European Research Council

Established by the European Commission

Molecular simulation and the numerical microscope

Gabriel STOLTZ

gabriel.stoltz@enpc.fr

Journées scientifiques Inria, Bordeaux, juin 2018

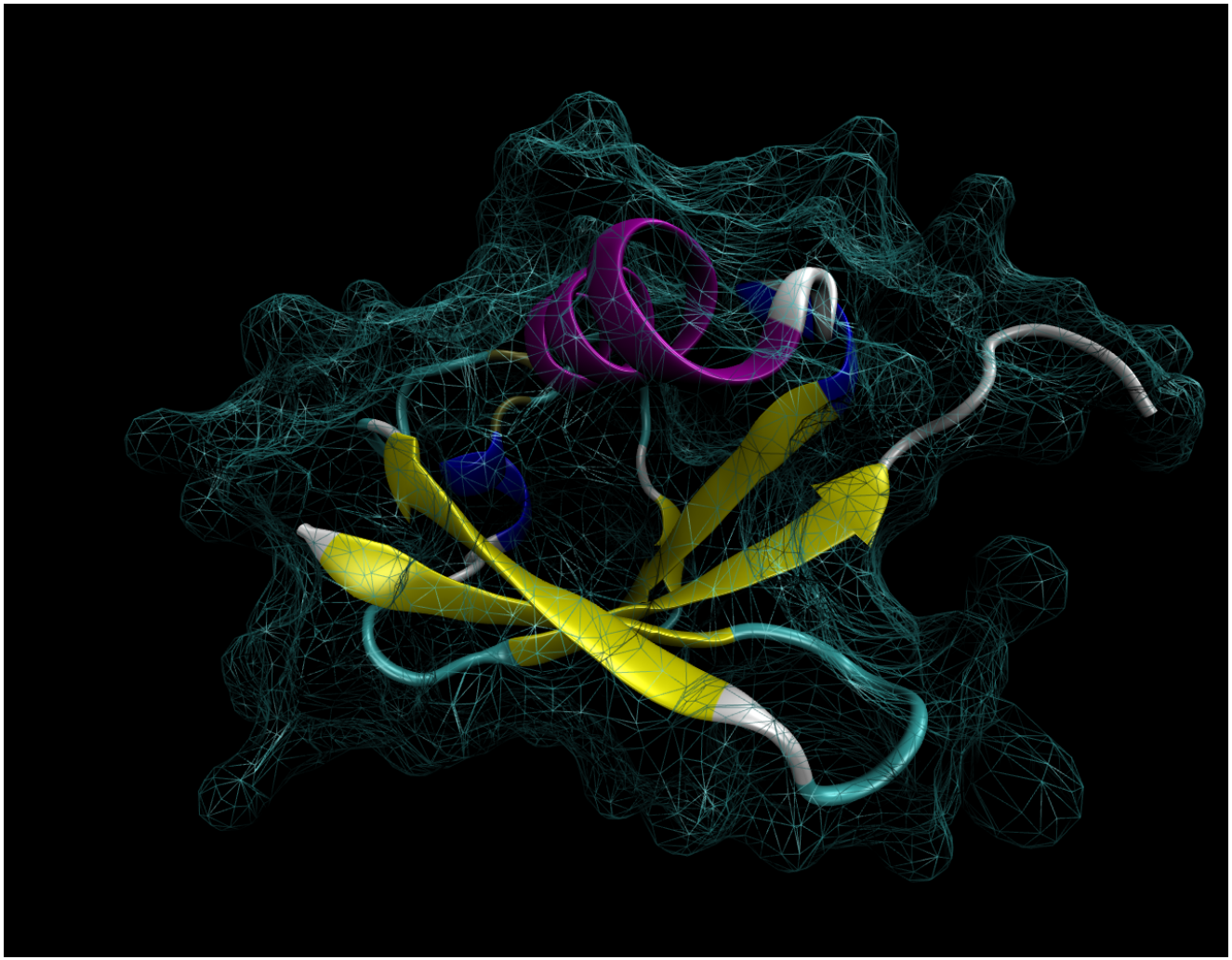
Computational Statistical Physics

- *Predict macroscopic properties of materials from their microscopic description*

- **Microstate**
  - positions $q = (q_1, \ldots, q_N)$ and momenta $p = (p_1, \ldots, p_N)$
  - energy of the system $H(q,p) = V(q) + \sum_{i=1}^{N} \frac{p_i^2}{2m_i}$
  - (almost) all the physics is in the choice of V...
- **Macrostate**
  - described by a probability measure μ
  - constraints fixed exactly or in average (number of particles, volume, energy)

- **Properties:** static (equation of state, heat capacities, etc) and dynamical (transport coefficient, transition pathway, etc)

Examples of molecular systems (1)

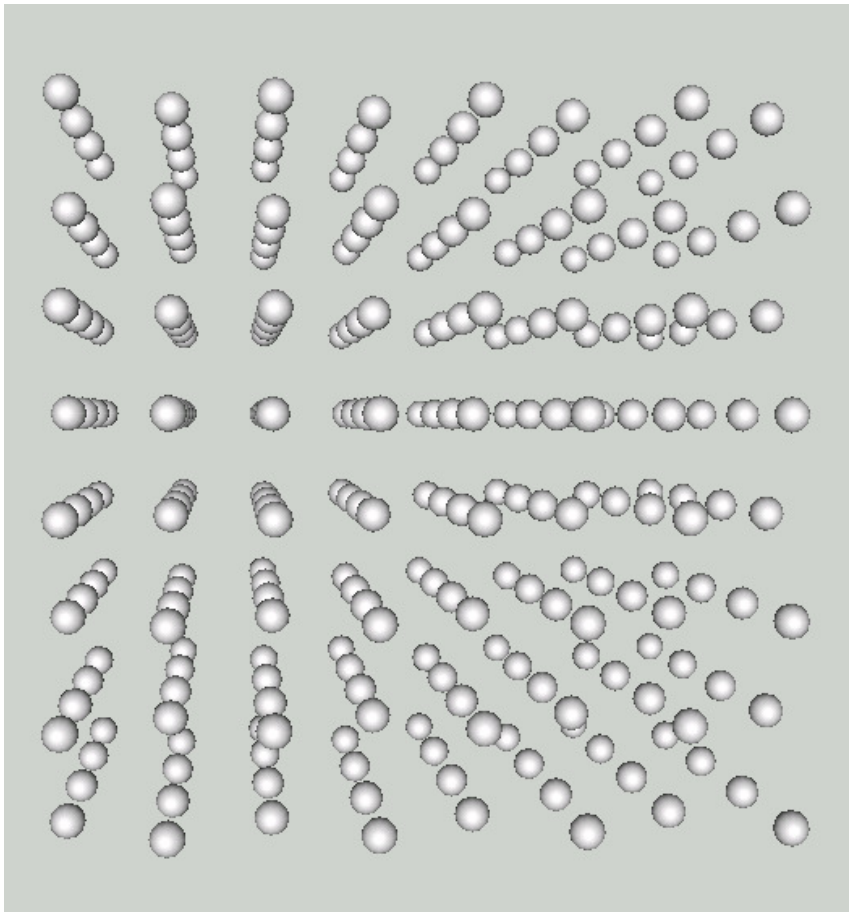


Ubiquitin (protein): structure? conformational changes?

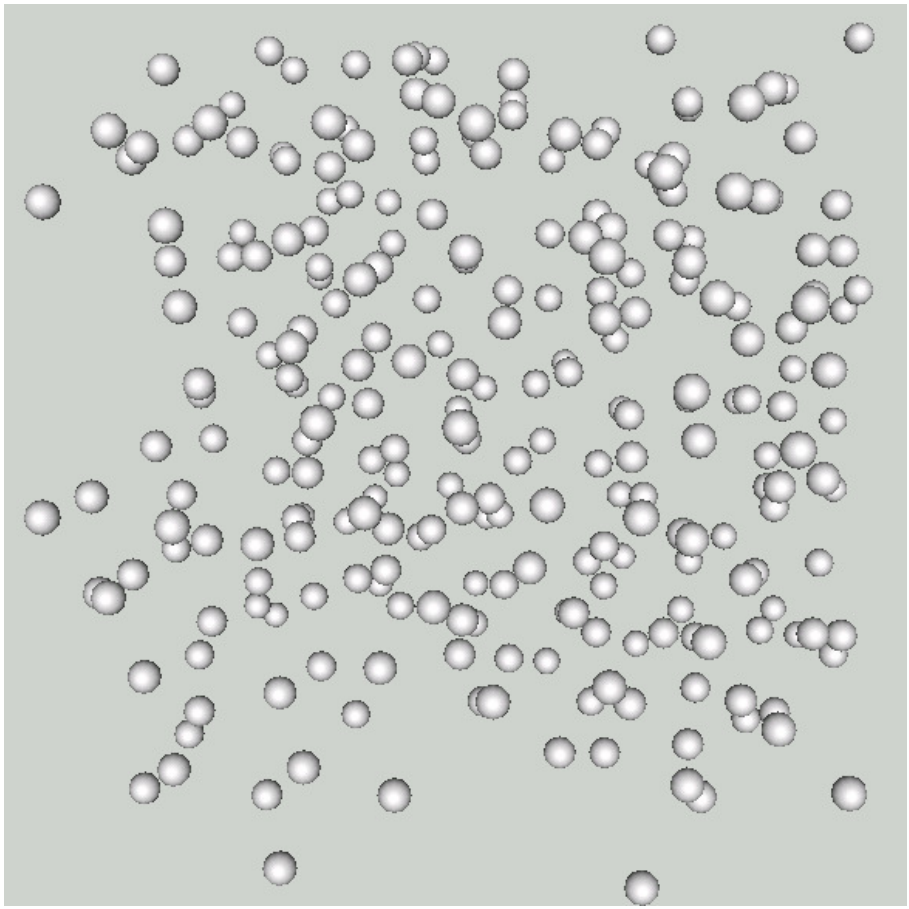
$\rightarrow$ In silico drug design

Examples of molecular systems (2)

What is the melting temperature of Argon?



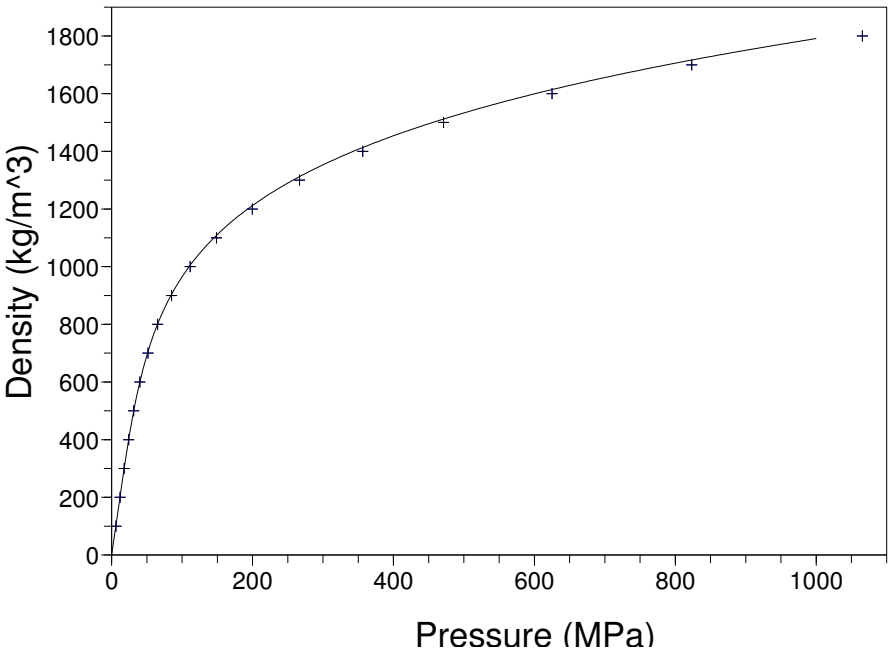
(a) Solid Argon (low temperature)

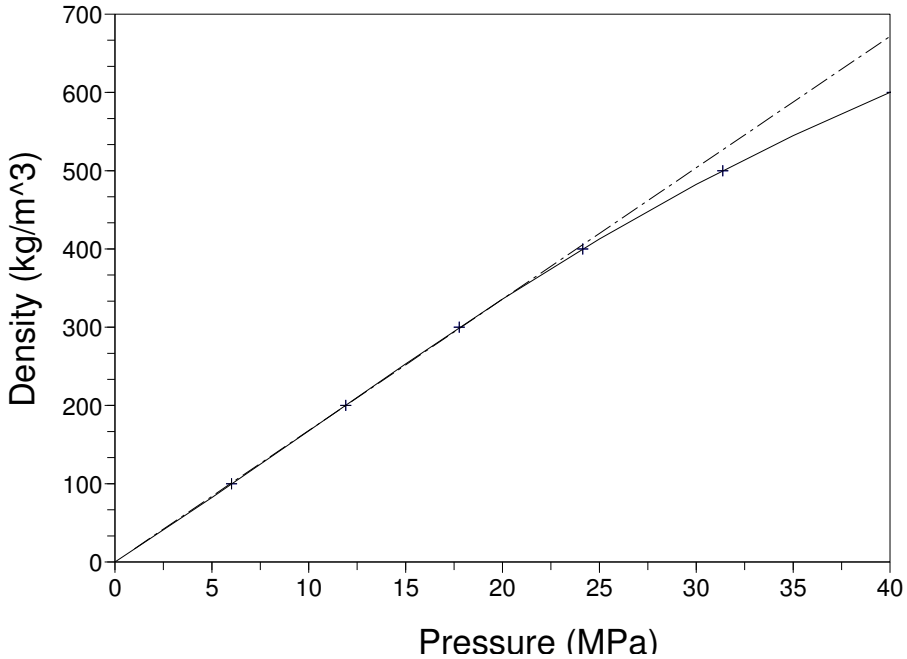


(b) Liquid Argon (high temperature)

Examples of molecular systems (3)

Equation of state of Argon: density as a function of pressure, $T = 300$ K (comparison with data of *National Institute of Standards and Technology*)





Explore extreme conditions of matter...

Some orders of magnitude...

- **Physical quantities**
  - distances ~ 1 Å $= 10^{-10}$ m
  - energy per particle $\sim k_{\mathrm{B}} T \sim 4 \times 10^{-21}$ J at 300 K
  - atomic masses $\sim 10^{-26}$ kg
  - typical times $\sim 10^{-15}$ s
  - number of particles $\sim \mathcal{N}_A = 6.02 \times 10^{23}$

- **"Standard" simulations**
  - 10^6 particles ["heroic": 10^9 particles and more]
  - total time: (fraction of) ns ["heroic": (fraction of) μs]

- **Analogy** to understand what such large numbers represent...
  - about 10^{22} moles of water on Earth
  - 10^6 moles of water ~ 1 m^3

Aims of computational statistical physics

- **"Numerical microscope"**
  - gaining some insight into physical mechanisms at the atomic scale
  - From the press release for the Nobel prize in Chemistry 2013 (Karplus/Levitt/Warshel)

    Today the computer is just as important a tool for chemists as the test tube. Simulations are so realistic that they predict the outcome of traditional experiments.

- **Computation of average properties** (static)

$$\langle A \rangle = \int_{\mathcal{E}} A(q,p)\,\mu(dq\,dp)$$

Computation of high dimensional integrals

- μ is a probability measure
- A is the observable

High performance computations

- About 40 % of running time in scientific computing centers devoted to molecular simulation (quantum + classical)

- **Parallelization strategies...**
  - spatial decomposition to compute forces and energies (every step)
  - time is intrinsically sequential...

- **Our scientific approach relies on**
  - testing new algorithms in home-made codes (e.g. SIMOL, co-developed with Inria)
  - implementing successfuls methods in codes targeted to specific applications (NAMD for computational biology, LAMMPS and STAMP for materials science, ...)

Numerical microscope

Hamiltonian dynamics and its integration (1)

- Integrate the dynamics of the system, and vizualize it...

Hamiltonian dynamics: fixed energy $H(q,p) = E$

$$\begin{cases} dq_t = M^{-1} p_t \, dt, \\ dp_t = -\nabla V(q_t) \, dt \end{cases}$$

- Many qualitative properties (time-reversibility, volume preservation, ...)
- Numerical integration
  - standard theory: fixed integration time, $\Delta t \to 0$
  - longtime integration with $\Delta t > 0$ fixed requires a dedicated treatment

Hamiltonian dynamics and its integration (2)

- Explicit Euler scheme $\begin{cases} q^{n+1} = q^n + \Delta t\, M^{-1} p^n \\ p^{n+1} = p^n - \Delta t\, \nabla V(q^n) \end{cases}$

- Symplectic Euler scheme $\begin{cases} q^{n+1} = q^n + \Delta t\, M^{-1} p^n \\ p^{n+1} = p^n - \Delta t\, \nabla V(q^{n+1}) \end{cases}$

- Mathematical understanding through backward numerical analysis

  Interpret the numerical solution of the exact problem as the exact solution of a modified problem

Here, for symplectic Euler: the numerical trajectory lies on the orbits of a modified Hamiltonian dynamics

$$H_{\Delta t}(q,p) = H(q,p) + \frac{\Delta t}{2} p^T M^{-1} \nabla V(q) + \mathrm{O}(\Delta t^2)$$

Computation of average properties (static)

Thermodynamic properties

- Canonical measure (Boltzmann–Gibbs): fixed volume and temperature T

$$\mu(dq\,dp) = Z^{-1}\,\mathrm{e}^{-H(q,p)/(k_{\mathrm{B}}T)}\,dq\,dp, \qquad Z = \int_{\mathcal{E}} \mathrm{e}^{-H/(k_{\mathrm{B}}T)}$$

Low energy states are more and more likely with lower temperatures

Langevin dynamics (friction $\gamma > 0$)

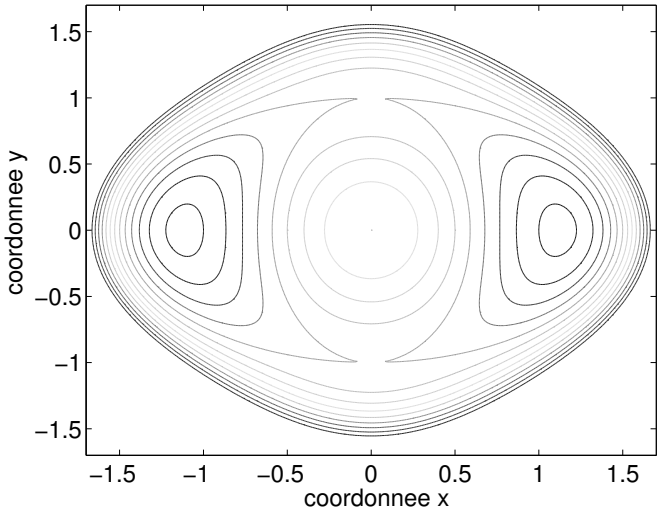
$$\begin{cases} dq_t = M^{-1} p_t\,dt, \\ dp_t = -\nabla V(q_t)\,dt - \gamma M^{-1} p_t\,dt + \sqrt{2\gamma k_{\mathrm{B}} T}\,dW_t \end{cases}$$

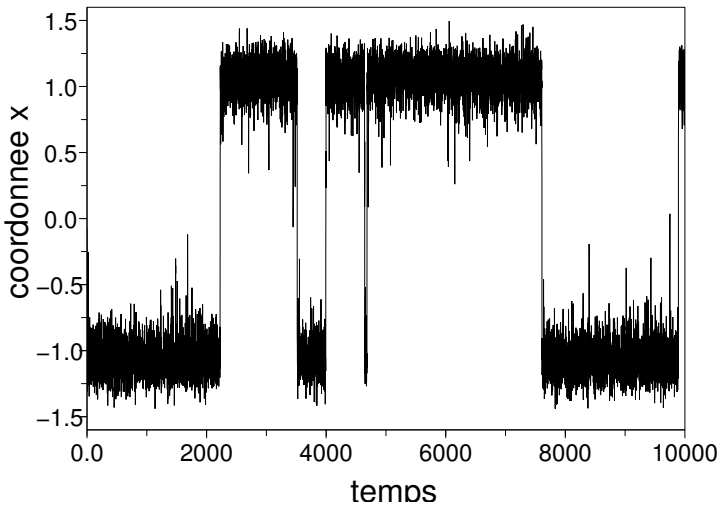
- Ergodic averages $\dfrac{1}{t}\displaystyle\int_0^t A(q_s, p_s)\,ds \xrightarrow[t\to+\infty]{} \langle A \rangle = \int_{\mathcal{E}} A\,d\mu$

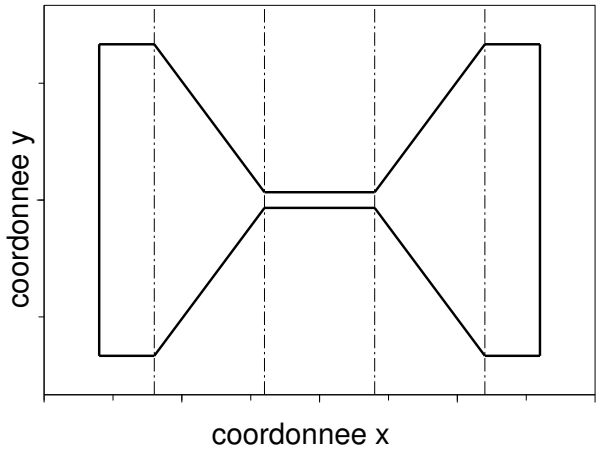
- Discretization of the stochastic differential equation $\to$ stability, bias, ...

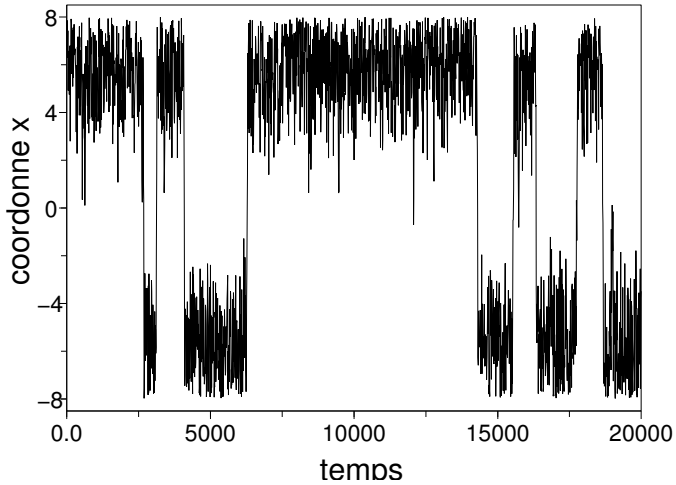
Emphasis here as well on **longtime properties** (invariant measure)

The actual mathematical challenge: metastability









Energetic vs. entropic barriers $\rightarrow$ Variance reduction techniques, e.g. modifying $-\nabla V$

Computation of dynamical properties

Transport coefficients

• Nonequilibrium driving mimicking macroscopic experimental setups, for instance additional non-gradient force

Langevin dynamics with $q \in \mathbb{T}^d$ (magnitude of the forcing η)

$$\begin{cases} dq_t = M^{-1} p_t \, dt, \\ dp_t = \left(-\nabla V(q_t) + \eta F\right) dt - \gamma M^{-1} p_t \, dt + \sqrt{2\gamma k_{\mathrm{B}} T} \, dW_t \end{cases}$$

• Linear response of an appropriate flux, here mobility α

$$\alpha = \lim_{\eta \to} \frac{\mathbb{E}_\eta(F^T M^{-1} p)}{\eta} = \frac{1}{k_{\mathrm{B}} T} \int_0^{+\infty} \mathbb{E}_0 \left[\left(F^T M^{-1} p_t \right) \left(F^T M^{-1} p_0 \right) \right] dt$$

• Variance reduction difficult since invariant measure is not known... and depends non-trivially on the dynamics!

Reference for all this talk: T. Lelièvre and G. Stoltz, Partial differential equations and stochastic methods in molecular dynamics, *Acta Numerica* **25**, 681-880 (2016)