

- 
- 
- 

# Dynamique des protéines, simulation moléculaire et physique statistique

Gerald R. Kneller

`kneller@llb.saclay cea.fr`, `kneller@cnrs-orleans.fr`

Université d'Orléans

Laboratoire Léon Brillouin, CEA Saclay

Centre de Biophysique Moléculaire, CNRS

# Intérêts de l'équipe

- Etude et modélisation de la relation dynamique-fonction des protéines
- Développement d'algorithmes de simulation (MD)
- Développement de logiciels pour la simulation moléculaire et leur analyse
- Spectroscopie neutronique, et autres techniques spectroscopiques (FCS, RMN,...)

# Méthodes

- Simulations de dynamique moléculaire, calcul de modes normaux
- Traitement numérique du signal
- Spectroscopie neutronique (+ FCS + RMN)
- Physique statistique hors équilibre

# Thématique actuelle

- ***Dynamique d'une protéine sous pression***
  - Thèse terminée sur la dynamique du lysozyme sous pression (V. Hamon, CBM Orléans)
  - Dynamique interne de protéines mésophiles et barophiles/-phobes sous l'influence de la pression - recherche d'une différence systématique (P. Calligari, LLB/ILL Grenoble)
- ***Développement d'une plate-forme de simulation à l'Institut Laue-Langevin***

# Thématiques à l'horizon

- ***Dynamique lente et fonction des protéines***
  - Etude par spectroscopie FCS / microscopie confocale (SOLEIL)
  - Développement de nouvelles méthodes stochastiques et quantiques de simulation (avec F. James, P.E. Jabin).
- ***Physico-chimie de la cellule***
  - Transport de particules virales
  - Polymérisation dans le cytosquelette

*Phase exploratoire avec l'Institut Pasteur*

- 
- 
- 

# Simulation MD

# MD classique

## Equations de Newton

$$m_{\alpha} \ddot{\mathbf{x}}_{\alpha} = \mathbf{F}_{\alpha}, \quad \alpha = 1, \dots, N.$$

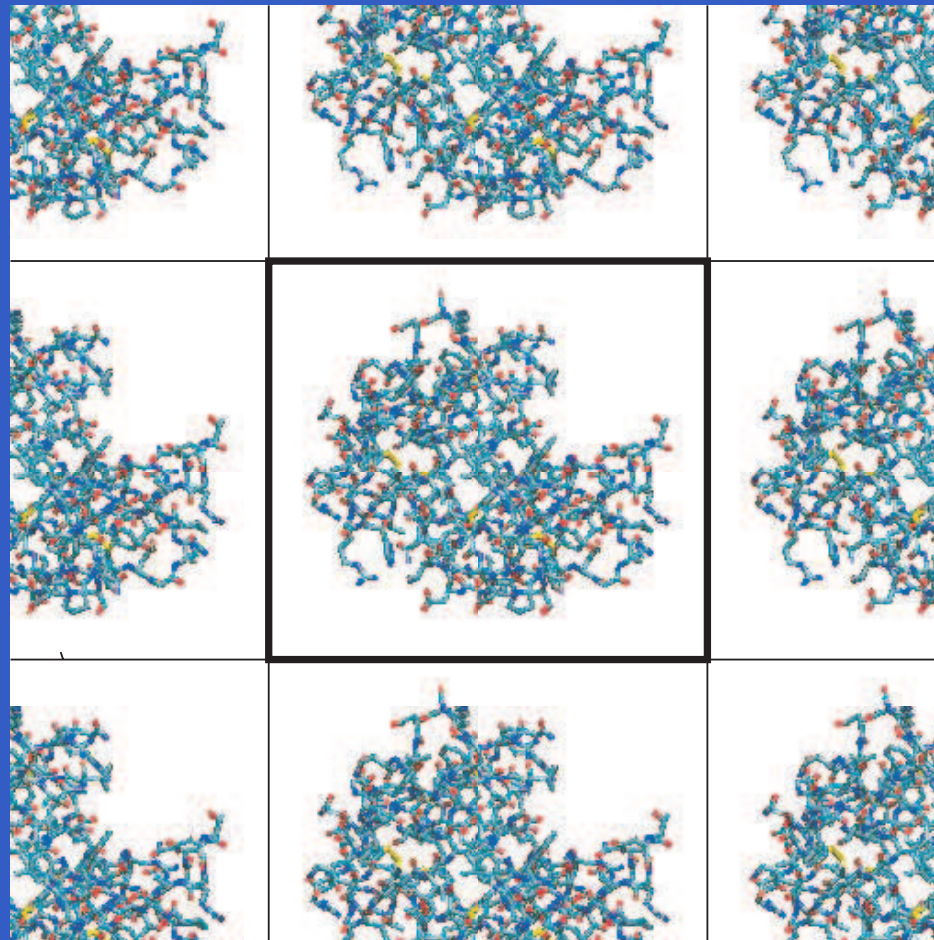
## Forces et potentiel / champ de force

$$\mathbf{F}_{\alpha} = - \frac{\partial U(\mathbf{x}_1, \dots, \mathbf{x}_N)}{\partial \mathbf{x}_{\alpha}}$$

- Chaque atome est une masse ponctuelle
- Potentiel effectif – pas d'électrons explicites
- Dynamique newtonienne des atomes

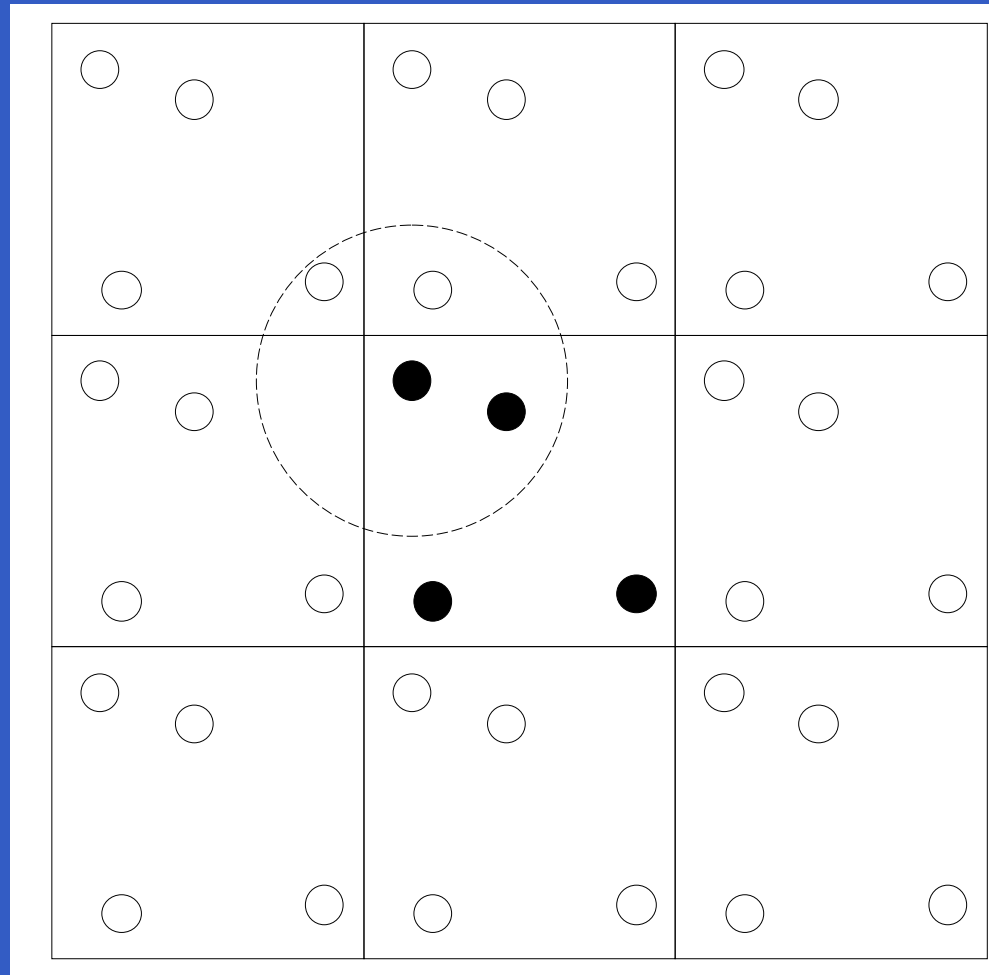
- 
- 
- 

# Conditions périodiques

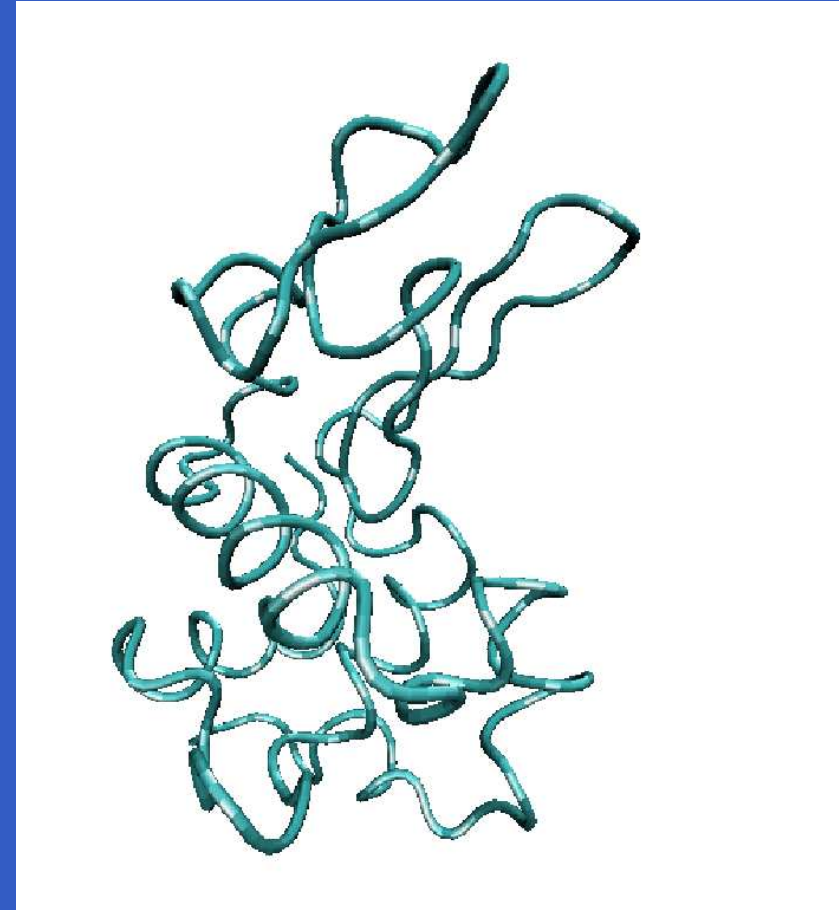
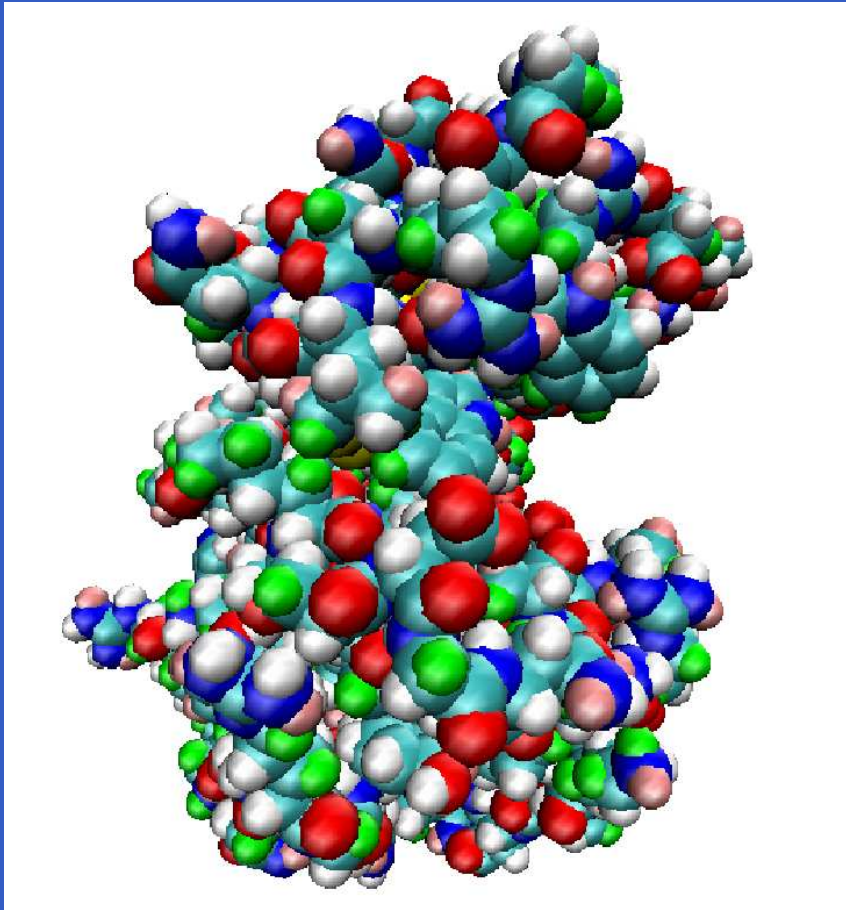




# Condition de "images minimale"

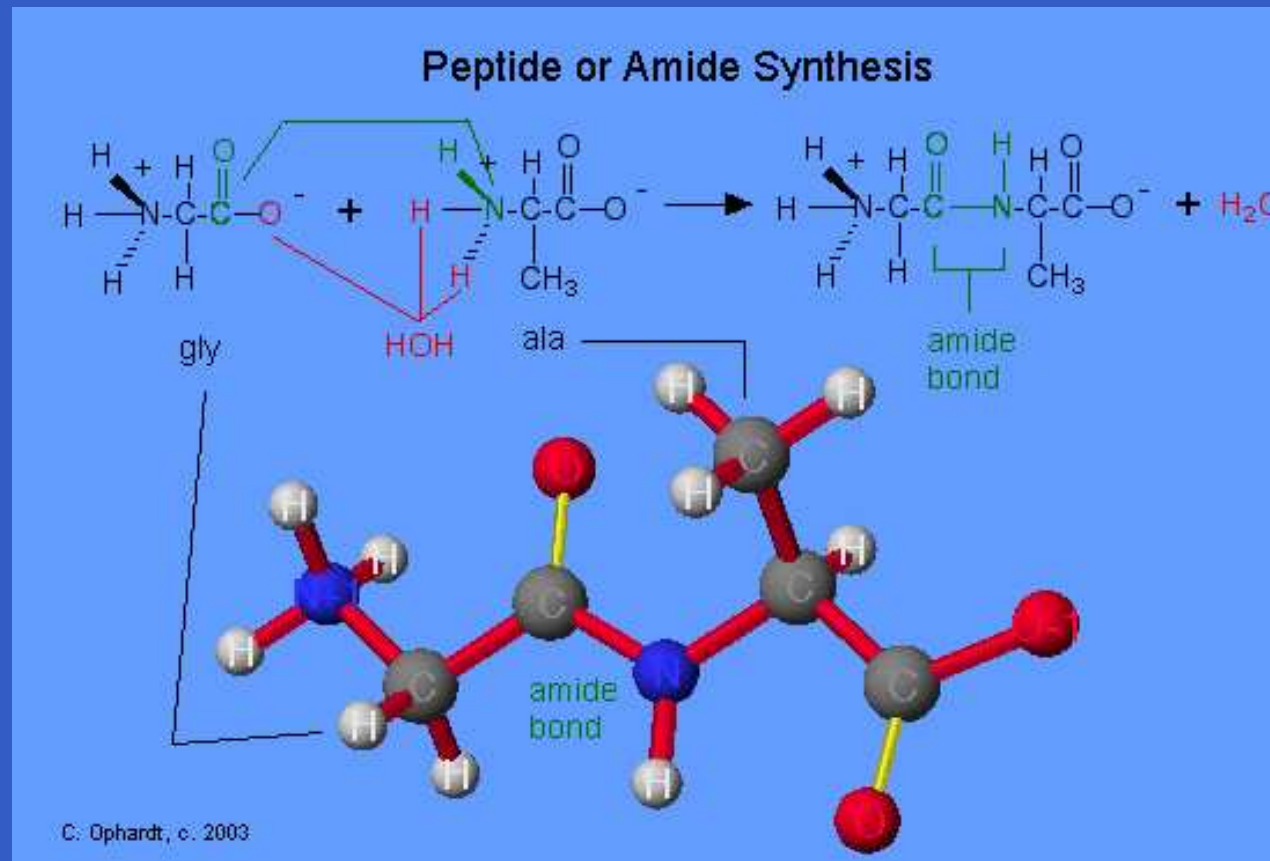


# Lyszoyne



A gauche : *Représentation atomique*. A droite : *squelette*.

# Chaînes polypeptidiques



# Champ de force (AMBER 94)

$$\begin{aligned}
 U = & \sum_{\text{liaisons } \alpha\beta} k_{\alpha\beta} \left( r_{\alpha\beta} - r_{\alpha\beta}^{(0)} \right)^2 \\
 & + \sum_{\text{angles } \alpha\beta\gamma} k_{\alpha\beta\gamma} \left( \phi_{\alpha\beta\gamma} - \phi_{\alpha\beta\gamma}^{(0)} \right)^2 \\
 & + \sum_{\text{dihèdres } \alpha\beta\gamma\delta} k_{\alpha\beta\gamma\delta} \cos \left( n_{\alpha\beta\gamma\delta} \theta_{\alpha\beta\gamma\delta} - \delta_{\alpha\beta\gamma\delta} \right) \\
 & + \left. \begin{aligned}
 & \sum_{\text{paires } \alpha\beta} 4\epsilon_{\alpha\beta} \left( \left[ \frac{\sigma_{\alpha\beta}}{r_{\alpha\beta}} \right]^{12} - \left[ \frac{\sigma_{\alpha\beta}}{r_{\alpha\beta}} \right]^6 \right) \\
 & + \sum_{\text{paires } \alpha\beta} \frac{q_{\alpha}q_{\beta}}{4\pi\epsilon_0 r_{\alpha\beta}}
 \end{aligned} \right\} \begin{array}{l} \text{non-} \\ \text{liées} \end{array}
 \end{aligned}$$

# Discrétisation

## Vitesse et accélération

$$\dot{\mathbf{x}}_{\alpha} \approx \frac{\mathbf{x}_{\alpha}(n+1) - \mathbf{x}_{\alpha}(n-1)}{2\Delta t},$$

$$\ddot{\mathbf{x}}_{\alpha} \approx \frac{\mathbf{x}_{\alpha}(n+1) - 2\mathbf{x}_{\alpha}(n) + \mathbf{x}_{\alpha}(n-1)}{\Delta t^2}.$$

## Algorithme de Verlet

$$\mathbf{x}_{\alpha}(n+1) = 2\mathbf{x}_{\alpha}(n) - \mathbf{x}_{\alpha}(n-1) + \frac{\Delta t^2}{m_{\alpha}} \mathbf{F}_{\alpha}(n).$$

# Caractéristiques du calcul numérique

- Nombre d'atomes  $N^2 \approx 10^4 \rightarrow 10^5$
- Nombre d'interactions  $\propto N^2$  (interactions non-liées)
- Nombre de pas d'intégration  $\approx 10^6 - 10^7$
- Adapté pour calcul parallèle :
  - décomposition en domaines
  - ou décomposition en sous-ensembles fixes
  - en tout cas : communication forte

# Echelles

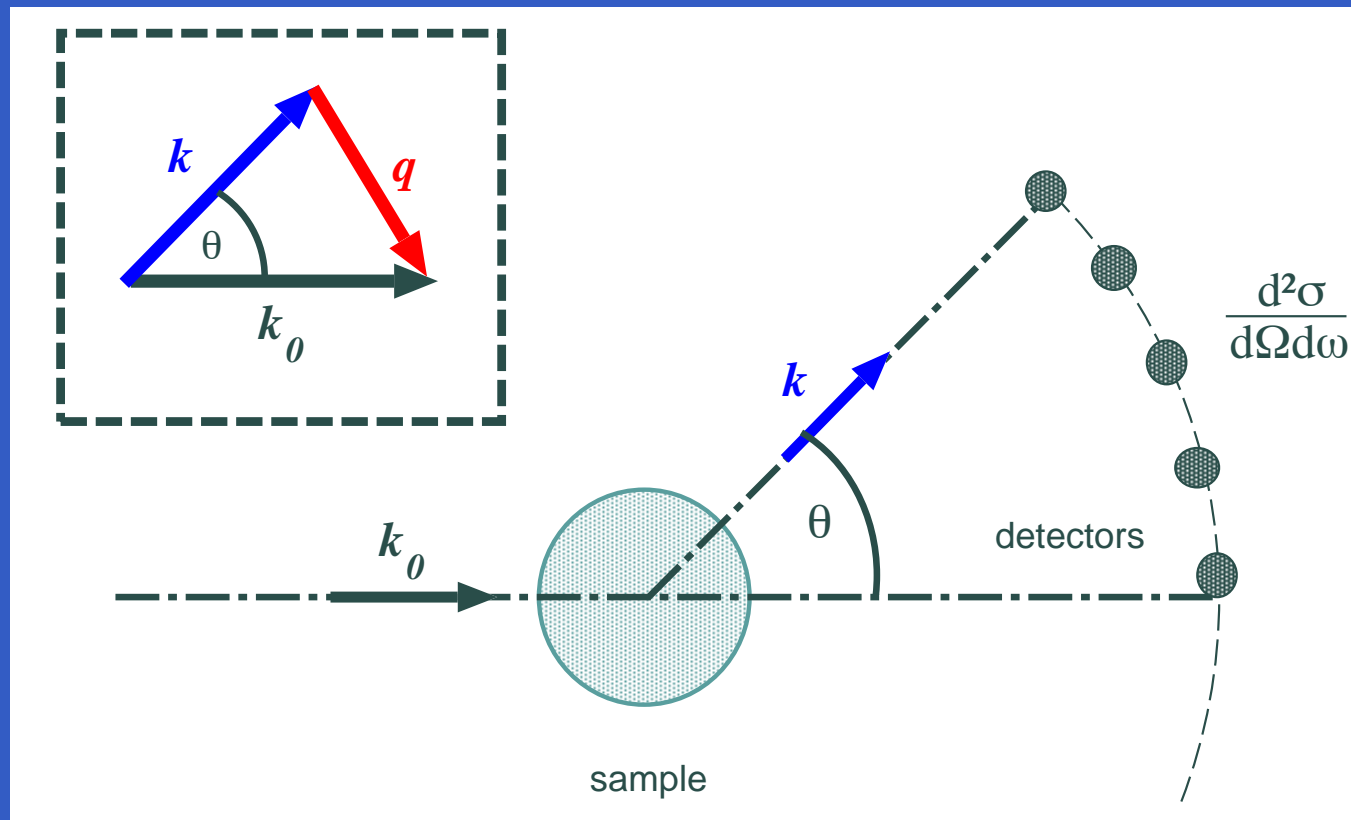
- Echelle spatiale :  $0.1 \text{ nm} \rightarrow \approx 10 \text{ nm}$
- Echelle de temps :  $0.1 \text{ ps} \rightarrow \approx 10 \text{ ns}$
- Comparables aux échelles accessibles par diffusion de neutrons thermiques
- Comparaison facilitée par l'interaction directe des neutrons avec des noyaux atomiques = objets simulés



# Diffusion de neutrons



# Expérience schématique



# Section efficace différentielle

$$\frac{d^2\sigma}{d\Omega d\omega} = \frac{|\mathbf{k}|}{|\mathbf{k}_0|} \mathcal{S}(\mathbf{q}, \omega)$$

Ici

$$\mathbf{q} = \mathbf{k}_0 - \mathbf{k}, \quad \omega = \frac{E_0 - E}{\hbar}$$

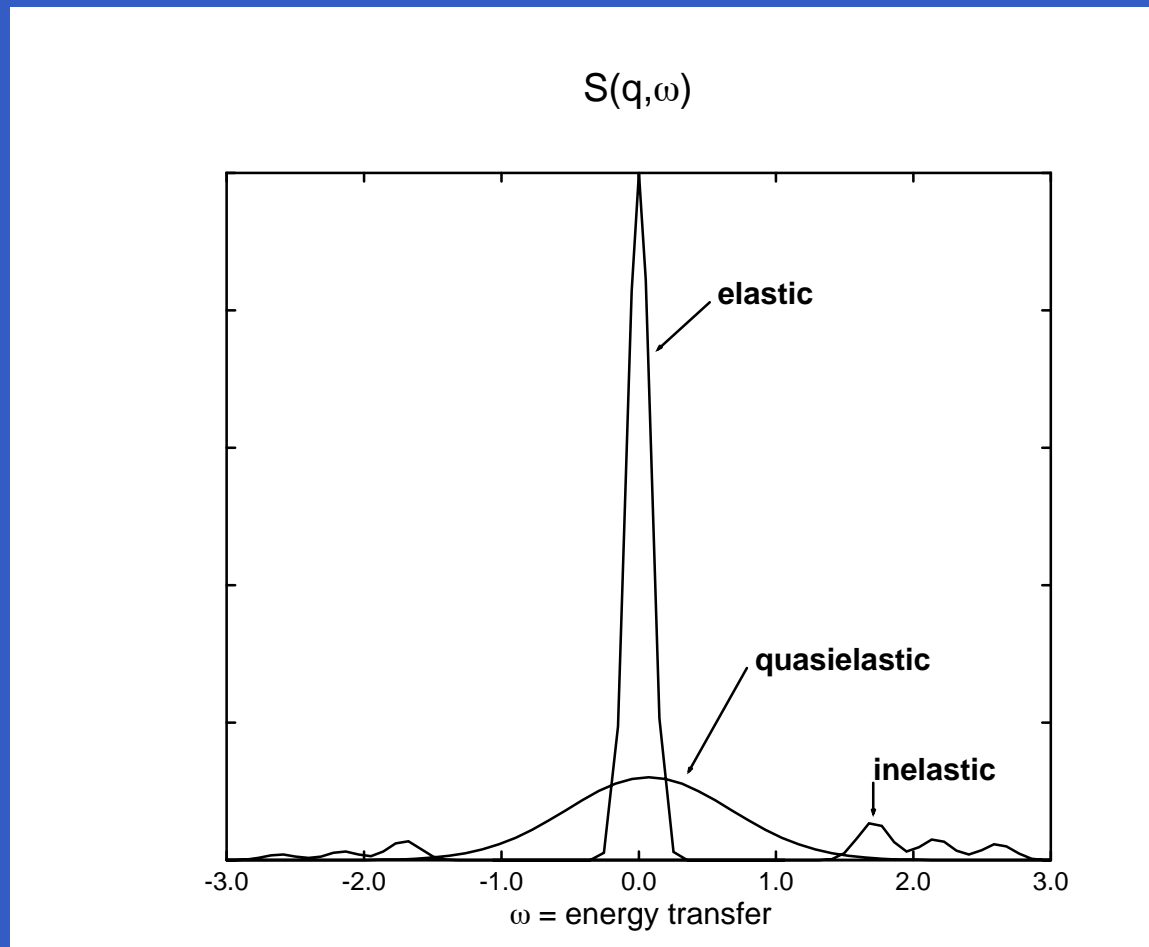
sont les transferts de **quantité de mouvement** et de **l'énergie**, respectivement, et  $\mathcal{S}(\mathbf{q}, \omega)$  est le **facteur de structure dynamique**.

# Facteur de structure dynamique

$$\mathcal{S}(\mathbf{q}, \omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dt \exp(-i\omega t) \phi(\mathbf{q}, t),$$
$$\phi(\mathbf{q}, t) = \sum_{\alpha, \beta} \overline{b_{\alpha} b_{\beta}} \left\langle e^{i\mathbf{q}^T \cdot \mathbf{R}_{\beta}(t)} e^{-i\mathbf{q}^T \cdot \mathbf{R}_{\alpha}(0)} \right\rangle.$$

$\overline{b_{\alpha} b_{\beta}}$  = moyenne sur isotopes et orientations relatives entre le spin du neutron et celui de l'atome cible.  $\phi(\mathbf{q}, t)$  est la *fonction intermédiaire de diffusion*.

# Spectre neutronique



- 
- 
- 

# Analyse des simulations

# Analyses simples

## Fonction de corrélation – variables continues

$$c_{AB}(t) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} d\tau A(\tau + t) B^*(\tau).$$

## Fonction de corrélation – variables discrètes

$$c_{AB}(n) = \frac{1}{N_t - |n|} \sum_{k=0}^{N_t - |n| - 1} A(k + n) B^*(k).$$

Utilisation intensive de FFT.

# Analyses sophistiquées

Modélisation d'une trajectoire MD par un processus stochastique *autoregressif* de l'ordre  $P$ ,

$$u(t) = \sum_{n=1}^P a_n^{(P)} u(t - n\Delta t) + \epsilon_P(t).$$

$\epsilon_P(t)$  est un *bruit blanc* de l'amplitude  $\sigma_P$ . Les coefficients  $\{a_n^{(P)}, \sigma_P\}$  sont ajustés aux données (algorithme de Burg).

# Modèle semi-analytique

Modèle analytique dans le plan complexe :

$$U(z) = \sum_{n=-\infty}^{+\infty} u(n)z^{-n} = \frac{\sigma_P}{1 - \sum_{j=1}^P a_j^{(P)} z^{-j}}$$

Fonction d'autocorrélation :

$$C(z) = \frac{\sigma_P^2}{\left(1 - \sum_{j=1}^P a_j^{(P)} z^{-j}\right) \left(1 - \sum_{l=1}^P a_l^{(P)} z^l\right)}.$$



# Racines caractéristiques

- All-pole form of  $C(z)$  :

$$C(z) = \frac{1}{a_P^{(P)}} \frac{-z^P \sigma_P^2}{\prod_{k=1}^P (z - z_k) \prod_{l=1}^P (z - z_l^{-1})}.$$

- Les  $\{z_k\}$  sont les racines de

$$p(z) = z^P - \sum_{k=1}^P a_k^{(P)} z^{P-k}.$$

# Estimation directe de spectres

- Spectre de puissance par  $C(z)$  :

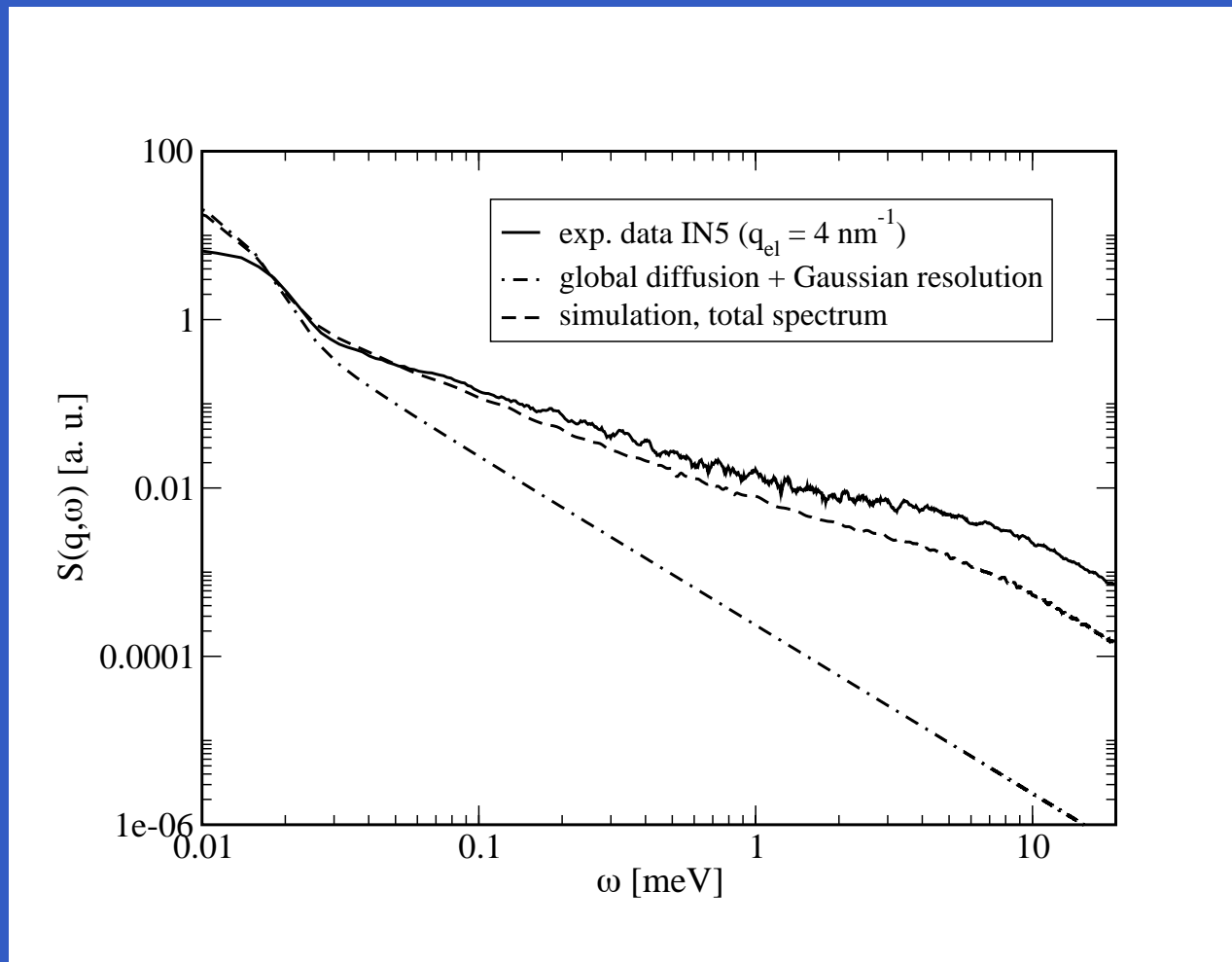
$$\tilde{c}(\omega) = C(\exp[i\omega\Delta t]).$$

- Fonction de corrélation multiexponentielle :

$$C(n) = \frac{1}{2\pi i} \oint_{\eta} dz z^{n-1} C(z) = \sum_{j=1}^P \beta_j z_j^{|n|}.$$

Here  $\beta_j \equiv \beta_j(\{z_k\}) = \text{const.}$ ,  $|z_j| < 1$ , and  
 $\eta : |z| > z_j, \forall j$ .

# Simulation / Expérience



# Caractéristiques du calcul numérique

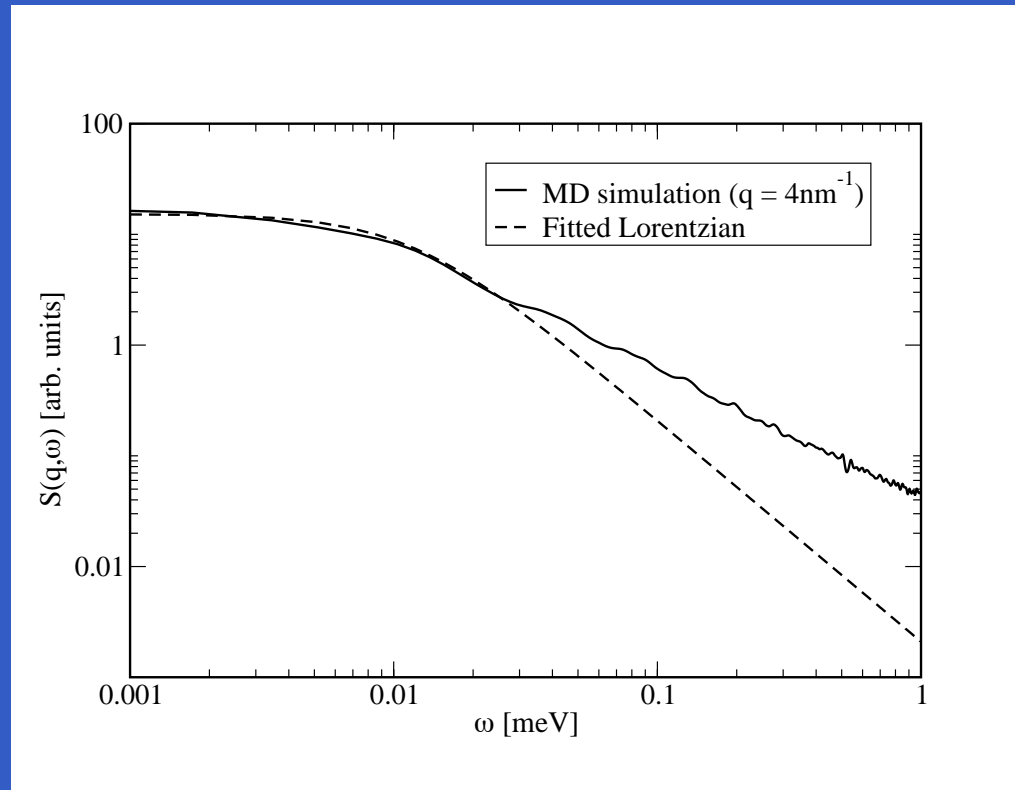
- Utilisation extensive de FFT
- Utilisation extensive de l'algorithme de Burg pour estimer les paramètres du modèle AR. Efficacité comparable à la FFT.
- Parallélisation triviale
  - Les fonctions de corrélation pour chaque atome peuvent être calculées indépendamment.
  - Idem pour différents vecteurs  $\mathbf{q}$  dans  $S(\mathbf{q}, \omega)$ .
  - $\implies$  communication  $\approx 0$ .



# Développement de modèles

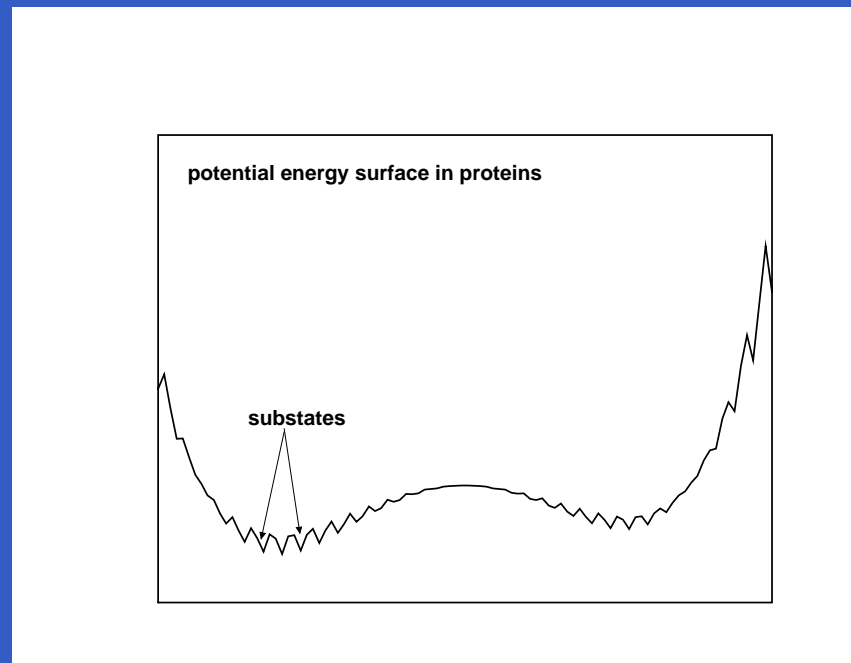


# Relaxation non-exponentielle



$S(q, \omega)$  simulé pour la dynamique interne du lysozyme et lorentzienne ajustée.

# Energie interne d'une protéine



- Surface rugeuse de  $E_{pot}$  d'après Frauenfelder (“sous-états conformationnels”).
- Beaucoup de degrés de liberté couplés, avec des échelles de temps très différentes.

# Fonction mémoire (Zwanzig, 1961)

- Equation de la fonction mémoire

$$\partial_t c(t) = - \int_0^t d\tau \xi(t - \tau) c(\tau)$$

- Dynamique Brownienne

$$\xi(t) = \gamma \delta(t) \quad \Longrightarrow \quad c(t) = c(0) \exp(-\gamma t).$$



# Dynamique brownienne fractionnaire

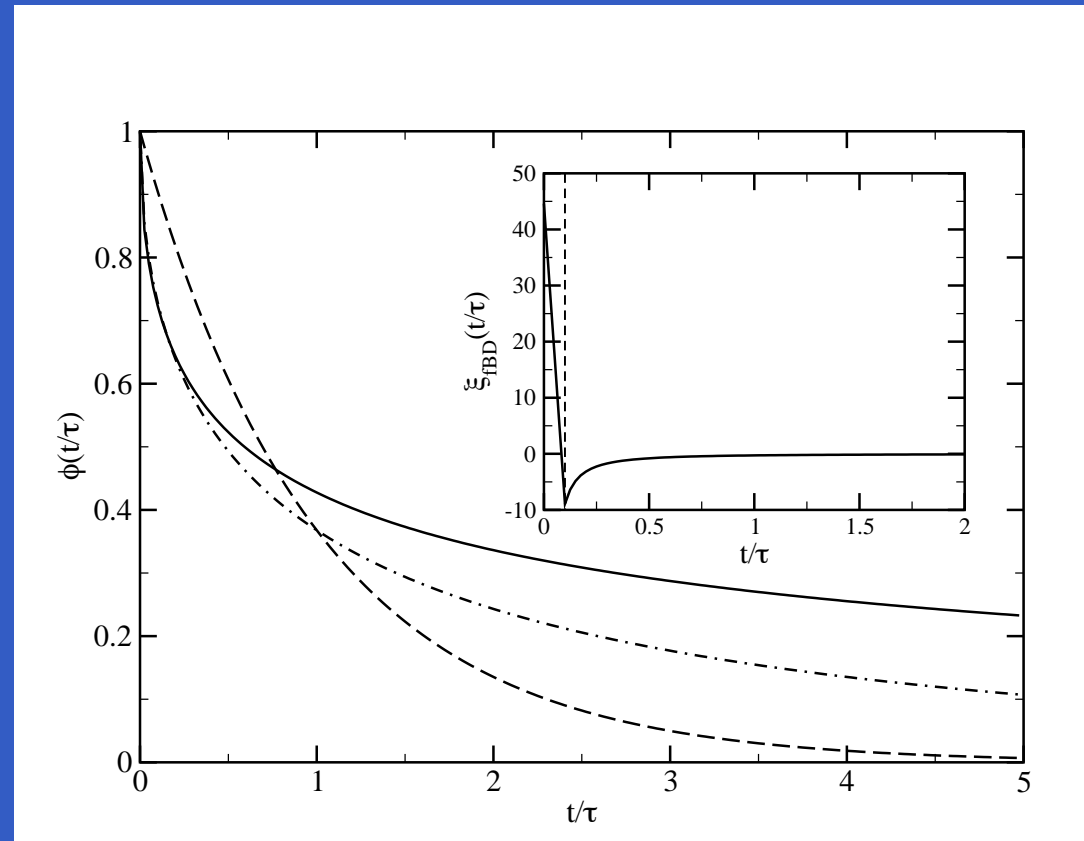
## Fonction de corrélation

$$c(t) = E_{\beta}(-(t/\tau)^{\beta}), \quad 0 < \beta \leq 1.$$

## Fonction de Mittag-Leffler

$$E_{\beta}(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(1 + \beta k)}.$$

# Fonction de corrélation



*Modèle DBF ( $\beta = 1/2$ ) = ligne continue, exponentielle étirée  
( $\beta = 1/2$ )=points et traits, exponentielle = traits.*

# Fonction mémoire et spectre

- **Fonction mémoire** ( $0 < \beta < 1$ )

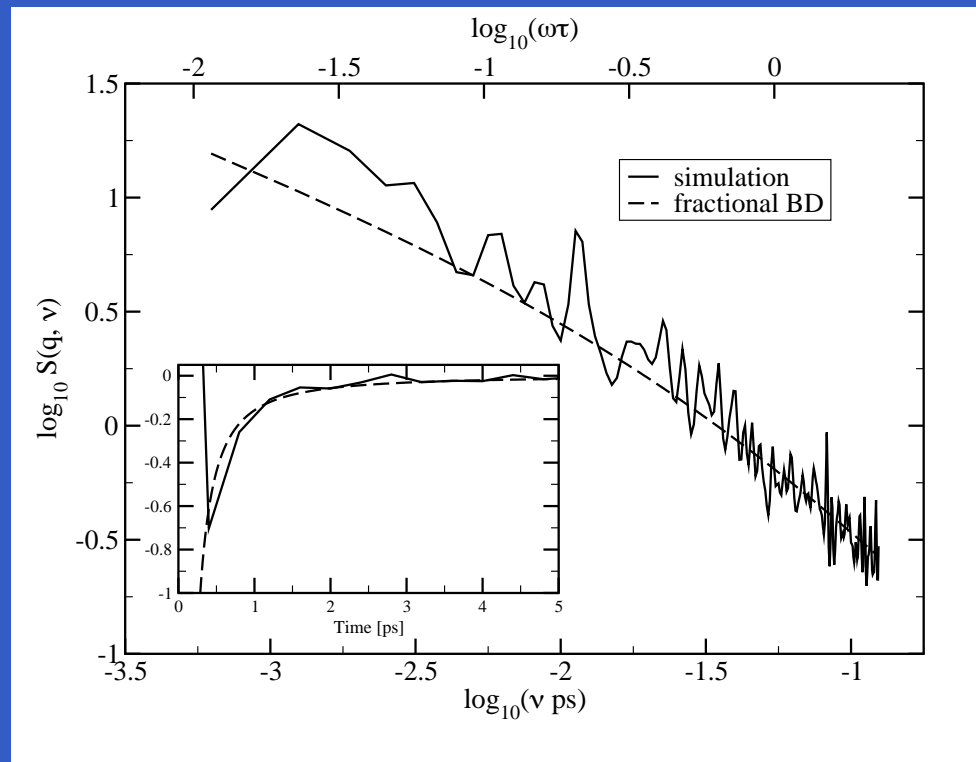
$$\xi(t) \propto (\beta - 1) \left( \frac{t}{\tau} \right)^{\beta-2}, \quad \int_0^{\infty} dt \xi(t) = 0.$$

- **Spectre de  $c(t)$**

$$S_{fBD}(\omega) = \frac{2\tau \sin(\beta\pi/2)}{|\omega\tau| (|\omega\tau|^\beta + 2 \cos(\beta\pi/2) + |\omega\tau|^{-\beta})}.$$

pour  $0 < \beta \leq 1$ . Lorentzienne pour  $\beta = 1$ .

# Modèle et simulation



Facteur de structure dynamique et fonction mémoire pour  
 $Lys$  à  $q = 10 \text{ nm}^{-1}$  par MD et modèle DBF.

- 
- 
- 

# Nos codes

# Molecular Modeling Toolkit

- Bibliothèque de simulation moléculaire
- Langages : Python et C
- Orienté objet
- Techniques implementées : Dynamique Moléculaire, Modes Normaux, Minimization d'énergie, etc.
- Parallélisé
- Développé depuis 1996 par K. Hinsen

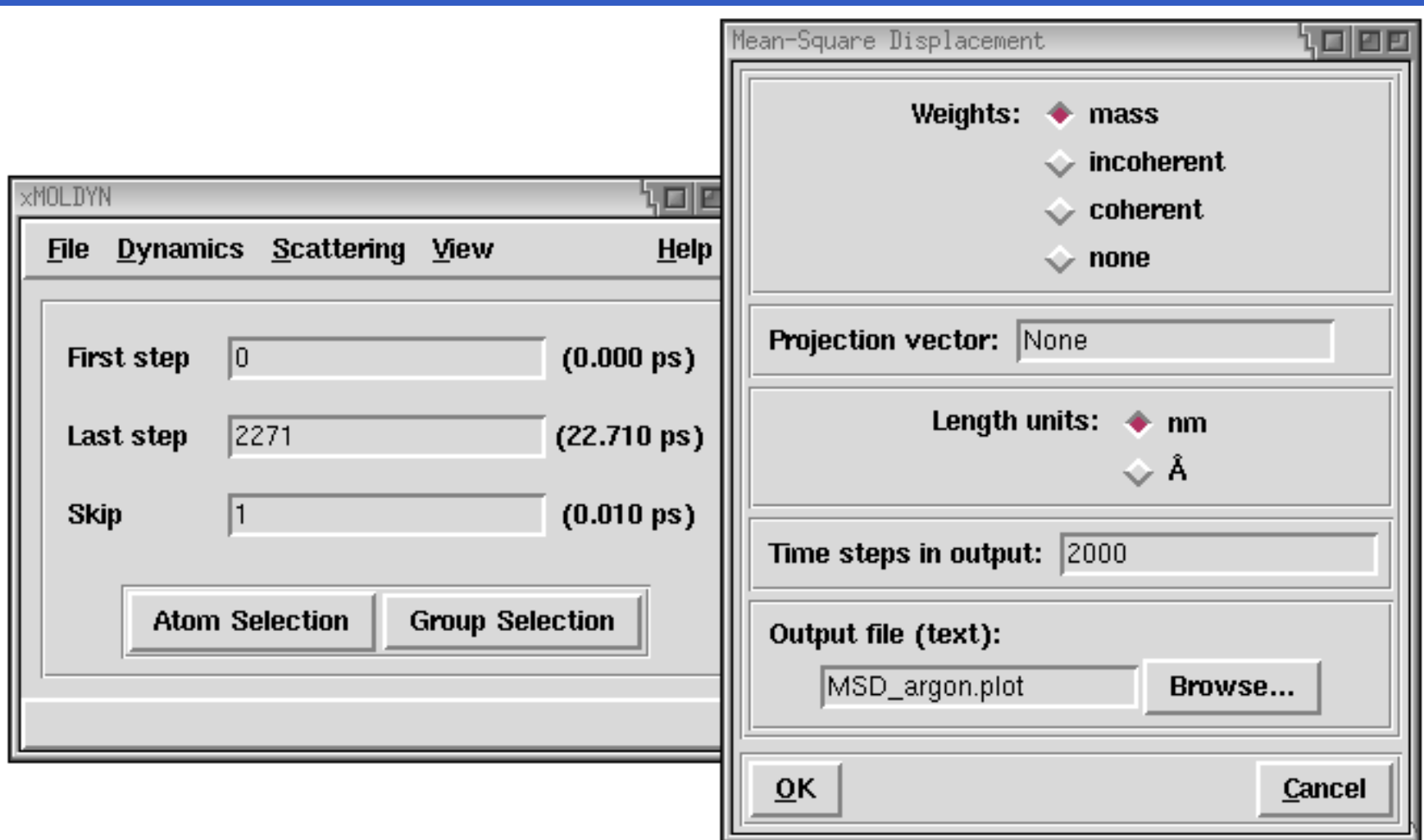
<http://dirac.cnrs-orleans.fr/MMTK/>

# nMOLDYN

- Analyse de trajectoires de Dynamique Moléculaire, en particulier pour comparer avec la diffusion de neutrons
- Interface graphique
- Basé sur MMTK
- Langage : 100% Python
- Développé depuis 1991 par G. Kneller et collaborateurs

**<http://dirac.cnrs-orleans.fr/nMOLDYN/>**

# nMOLDYN





- 
- 
- 

# PPF et équipement

# Nos intérêts

- Mise à disposition de nos codes par une interface Web. Prévoir échange de volumes de données importants (trajectoires MD).
- Les analyses par nMoldyn sont idéales pour un calcul distribué.

# Equipement

- Saclay : 8 noeuds Opteron / partie d'une grappe de 64 noeuds du DRECAM / CEA Saclay.
- Orléans : Grappe de 20 PC bi-proc Pentium III "en fin de vie".
- + PCs de bureau.

# Remerciements

- K. Hinsén, CBM Orléans → LLB Saclay
- M.C. Bellissent-Funel, LLB Saclay Saclay
- V. Hamon, CBM Orléans
- G. Sutmann, FZ Jülich (D)
- T. Rog & K. Murzyn, Krakow (PL)