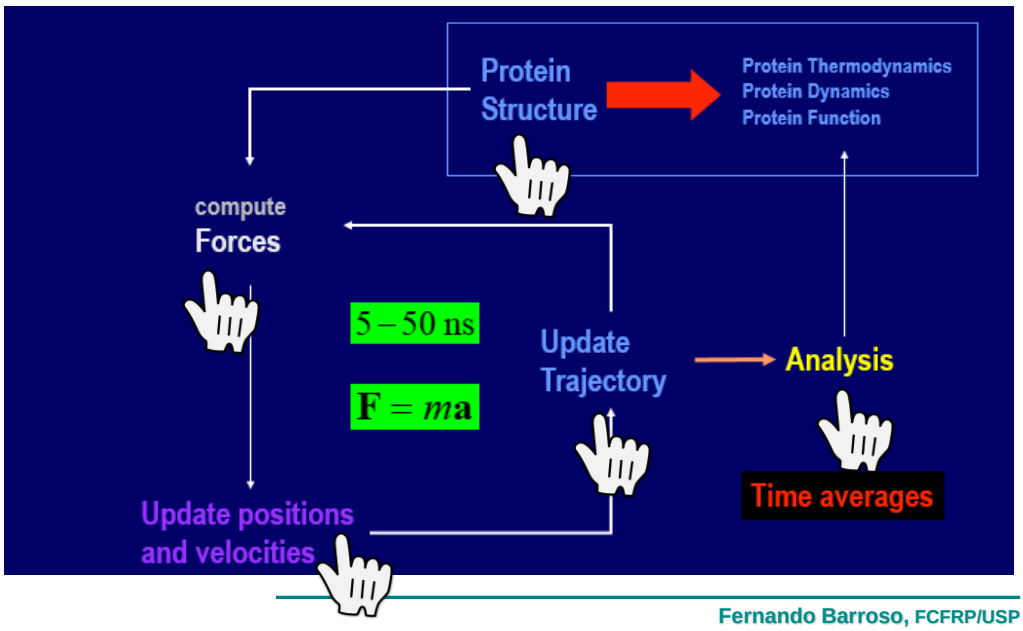


$$\vec{v}_i=\frac{\vec{v}_i\left(t+\frac{\Delta t}{2}\right)+\vec{v}_i\left(t-\frac{\Delta t}{2}\right)}{2}$$

79

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br

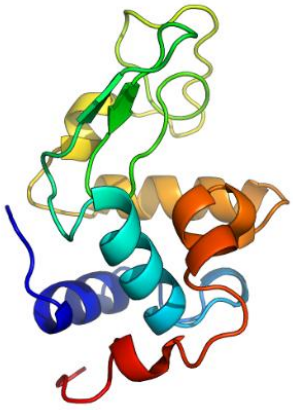
Dinâmica Molecular



GROMACS Tutorial

Lysozyme in Water

Justin A. Lemkul, Ph.D.
Virginia Tech Department of Biochemistry



Laboratory of Computational Biophysical Chemistry

- ✓ Gromacs
- ✓ Charmm
- ✓ Amber
- ✓ Discover, Insight
- ✓ Sigma, Tripos, ...
- ✓ NAMD
- ✓ Tinker
- ✓ Lammps...

GROMACS FAST. FLEXIBLE. FREE.

LAMMPS

http://www.mdtutorials.com/gmx/lysozyme/index.html

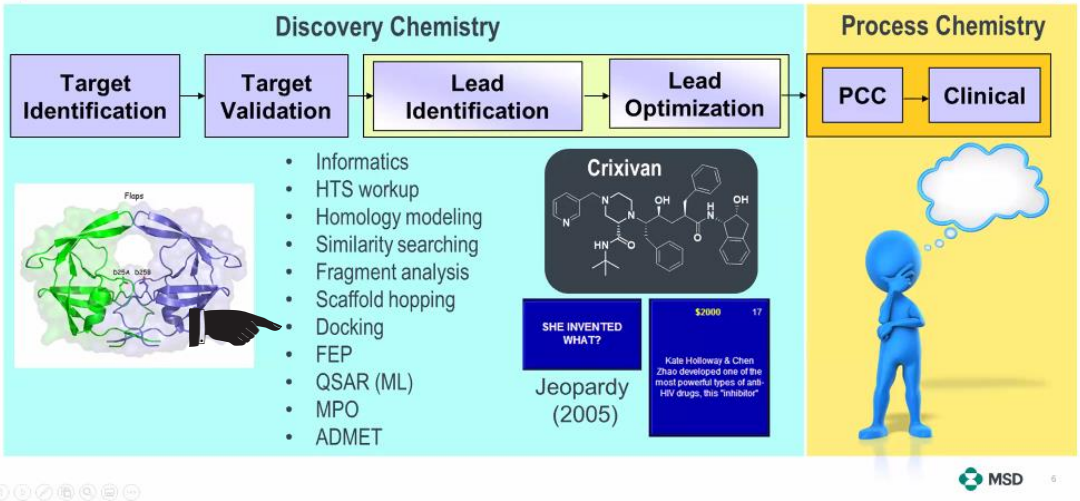
Laboratory of Computational Biophysical Chemistry

Bons livros para DM

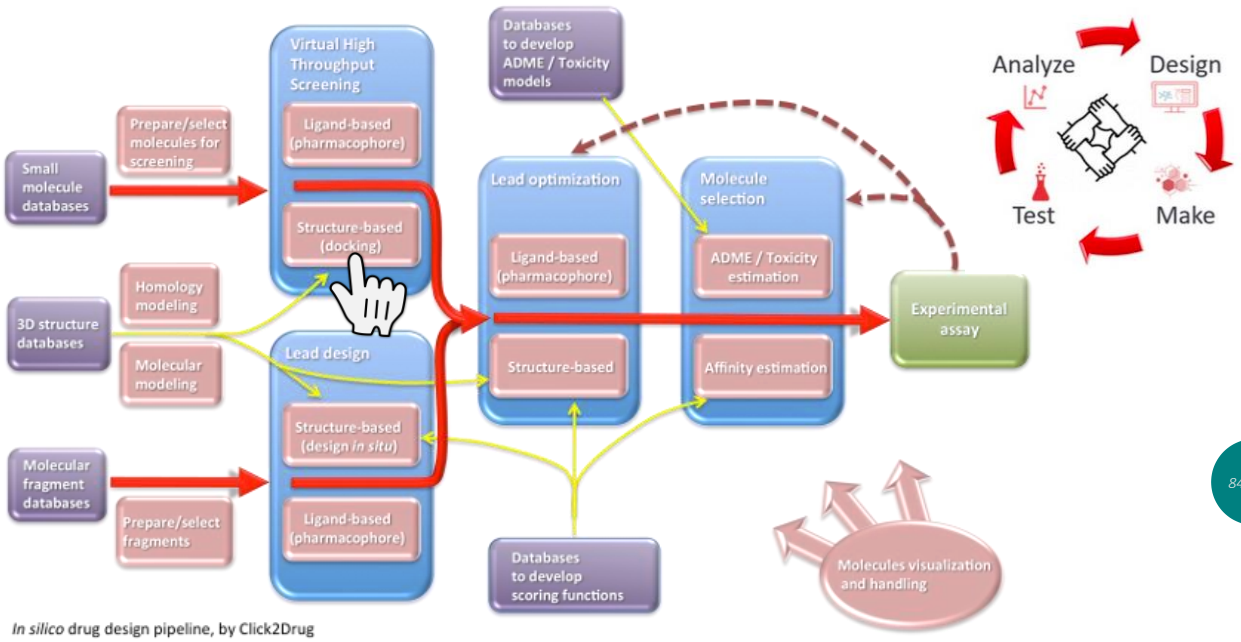
1. M.P.Allen e D.J. Tildesley, *Computer Simulation of Liquids*, (Oxford University Press, 1989).
2. D. Frenkel e B. Smid, *Understanding Molecular Simulation*, (Academic Press, 2001).
3. T. Schlick, *Molecular Modeling and Simulation*, (Springer, 2010).
4. J.M. Haille, *Molecular Dynamics Simulation: Elementary Methods*, (Wiley-Interscience, 1997).
5. D.C. Rappaport, *The Art of Molecular Dynamics Simulation*, (Cambridge University Press, 1997).

Fernando Barroso, FCFRP/USP

Research pipeline

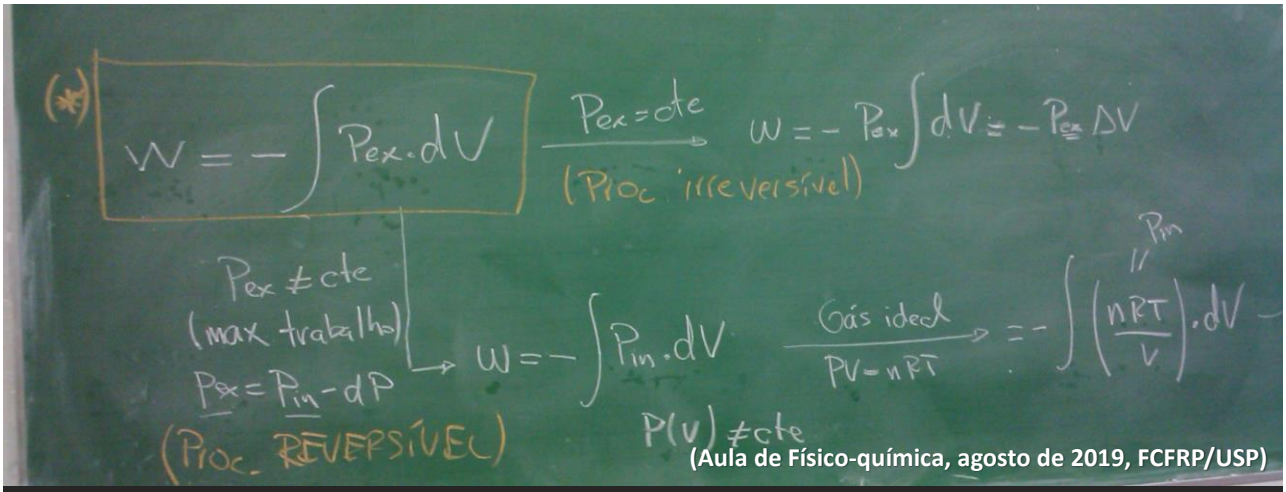


Fernando Barroso, FCFRP/USP



In silico drug design pipeline, by Click2Drug

© Fernando Barroso, FCFRP/USP
flbarroso@usp.br



(Aula de Físico-química, agosto de 2019, FCFRP/USP)

Obrigado!

Fernando Barroso (flbarroso@usp.br)