Computational Drug Design & Molecular Dynamics : an HPC perspective

Jean-Philip Piquemal

Sorbonne Université

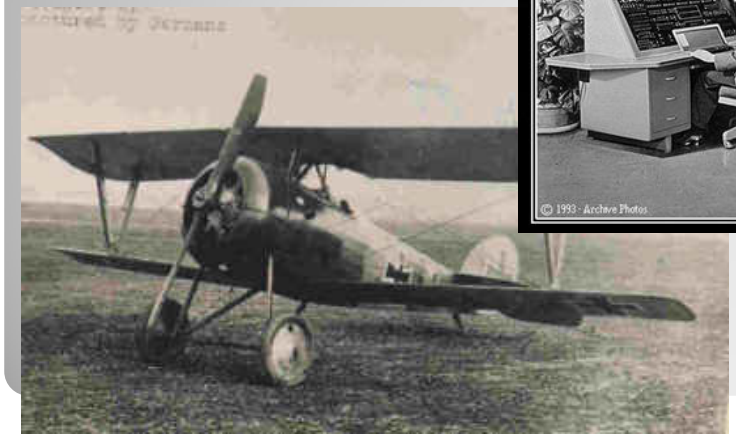
& Institut Universitaire de France

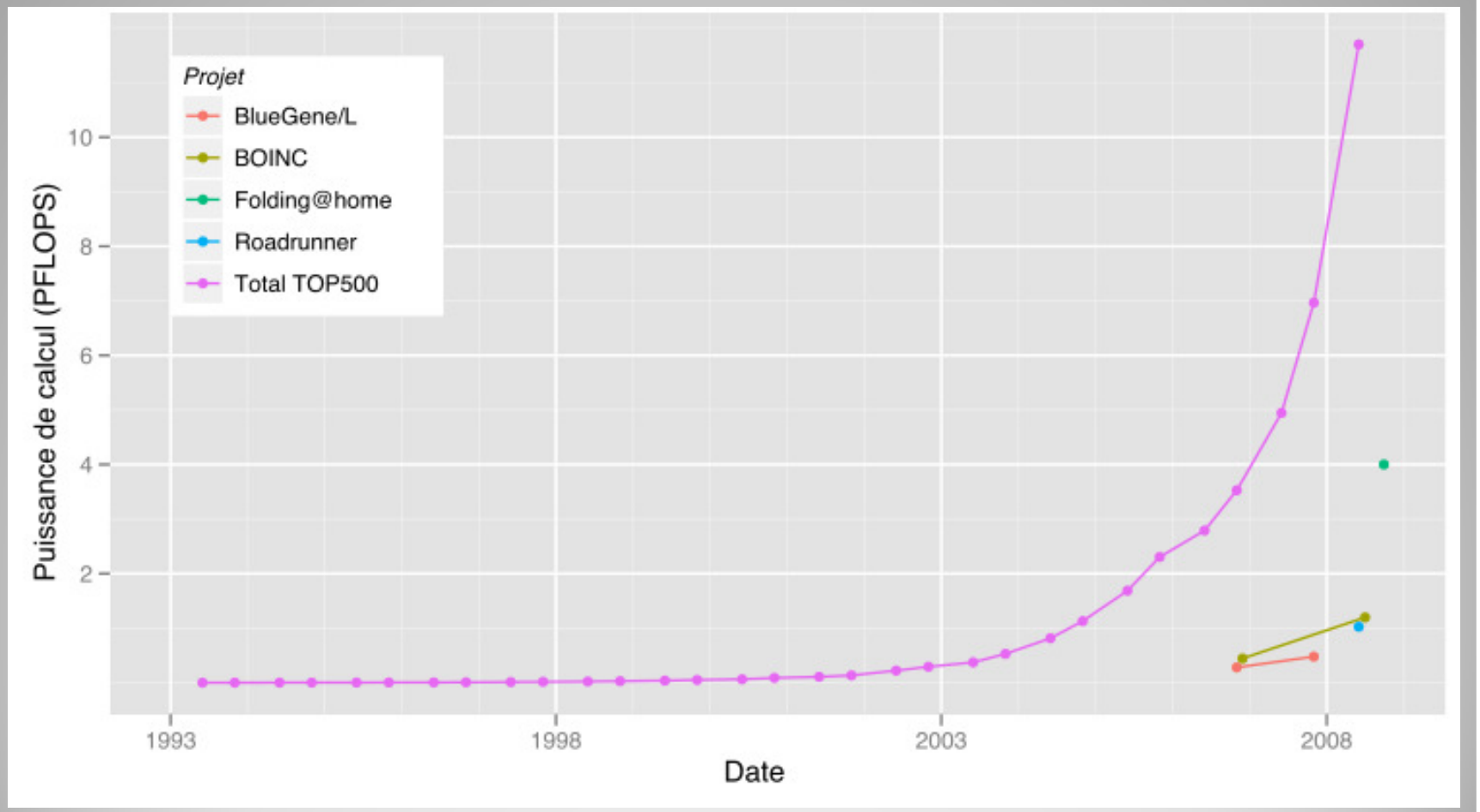
& University of Texas at Austin

Le Lab Quantique

02/04/2020

Drug Design & High-Performance Computing

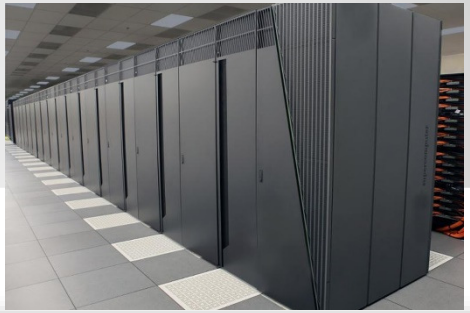




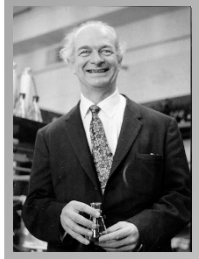
Source top 500
https://commons.wikimedia.org/w/index.php?curid=8740585

The situation changed…

The huge increase of HPC capabilities now enable to use advanced computational chemistry approaches.

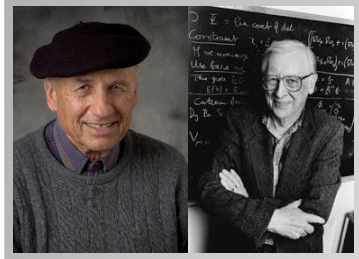


Computational Chemistry



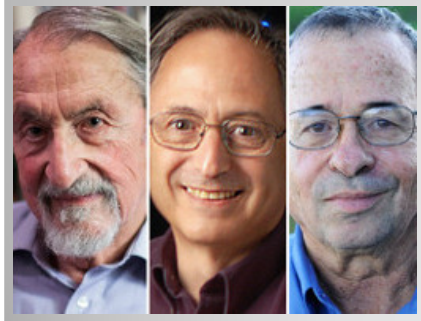
Linus Pauling (1954):

- The nature of the chemical bond



Walter Kohn & **John Pople** (1998):

- development of computational quantum chemistry
- Density Functional Theory (DFT)

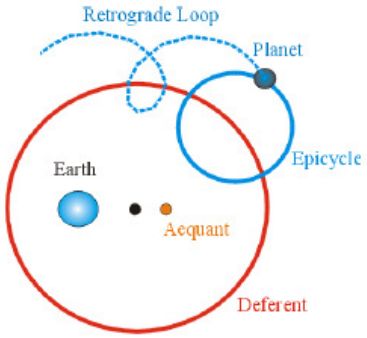


Martin Karplus, **Michael Levitt** et **Arieh Warshel** (2013):

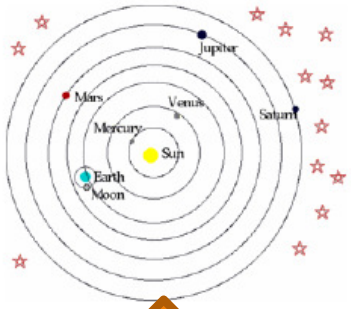
- Molecular Dynamics
- Chemical reactivity Modeling

Atomistic simulations:
Quantum or Classical Mechanics?

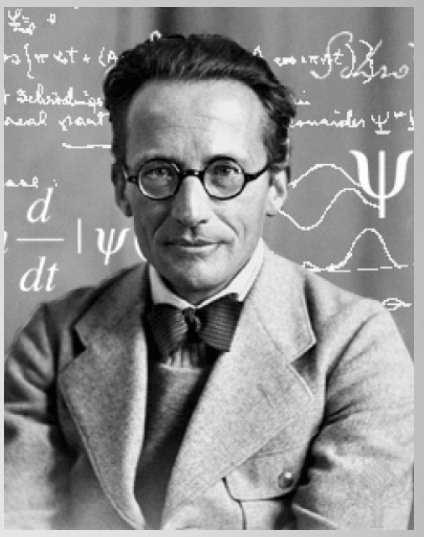
Classical mechanics: Molecular dynamics



$$E_{total} = E_{bond} + E_{angle} + E_{bond\text{-}angle} + E_{torsion} + E_{repulsion} + E_{dispersion} + E_{charge} + E_{polarization} + E_{solvation} + E_{cross\text{-}terms} + E_{????}$$



$$\hat{H}\Psi = E\Psi$$



Quantum Mechanics: Quantum Chemistry

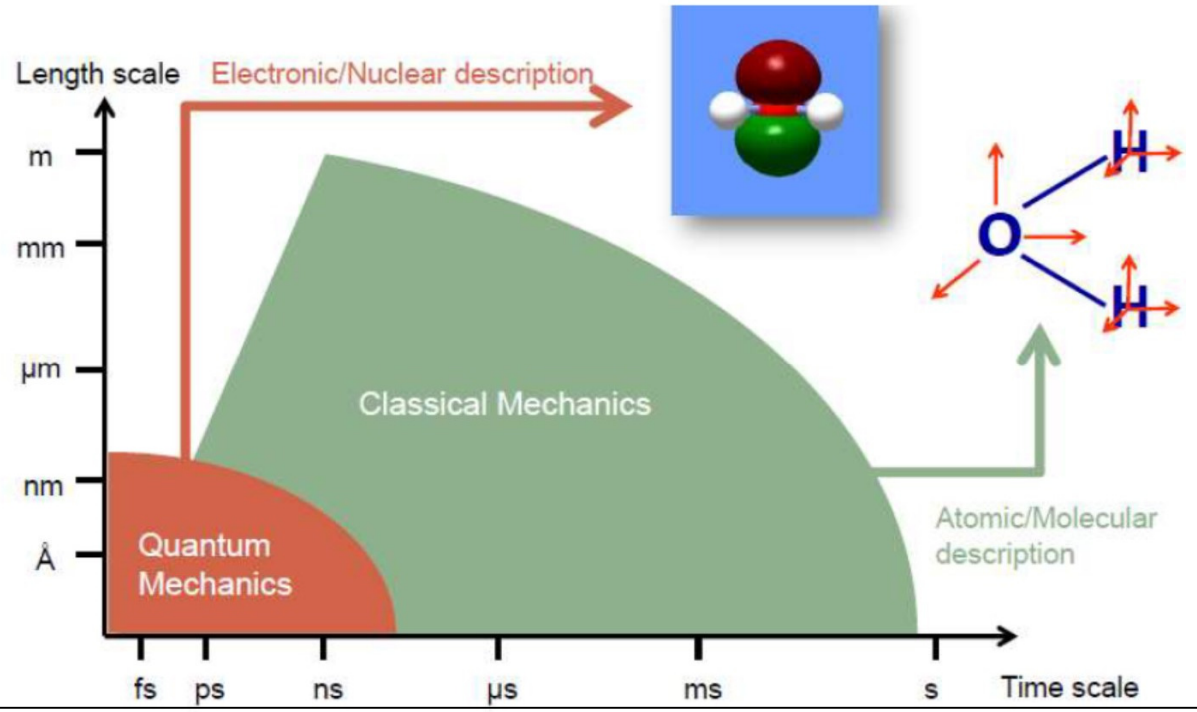
Timescales…and Biology

Bond vibration - 1 fs
Collective vibrations- 1 ps
Conformational transition- ps or more
Enzymatic catalysis- microsecond/millisecond
Docking (ligand biding to a target) - micro/millisecond
Protein folding - millisecond/second

Molecular Dynamics:

Integration timestep - 1 femtoseconde
Accessible timescale about a few milliseconds
(DE. Shaw research).

Timescales (2)…



Molecular Dynamics:

Integration timestep - 1 femtoseconde

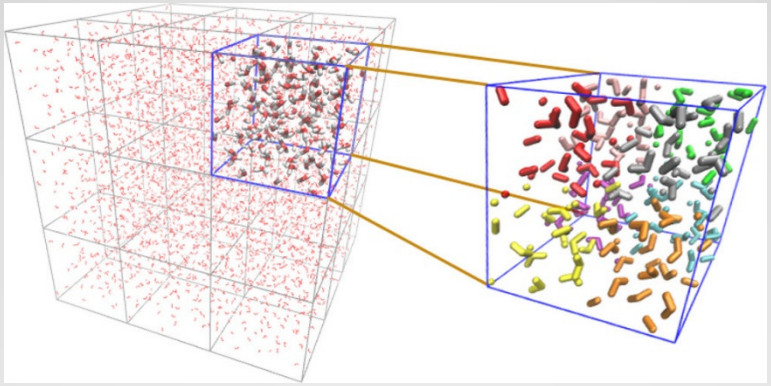
Accessible timescale about a few milliseconds (DE. Shaw research).

MOLECULAR SIMULATION AT A GLANCE

Molecular Dynamics notably allows one to :

study complex molecular processes with atomic-scale space-time resolution

virtual microscope

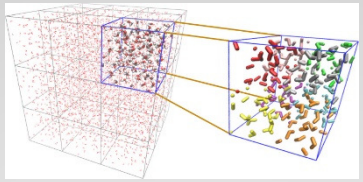


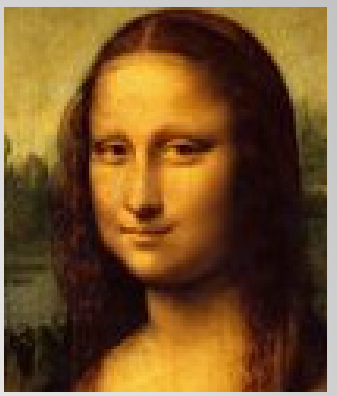
predict the properties of new molecules, materials, and nanodevices

atomic-scale computer-aided design and engineering

INTRODUCTION TO NEW GENERATION MOLECULAR DYNAMICS

TO BE QUANTITATIVE/PREDICTIVE,
SUCH SIMULATIONS MUST RELY ON MORE PHYSICS

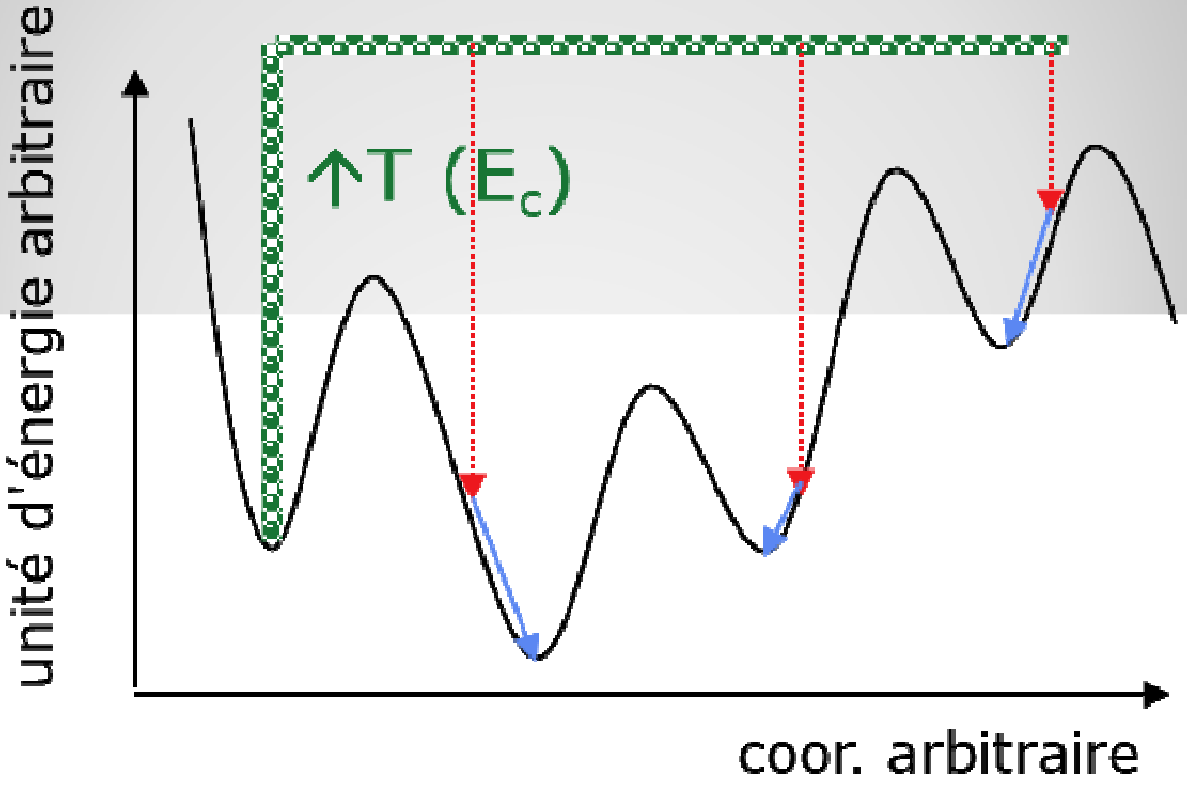




CLASSICAL FORCE FIELDS (2-BODY PHYSICS)

POLARIZABLE FORCE FIELDS (N-BODY PHYSICS)

Molecular dynamics



1st simulations (1957/59)

Phase Transition for a Hard Sphere System

B. J. Alder and T. E. Wainwright
University of California Radiation Laboratory, Livermore, California
(Received August 12, 1957)

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)

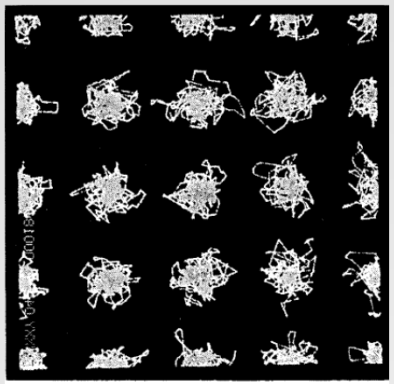
A method is outlined by which it is possible to calculate exactly the behavior of several hundred interacting classical particles. The study of this many-body problem is carried out by an electronic computer which solves numerically the simultaneous equations of motion. The limitations of this numerical 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.

Hard spheresdures

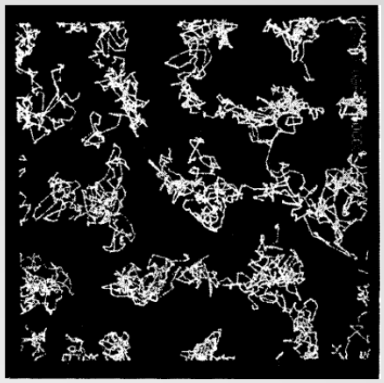
$$u_{ij}(r) = \begin{cases} 0 & r > d \\ \infty & r \le d \end{cases}$$

(computations of the collision time)

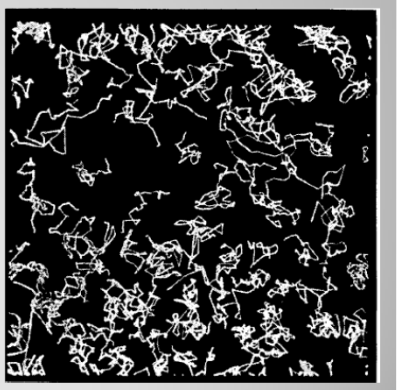
IBM-704:



Solide pahse



Liquide phase



Interface liquid-vapor

N=32: 7000 collisions / h
N=500: 500 collisions / h

Production time
~20000 step

N=32 → 6.5×10^5 coll. → 4 days
N=500 → 10^7 coll. → 2.3 years

First properties (1964)

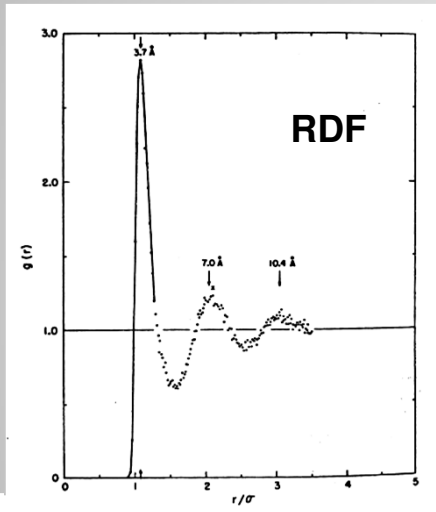
CDC-3600

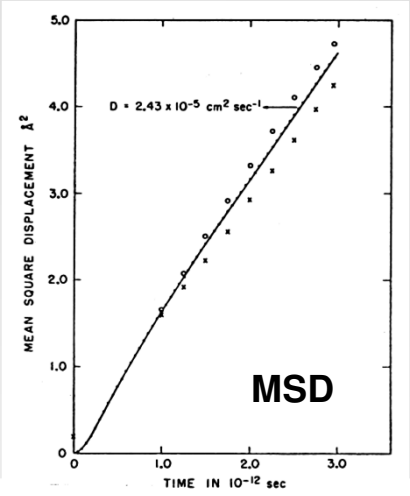
PHYSICAL REVIEW VOLUME 136, NUMBER 2A 19 OCTOBER 1964

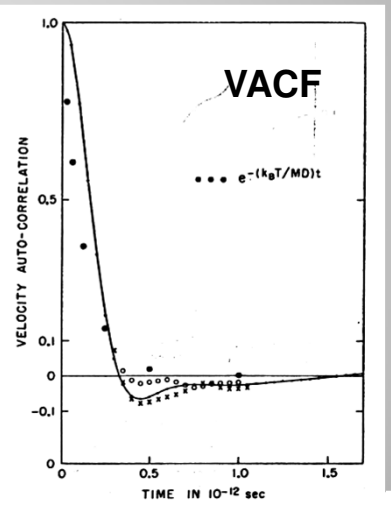
Correlations in the Motion of Atoms in Liquid Argon*

A. RAHMAN
Argonne National Laboratory, Argonne, Illinois
(Received 6 May 1964)

A system of 864 particles interacting with a Lennard-Jones potential and obeying classical equations of motion has been studied on a digital computer (CDC 3600) to simulate molecular dynamics in liquid argon at 94.4°K and a density of 1.374 g cm^{-3}. The pair-correlation function and the constant of self-diffusion are found to agree well with experiment; the latter is 15% lower than the experimental value. The spectrum of the velocity autocorrelation function shows a broad maximum in the frequency region $\omega=0.25(k_BT/\hbar)$. The shape of the Van Hove function $G_s(r,t)$ attains a maximum departure from a Gaussian at about $t=3.0\times10^{-12}$ sec and becomes a Gaussian again at about 10^{-11} sec. The Van Hove function $G_d(r,t)$ has been compared with the convolution approximation of Vineyard, showing that this approximation gives a too rapid decay of $G_d(r,t)$ with time. A delayed-convolution approximation has been suggested which gives a better fit with $G_d(r,t)$; this delayed convolution makes $G_d(r,t)$ decay as t^4 at short times and as t at long times.



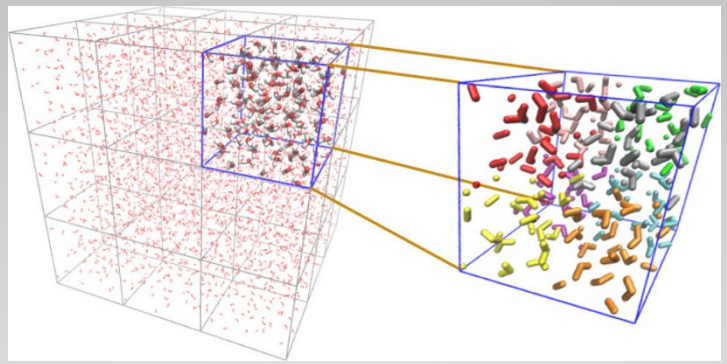


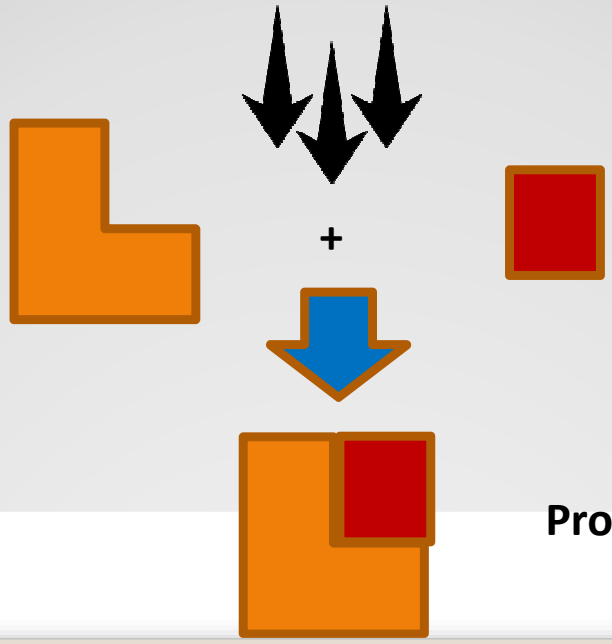


864 particules
Time per step ~ 45s

Production time ~20000 steps ≅ 10 days!
(today a few secondes on a cellphone ;)…)

Docking & drug design: MD for drug discovery





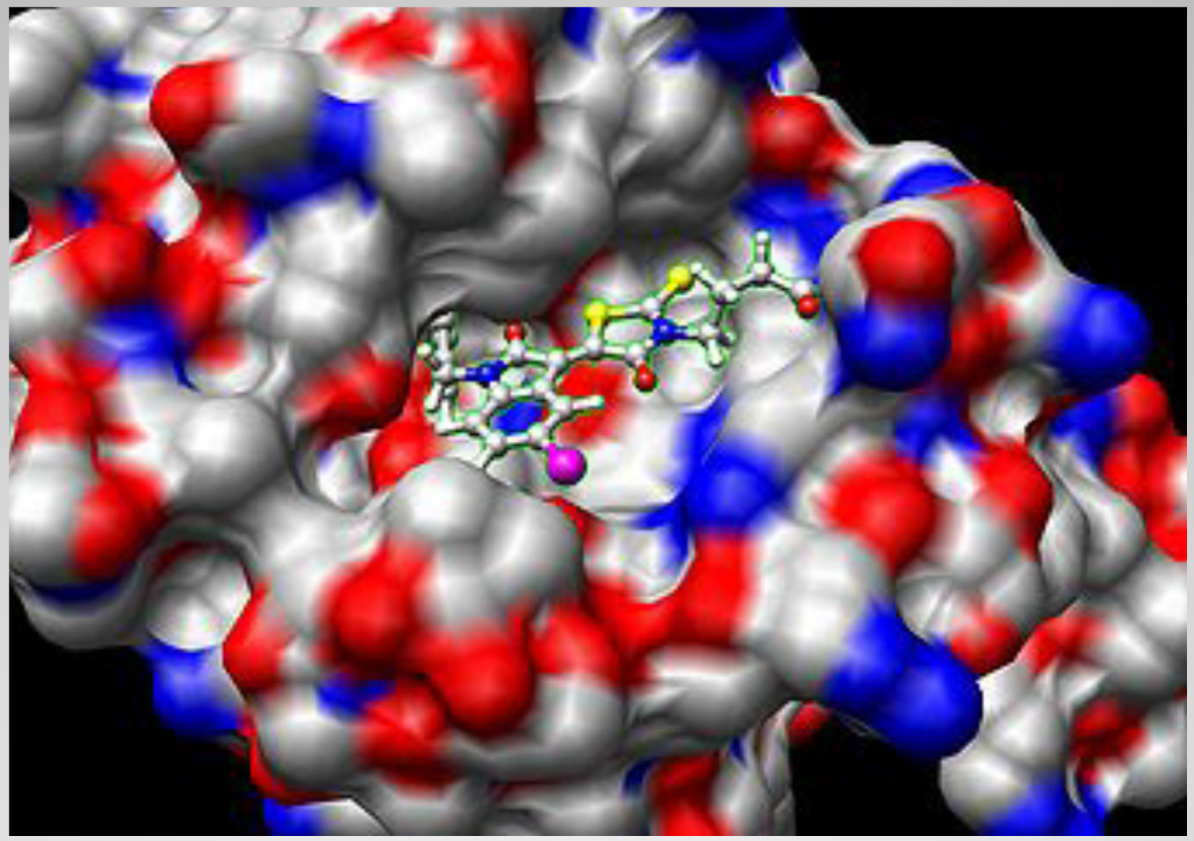
Target: protein (or enzyme)

+

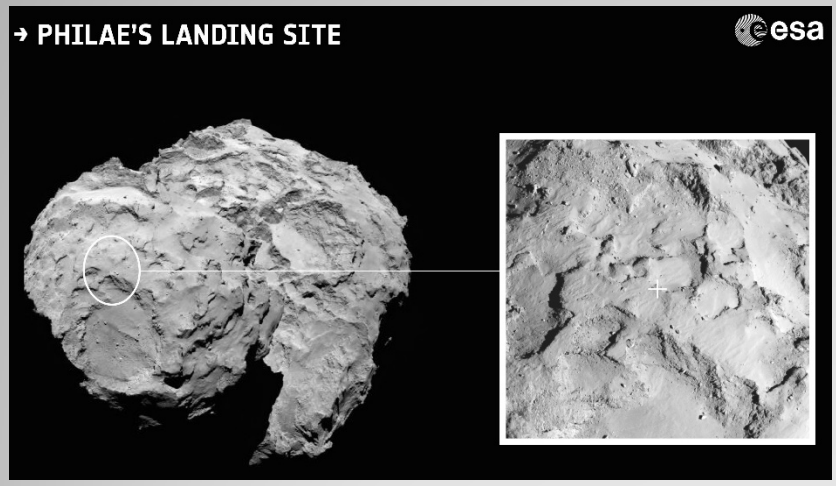
Ligand (drug)

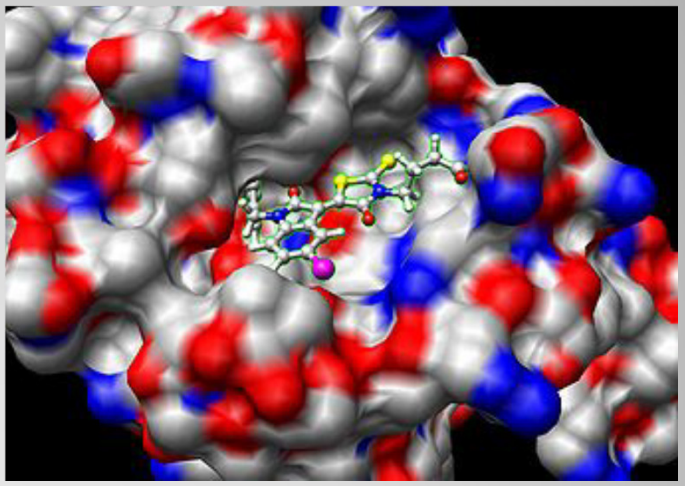
Protein-ligand complex

Docking & drug design: MD for drug discovery



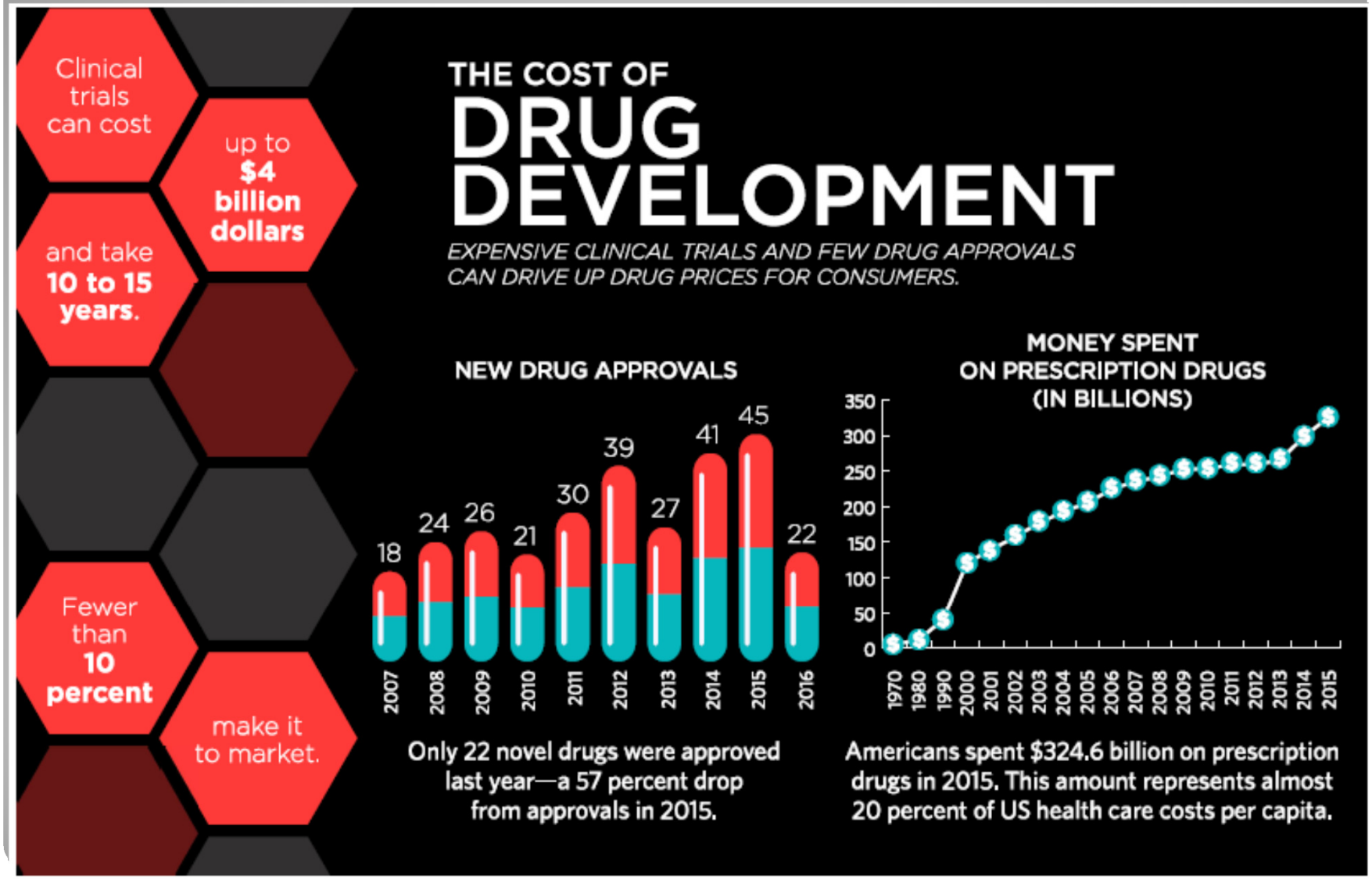
Docking & drug design: MD for drug discovery





Computing the most favourable location for a drug interacting with a protein is "as hard" as landing a probe on a comet (i.e. ESA Rosetta mission).

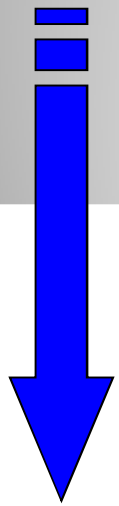
Great need of computational precision !



THE SCIENTIST STAFF; DATA FROM FDA AND CMS

Discovery of a new drug: a long process…

Research



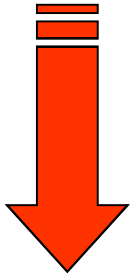
Target Selection

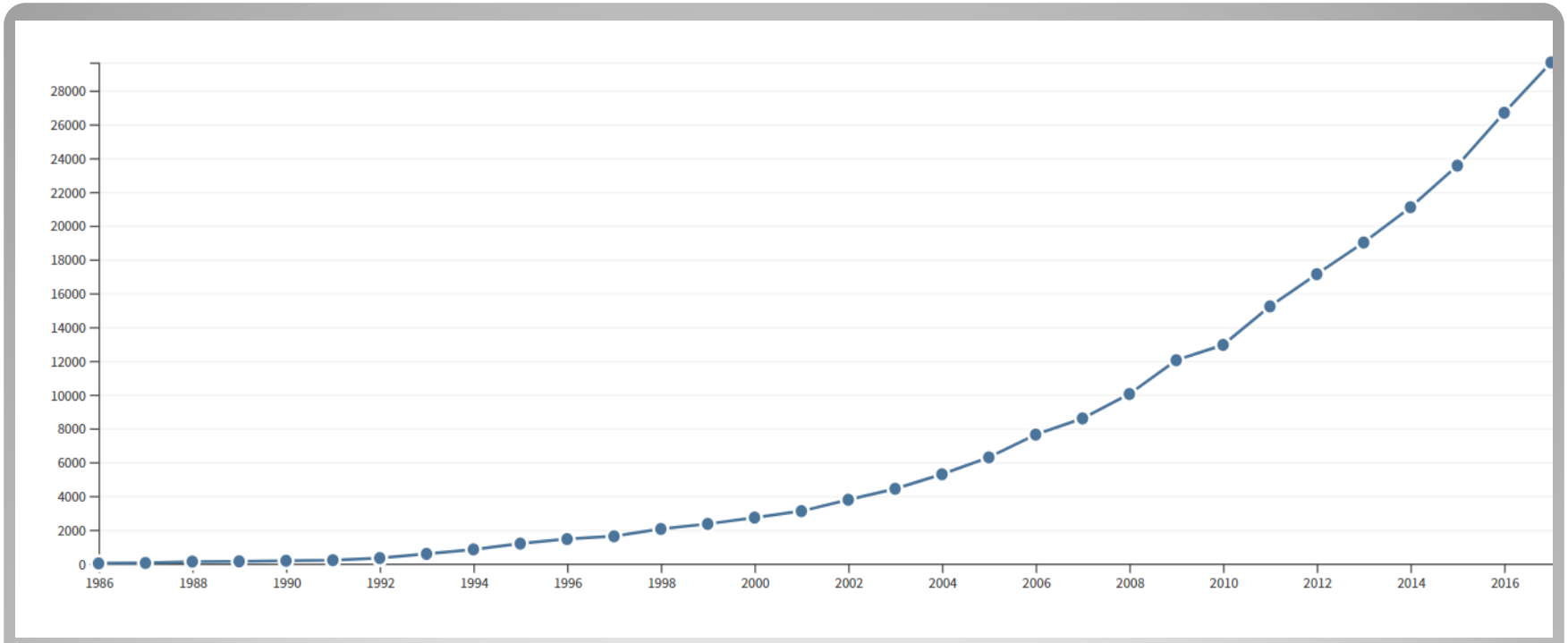
Hit finding (Virtual screening, databases…)

Hit to lead **(progressive elimination)**
docking, sélectivity

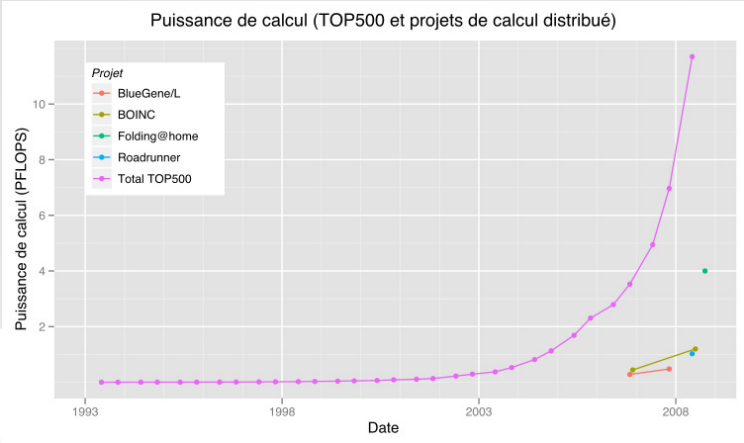
Lead Optimization **(best hit optimization)**

Drug Development





Rational design and molecular modelling in drug discovery.



Rational design in the industry

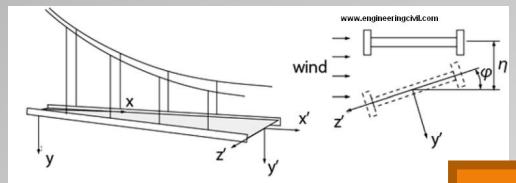


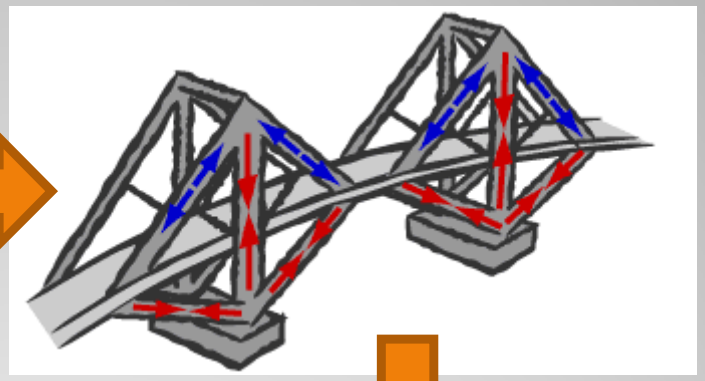
Figure 1. Theoretical model of suspension bridge.

Hirai's research on lateral torsional buckling of suspension bridge starts at the Equation 1.

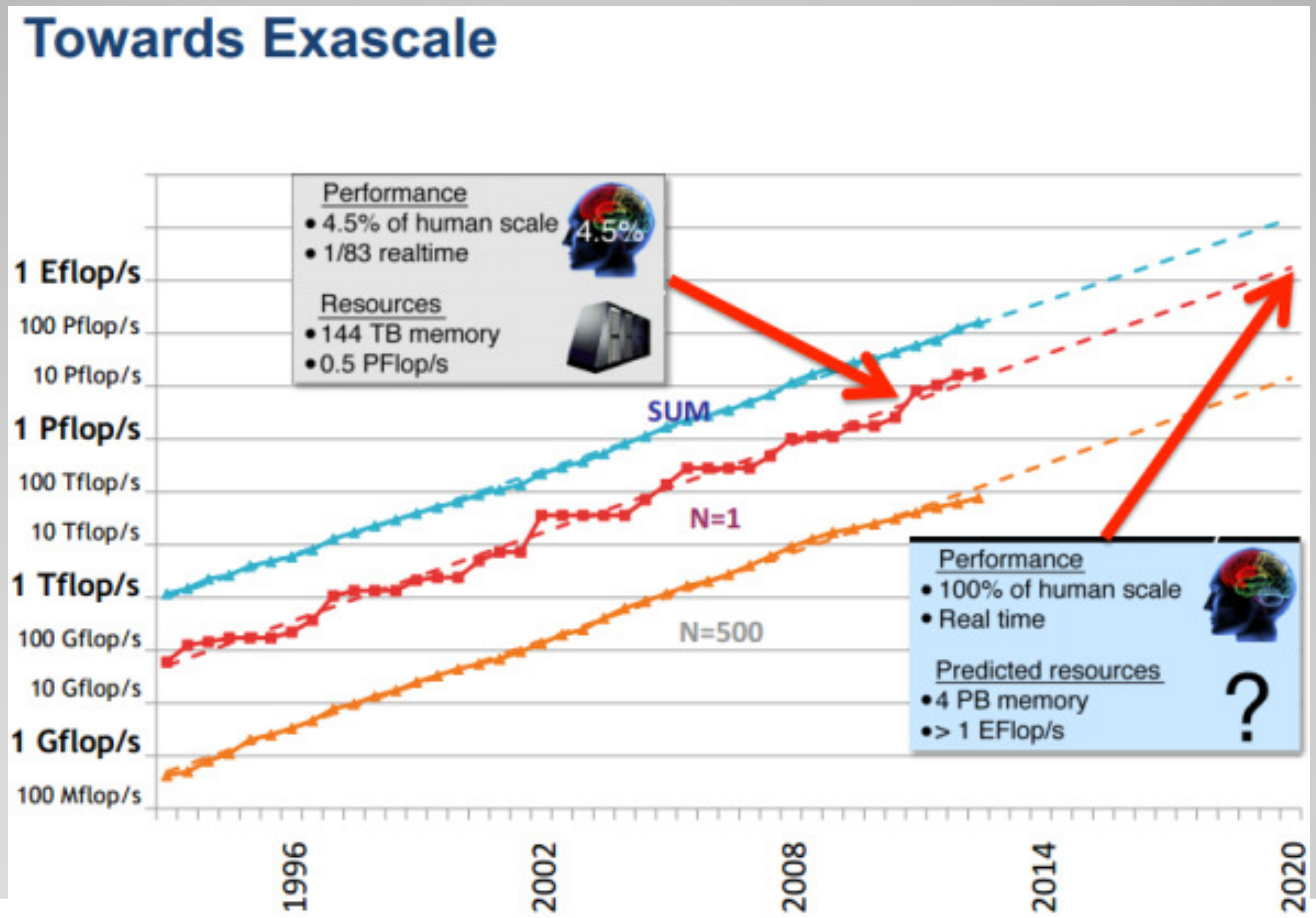
$$EI\frac{d^4\eta}{dx^4} - 2H_\omega\frac{d^4\eta}{dx^4} - 2h_1\frac{d^2y}{dx^2} + \frac{d^2}{dx^2}(\mathrm{M}\varphi) - (S + (C_d))pb\varphi = 0$$

$$\mathrm{M}\frac{d^2\eta}{dx^2} - EC_\omega\frac{d^4\eta}{dx^4} - \left(GK + \frac{H_\omega b^2}{2}\right)\frac{d^2\eta}{dx^2} - bh_2\frac{d^2y}{dx^2} - S_1pb\varphi^2 = 0 \quad (1)$$

Where, η and φ mean main girder's buckling displacement in vertical and torsional



Rational design: the need for physics, mathematics and computer sciences



Crédits: Extreme tech

To perform new generation MD simulations: you need advanced softwares able to use modern hardware

3 steps :

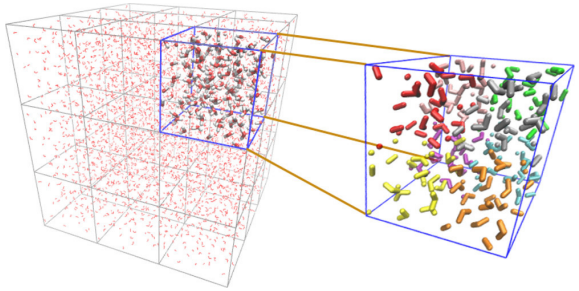
1. thermalization
2. Equilibration
3. Production

CPUs (massive parallelism)
GPUs (massive parallelism)
OPUs/TPUs
Quantum Computing

Dedicated Hardware (DE. Shaw) /FGPAs

SCIENCES SORBONNE UNIVERSITÉ

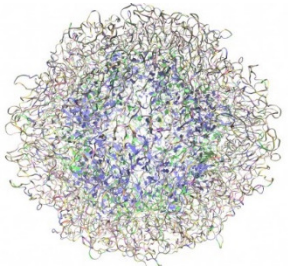
TINKER-HP



Tinker-HP
high performance and high precision

Tinker-HP is a new molecular modeling platform allowing for molecular dynamics simulations using highly precise new methods (new generation polarizable force fields).

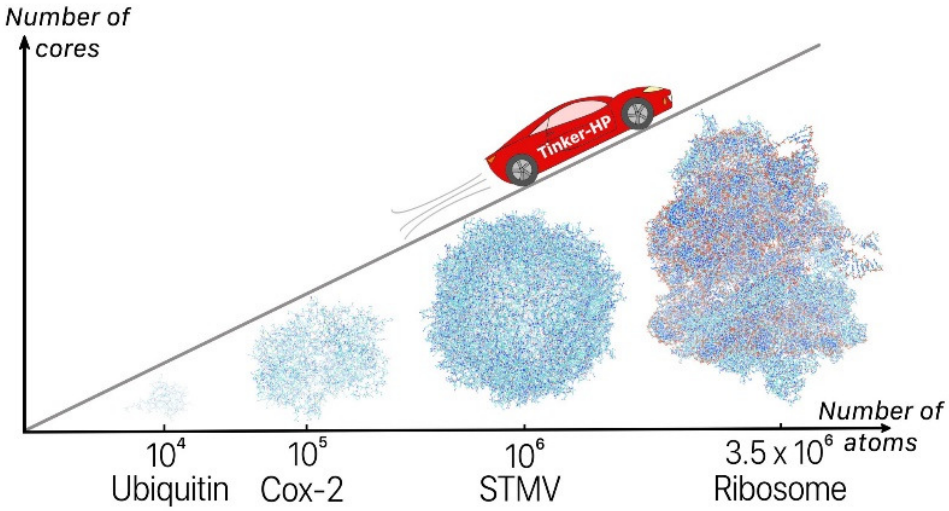
Massively parallel MPI implementation on CPUs and GPUs.



- Objectives : to simulate the dynamical time evolution of a molecular system thanks to Newton laws.

- **Such methodology is million times faster than quantum mechanics but remains accurate thanks to new generation force fields grounded on quantum chemistry and including many-body effects.**

Grand challenge: optimization and applications of Tinker-HP

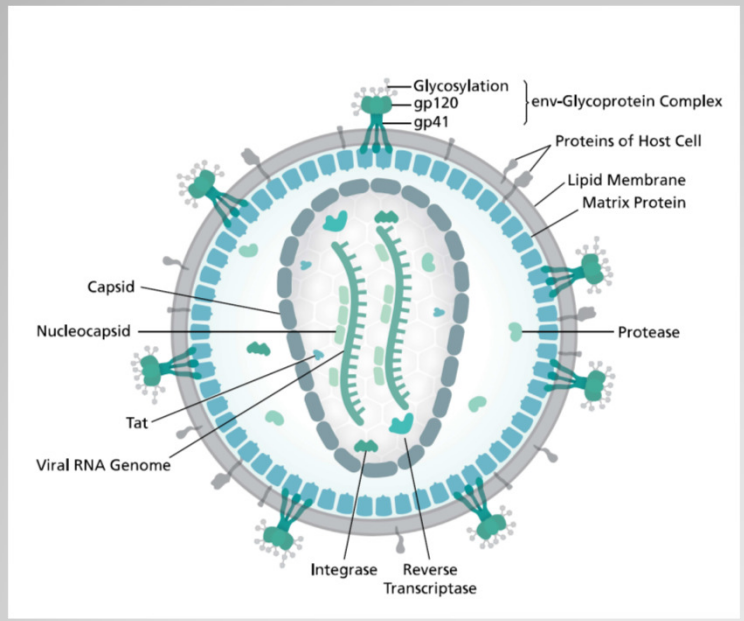


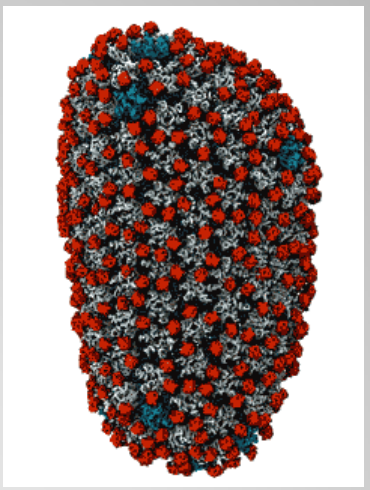
Tinker-HP

L. Lagardère, L.-H. Jolly, F. Lipparini, F. Aviat, B. Stamm, Z. F. Jing, M. Harger, H. Torabifard, G. A. Cisneros, M.J. Schnieders, N. Gresh, Y. Maday, P. Ren, J.. W. Ponder, J.-P. Piquemal,
Chemical Science, **2018**, 9, 956-972 (open access),
DOI: 10.1039/C7SC04531J

Website: http://tinker-hp.ip2ct.upmc.fr/

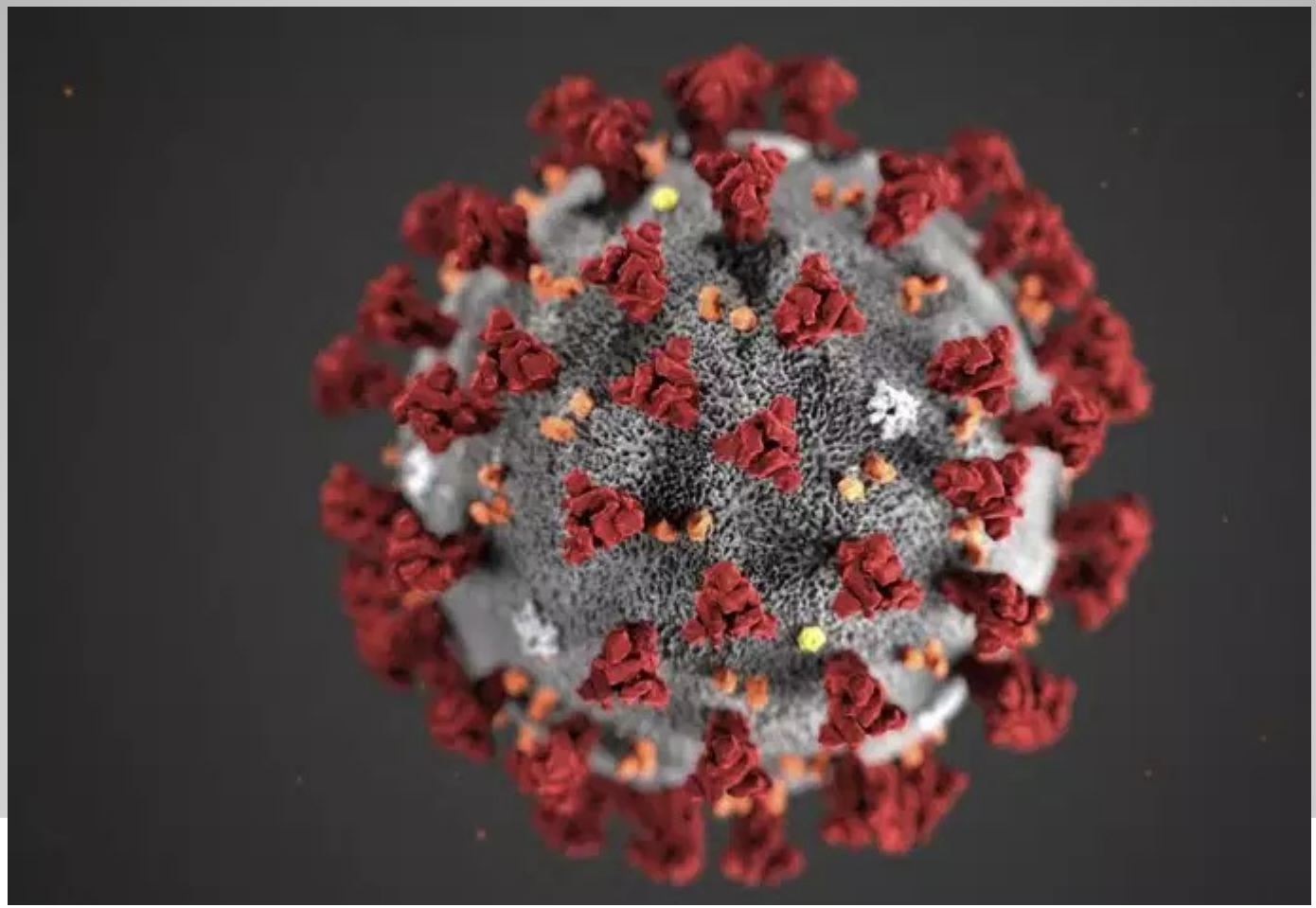
Perspectives: HIV-1 nucleocapsid & Capsid





3.5 millions atoms possible with petascale machines.
Full virus using pre-exascale (EMC2 ERC project)
Mixed-precision: multi-GPUs/GPUs-CPUs.

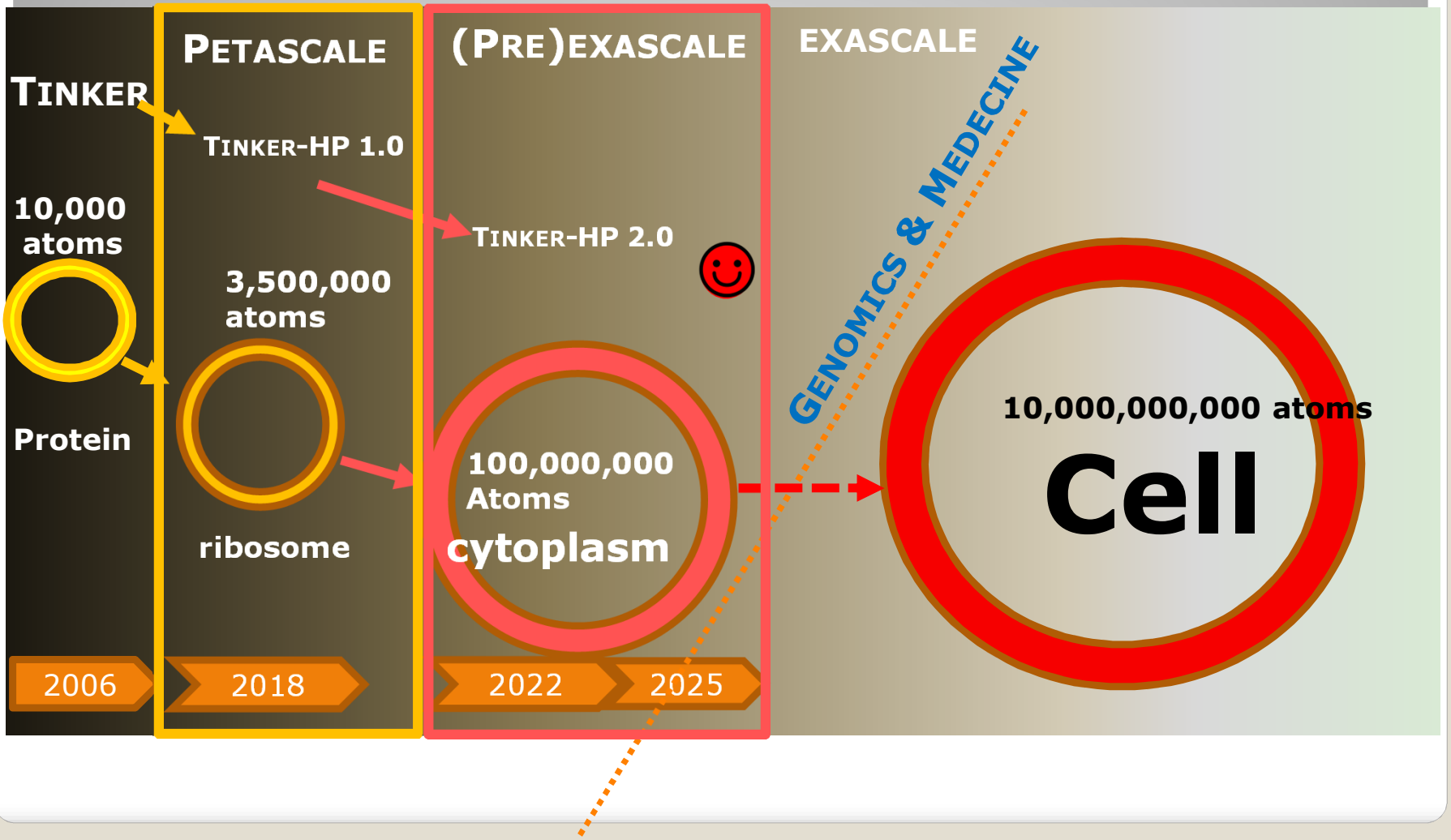
But a coranovirus decided to change our plans…



Towards a cell (or virus)…

ERC Synergy EMC2(grant agreement No 810367)
Extreme-scale Mathematically-based Computational Chemistry

European Research Council



QUANTONATION