

Simulations moléculaires et leur analyse

G.R. Kneller^{1,2}

¹ *Laboratoire Léon Brillouin, CEA Saclay, 91191 Gif-sur-Yvette, France*

² *Centre de Biophysique Moléculaire, CNRS, rue Charles Sadron, 45071 Orléans, France*

Résumé. Ce cours explique les concepts de la simulation moléculaire et son utilisation pour l'étude de la dynamique des macromolécules biologiques, telle qu'elle est étudiée par diffusion de neutrons. La complémentarité des deux techniques permet de comprendre des détails de la dynamique d'un système moléculaire complexe qui sont inaccessibles aux expériences seules. On peut ainsi mieux interpréter les données expérimentales et également développer des modèles physiques qui donnent une vue cohérente des observations. Ce dernier point est illustré par une étude du lysozyme en solution. En utilisant des méthodes provenant du traitement numérique du signal, on peut extraire des spectres de temps de relaxation qui illustrent que la dynamique des protéines est caractérisée par des processus de relaxation multi-échelles.

Table des matières

1	Introduction	155
2	Simulation de dynamique moléculaire	156
2.1	Le principe	156
2.2	Conditions périodiques	157
2.3	Ensemble microcanonique	158
2.4	Ensembles thermodynamiques non-standard	160
3	Analyse simple de simulations	161
3.1	Fonction de corrélation des vitesses et son spectre de Fourier	161
3.2	Déplacement carré moyen	163
3.3	Fonctions de diffusion intermédiaires et facteurs de structure dynamiques	165
3.4	Approximation gaussienne	166
3.5	EISF	168
3.6	Limite classique	169
4	Relaxation multi-échelles dans une protéine	171
4.1	Équation de Langevin généralisé – fonction mémoire	171
4.2	Dynamique brownienne fractionnaire	172
4.3	Approche numérique	172
4.4	Dynamique du lysozyme	175
A	Distribution de Dirac	175

1. INTRODUCTION

Il est maintenant un fait accepté que la flexibilité et la dynamique d'une protéine jouent un rôle aussi important pour son fonctionnement que sa structure. En utilisant la spectroscopie de fluorescence sur une molécule isolée on peut aujourd'hui même suivre une réaction enzymatique en temps réel [1]. La combinaison de la spectroscopie de fluorescence sur une molécule isolée avec un transfert d'électron

photo-induit permet de suivre la dynamique d'une protéine sur une vaste échelle de temps [2]. Ces expériences confirment la présence d'une gamme très large de temps de relaxation qui peuvent atteindre une milli-seconde, voire une seconde. Nous montrerons dans ce cours que les modèles théoriques pour ces observations (voir par exemple [3]) peuvent aussi être appliqués afin de décrire la dynamique d'une protéine à l'échelle de la nano-seconde, tel qu'elle est vue par diffusion de neutrons et simulation de dynamique moléculaire (MD = *Molecular Dynamics*). D'abord sera donné une introduction à la technique de simulation MD, ainsi qu'une description du calcul de certaines observables qui sont accessibles par diffusion de neutrons. La combinaison de ces deux méthodes complémentaires est une approche performante pour l'étude de la dynamique et de la structure d'un système moléculaire à l'échelle atomique. Les simulations MD ne permettent pas seulement d'étudier un système moléculaire aux mêmes échelles spatiales et temporelles que la diffusion de neutrons thermiques (environ 1 Å – 100 Å, 0.1 ps – 10 ns), mais les deux méthodes suivent également la dynamique des mêmes objets, qui sont les noyaux des atomes dans le système étudié.

2. SIMULATION DE DYNAMIQUE MOLÉCULAIRE

2.1 Le principe

Dans une simulation de dynamique moléculaire chaque atome est représenté par une masse ponctuelle qui interagit avec tous les autres atomes dans le système par un potentiel effectif. On suppose que la dynamique des atomes peut être décrite par les lois de la dynamique classique. Dans ce cas la dynamique du système est décrite par les équations de Newton,

$$m_\alpha \ddot{\mathbf{R}}_\alpha = \mathbf{F}_\alpha, \quad \alpha = 1, \dots, N \quad (2.1)$$

Ici m_α est la masse de l'atome α , \mathbf{F}_α est la force totale sur cet atome et N est le nombre d'atomes. Dans le cadre d'une simulation MD la force sur chaque atome dérive d'un potentiel:

$$\mathbf{F}_\alpha = - \frac{\partial U(\mathbf{R}_1, \dots, \mathbf{R}_N)}{\partial \mathbf{R}_\alpha} \quad (2.2)$$

Le potentiel U dépend *a priori* des positions $\mathbf{R}_1, \dots, \mathbf{R}_N$ de tous les N atomes du système. Au cours des 20 dernières années plusieurs modèles pour U ont été développés pour la simulation des macromolécules biologiques. La forme générale est toujours similaire [4–7], et on donne ci-dessous la forme du champ de force AMBER94 [7]:

$$\begin{aligned}
 U = & \sum_{\text{liaisons } \alpha\beta} k_{\alpha\beta} \left(r_{\alpha\beta} - r_{\alpha\beta}^{(0)} \right)^2 \\
 & + \sum_{\text{angles } \alpha\beta\gamma} k_{\alpha\beta\gamma} \left(\phi_{\alpha\beta\gamma} - \phi_{\alpha\beta\gamma}^{(0)} \right)^2 \\
 & + \sum_{\text{dihèdres } \alpha\beta\gamma\delta} k_{\alpha\beta\gamma\delta} \cos \left(n_{\alpha\beta\gamma\delta} \theta_{\alpha\beta\gamma\delta} - \delta_{\alpha\beta\gamma\delta} \right) \\
 & + \left. \begin{aligned} & \sum_{\text{paires } \alpha\beta} 4\epsilon_{\alpha\beta} \left(\left[\frac{\sigma_{\alpha\beta}}{r_{\alpha\beta}} \right]^{12} - \left[\frac{\sigma_{\alpha\beta}}{r_{\alpha\beta}} \right]^6 \right) \\ & + \sum_{\text{paires } \alpha\beta} \frac{q_\alpha q_\beta}{4\pi\epsilon_0 r_{\alpha\beta}} \end{aligned} \right\} \begin{array}{l} \text{non-} \\ \text{liées} \end{array} \quad (2.3)
 \end{aligned}$$

On distingue deux groupes de termes: les *interactions liées* qui sont dues aux liaisons covalentes dans la molécule et les *interactions non-liées* qui décrivent la répulsion entre atomes proches, les interactions entre deux atomes dues aux dipôles mutuellement induits, et les interactions coulombiennes.

Afin d'obtenir un schéma numérique pour l'intégration des équations de Newton, la vitesse et l'accélération de l'atome α sont souvent approchées en utilisant les différences centrées:

$$\dot{\mathbf{R}}_{\alpha} \approx \frac{\mathbf{R}_{\alpha}(n+1) - \mathbf{R}_{\alpha}(n-1)}{2\Delta t}, \quad (2.4)$$

$$\ddot{\mathbf{R}}_{\alpha} \approx \frac{\mathbf{R}_{\alpha}(n+1) - 2\mathbf{R}_{\alpha}(n) + \mathbf{R}_{\alpha}(n-1)}{\Delta t^2}. \quad (2.5)$$

Avec (2.5) les équations de Newton prennent la forme

$$\mathbf{R}_{\alpha}(n+1) = 2\mathbf{R}_{\alpha}(n) - \mathbf{R}_{\alpha}(n-1) - \frac{\Delta t^2}{m_{\alpha}} \frac{\partial U(\mathbf{R}_1, \dots, \mathbf{R}_N)}{\partial \mathbf{R}_{\alpha}} \quad (2.6)$$

Ici Δt est le pas d'intégration qui est typiquement de l'ordre d'une femtoseconde (10^{-15} s). On remarque que $\mathbf{R}_{\alpha}(n) \equiv \mathbf{R}_{\alpha}(n\Delta t)$ est la position de l'atome α à $t = n\Delta t$. L'équation (2.6) est effectivement un schéma itératif pour la solution numérique des équations de Newton. Afin d'obtenir les nouvelles positions $\mathbf{R}_{\alpha}(n+1)$ on calcule d'abord les forces $\mathbf{F}_{\alpha} = -\frac{\partial U}{\partial \mathbf{R}_{\alpha}}$ qui dépendent des positions à $t = n\Delta t$ et propage la solution selon (2.6). Cette procédure est répétée autant que nécessaire afin d'obtenir une trajectoire de la longueur T souhaitée. L'algorithme (2.6) est connu sous le nom "algorithme de Verlet" [8, 9], et malgré sa simplicité il est toujours l'algorithme le plus utilisé pour les simulations MD. Son avantage principal est qu'il est *stable*, car il ne génère pratiquement pas des erreurs systématiques si Δt reste suffisamment petit. De ce point de vue il est supérieur à d'autres algorithmes beaucoup plus sophistiqués, comme les méthodes du type prédicteur-correcteur [10].

2.2 Conditions périodiques

Comme un système simulé ne contient qu'une fraction infinitésimale d'un vrai système macroscopique – typiquement 10^3 à 10^5 atomes devant $\approx 10^{23}$ atomes – on doit appliquer des conditions périodiques afin de limiter des effets de borne. La figure 1 montre le principe des conditions périodiques. Tout atome qui quitte la boîte de simulation d'un côté rentre du côté opposé. Il est évident que la densité moyenne

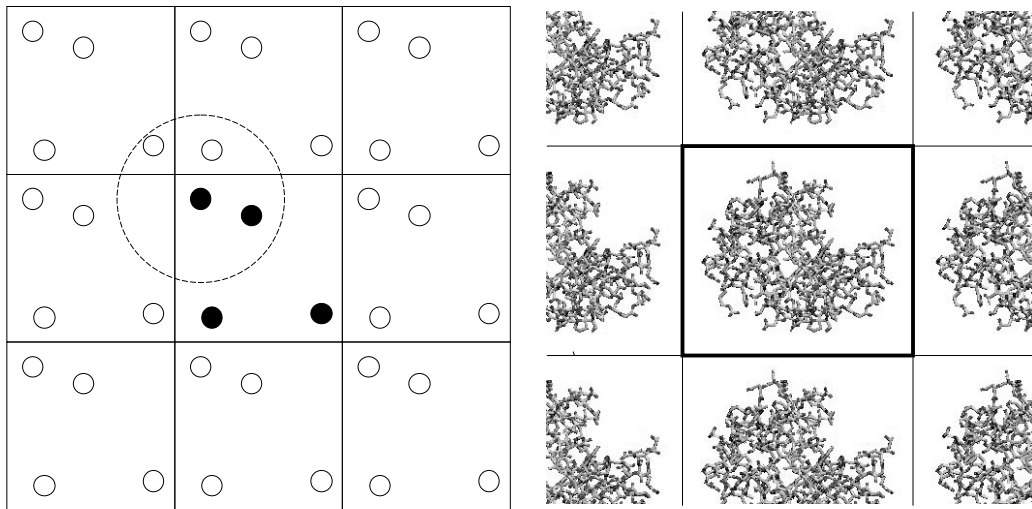


Figure 1. **A gauche:** application de conditions périodiques dans une simulation MD. Le cercle illustre le principe de la convention M.I. (*minimum image*). **A droite:** boîte de simulation pour une protéine en solution et indication des images périodiques. Le solvant n'est pas montré.

d'un système moléculaire est conservée de cette manière et on ne simule pas une goutte infinitésimale d'un vrai système dans laquelle la tension de surface domine les mouvements des atomes à l'intérieur. Afin d'éviter des sauts de l'énergie potentielle au moment où un partenaire dans une interaction entre deux atomes α et β traverse une surface de la boîte de simulation, on applique le principe de "l'image minimale" (M.I. = *minimum image*). Au lieu d'utiliser les coordonnées de l'atome β dans le calcul de la force d'interaction, on utilise celles de l'image β' la plus proche, y compris $\beta' = \beta$. Supposons que la boîte de simulation est cubique de longueur L . Dans ce cas on construit

$$\begin{aligned}x'_{\alpha\beta} &= \min(x_\alpha - x_\beta, x_\alpha - x_\beta + L, x_\alpha - x_\beta - L), \\y'_{\alpha\beta} &= \min(y_\alpha - y_\beta, y_\alpha - y_\beta + L, y_\alpha - y_\beta - L), \\z'_{\alpha\beta} &= \min(z_\alpha - z_\beta, z_\alpha - z_\beta + L, z_\alpha - z_\beta - L),\end{aligned}$$

et on utilise la distance

$$r_{\alpha\beta} = \sqrt{x'_{\alpha\beta}{}^2 + y'_{\alpha\beta}{}^2 + z'_{\alpha\beta}{}^2}$$

pour le calcul des forces d'interactions. La convention M.I. ne peut être utilisée que pour les forces à courte portée, comme les forces du type Lennard-Jones (voir l'expression (2.3)):

$$U_{LJ}(\mathbf{R}_1, \dots, \mathbf{R}_N) = \sum_{\alpha, \beta} 4\epsilon_{\alpha\beta} \left(\left[\frac{\sigma_{\alpha\beta}}{r_{\alpha\beta}} \right]^{12} - \left[\frac{\sigma_{\alpha\beta}}{r_{\alpha\beta}} \right]^6 \right). \quad (2.7)$$

Afin de limiter le calcul des forces à courte portée à un strict minimum on utilise dans la plupart des cas un rayon de coupure $r_c < L/2$, ne tenant compte que des interactions pour lesquelles $r_{\alpha\beta} \leq r_c$.

Les forces coulombiennes qui décroissent $\propto r^{-2}$ ne peuvent pas être calculées de la manière décrite ci-dessus, parce qu'on ne peut pas négliger leurs contributions si r atteint les dimensions de la boîte de simulation – typiquement 5 à 10 nm. Comme on considère un pseudo-cristal à cause des conditions périodiques qui sont appliquées dans une simulation MD on peut utiliser la sommation Ewald pour calculer l'énergie et les forces électrostatiques [9]. P. Ewald avait proposé cette méthode dans sa thèse de 1921 afin de calculer l'énergie électrostatique d'un cristal. En appliquant la sommation Ewald on doit être conscient que la boîte de simulation n'est pas la maille d'un vrai cristal. Toute inhomogénéité dans la distribution de charges est amplifiée par la répétition périodique de la boîte de simulation. En revanche, si la distribution de charges est plutôt homogène on peut utiliser une variante simplifiée de la sommation Ewald qui est basée sur le fait que les interactions électrostatiques sont effectivement écrantées dans ce cas [11].

2.3 Ensemble microcanonique

Du point de vue thermodynamique un système moléculaire qui est simulé par la dynamique moléculaire classique est un ensemble microcanonique, aussi appelé ensemble NVE . Ici le nombre de particules (atomes) N , le volume V du système et l'énergie totale sont constants. La constance de l'énergie vient du fait que les forces dérivent d'un potentiel. Il suit d'abord des équations de Newton que

$$\sum_{\alpha=1}^N \dot{\mathbf{R}}_\alpha \cdot (m_\alpha \ddot{\mathbf{R}}_\alpha - \mathbf{F}_\alpha) = 0.$$

Comme \mathbf{F}_α est le gradient négatif du potentiel U , on a

$$\sum_{\alpha=1}^N \dot{\mathbf{R}}_\alpha \cdot \mathbf{F}_\alpha = - \sum_{\alpha=1}^N \dot{\mathbf{R}}_\alpha \cdot \frac{\partial U}{\partial \mathbf{R}_\alpha} = - \frac{dU}{dt}.$$

En introduisant l'énergie totale du système,

$$E = \sum_{\alpha=1}^N \frac{1}{2} m_{\alpha} \dot{\mathbf{R}}_{\alpha}^2 + U(\mathbf{R}_1, \dots, \mathbf{R}_N) \quad (2.8)$$

on obtient alors

$$\frac{dE}{dt} = 0 \implies E = cste. \quad (2.9)$$

La relation entre la mécanique classique, qui décrit un système au niveau microscopique, et la thermodynamique, qui décrit un système au niveau macroscopique, est le sujet de la mécanique statistique [12–14]. Ici on part de la *fonction de Hamilton*, ou tout court *l'hamiltonien*, qui est défini par

$$\mathcal{H} = \sum_{\alpha=1}^N \frac{\mathbf{P}_{\alpha}^2}{2m_{\alpha}} + U(\mathbf{R}_1, \dots, \mathbf{R}_N) \quad (2.10)$$

où $\{\mathbf{P}_{\alpha}\}$ est la quantités du mouvement de l'atome α ,

$$\mathbf{P}_{\alpha} = m_{\alpha} \dot{\mathbf{R}}_{\alpha}. \quad (2.11)$$

L'hamiltonien décrit la dynamique d'un système dans *l'espace de phases* $6N$ -dimensionnel, qui est engendré par les positions \mathbf{R}_{α} et les quantités du mouvement associées \mathbf{P}_{α} ($\alpha = 1, \dots, N$). Les équations de Hamilton ont la forme symétrique

$$\dot{\mathbf{R}}_{\alpha} = \frac{\partial \mathcal{H}}{\partial \mathbf{P}_{\alpha}}, \quad (2.12)$$

$$\dot{\mathbf{P}}_{\alpha} = -\frac{\partial \mathcal{H}}{\partial \mathbf{R}_{\alpha}}. \quad (2.13)$$

Avec la forme (2.10) pour l'hamiltonien on obtient alors

$$\dot{\mathbf{R}}_{\alpha} = \frac{\mathbf{P}_{\alpha}}{m_{\alpha}}, \quad (2.14)$$

$$\dot{\mathbf{P}}_{\alpha} = \mathbf{F}_{\alpha}. \quad (2.15)$$

On reconnaît facilement l'équivalence avec les équations de Newton et également que $\mathcal{H} = E$ dans un ensemble microcanonique.

La valeur moyenne thermodynamique d'une variable A est définie comme moyenne sur l'espace de phases. On utilisere la définition $\Gamma = (\mathbf{P}_1, \dots, \mathbf{P}_N; \mathbf{R}_1, \dots, \mathbf{R}_N)$ pour un point dans l'espace de phases afin de garder une notation compacte. Avec ceci la moyenne d'une variable dynamique A est donnée par

$$\langle A \rangle = \int_{\Gamma} d\Gamma A(\Gamma) f(\Gamma) \quad (2.16)$$

Ici $f(\Gamma)$ est la fonction de distribution dans l'espace de phases qui a la forme suivante pour un ensemble microcanonique:

$$f_{NVE}(\Gamma) = \frac{\delta(\mathcal{H}(\Gamma) - E)}{\int_{\Gamma} d\Gamma \delta(\mathcal{H}(\Gamma) - E)} \quad (2.17)$$

La distribution de Dirac $\delta(\mathcal{H}(\Gamma) - E)$ décrit la contrainte $\mathcal{H}(\Gamma) = E$ qui restreint les points accessibles dans l'espace de phase à une sphère $(6N - 1)$ -dimensionnelle. Sur cette sphère aucun point n'est

privilegié par rapport à un autre. Les propriétés essentielles de la distribution de Dirac sont listées dans l'annexe.

2.4 Ensembles thermodynamiques non-standard

L'ensemble microcanonique ne correspond pas aux conditions physiologiques des macromolécules *in vivo*. Ces conditions sont caractérisées par une température et une pression constante et donc par l'ensemble NpT . Il faut préciser que ni la température microscopique,

$$T(t) = \frac{1}{3Nk_B} \sum_{\alpha=1}^N \frac{\mathbf{P}_\alpha^2}{m_\alpha}, \quad (2.18)$$

ni la pression microscopique,

$$p(t) = \frac{1}{3V} \left\{ \sum_{\alpha=1}^N \frac{\mathbf{P}_\alpha^2}{m_\alpha} - \sum_{\alpha=1}^N \mathbf{R}_\alpha \cdot \mathbf{F}_\alpha \right\} \quad (2.19)$$

sont constantes dans cet ensemble, mais leur valeurs moyennes qui correspondent aux paramètres macroscopiques de l'ensemble. Ici k_B est la constante de Boltzmann¹. On note que la distribution d'équilibre de l'ensemble NpT a la forme

$$f_{NpT}(\Gamma) = \frac{\int_0^\infty dE \int_0^\infty dV \exp(-\beta[E + pV]) \delta(\mathcal{H}(\mathbf{E}, V) - E)}{\int_\Gamma d\Gamma \int_0^\infty dE \int_0^\infty dV \exp(-\beta[E + pV]) \delta(\mathcal{H}(\mathbf{E}, V) - E)} \quad (2.20)$$

où $\beta = (k_B T)^{-1}$. Afin de simuler la dynamique d'un système mécanique dans l'ensemble NpT , Andersen and Nosé ont introduit le principe du système étendu [15–17]. On considère la dynamique d'un système composé du système physique auquel on ajoute des degrés de liberté supplémentaires qui représentent, respectivement, un thermostat et un barostat. Ces variables dynamiques supplémentaires agissent comme variables d'échelle qui modifient les vitesses et les positions du système physique, tel que le système physique suit une dynamique dans l'ensemble souhaité. Les équations du mouvement pour l'ensemble NpT ont la forme suivante (comparer éqs. (2.14) et (2.15)):

$$\dot{\mathbf{R}}_\alpha = \frac{\dot{V}}{3V} \mathbf{R}_\alpha + \frac{\mathbf{P}_\alpha}{m_\alpha}, \quad (2.21)$$

$$\dot{\mathbf{P}}_\alpha = \mathbf{F}_\alpha - \frac{\dot{V}}{3V} \mathbf{P}_\alpha - \zeta \mathbf{P}_\alpha, \quad (2.22)$$

$$\dot{V} = \frac{pV}{W_V}, \quad (2.23)$$

$$\dot{p}_V = \frac{1}{3V} \left\{ \sum_{\alpha=1}^N \frac{\mathbf{P}_\alpha^2}{m_\alpha} - \sum_{\alpha=1}^N \mathbf{R}_\alpha \cdot \mathbf{F}_\alpha \right\} - p - \zeta p_V, \quad (2.24)$$

$$\dot{\zeta} = \frac{1}{W_S} \left\{ \sum_{\alpha=1}^N \frac{\mathbf{P}_\alpha^2}{m_\alpha} + \frac{p_V^2}{2W_V} - (3N + 1)k_B T \right\}. \quad (2.25)$$

Ici W_S et W_V sont des pseudo-masses qui décrivent, respectivement, l'inertie du thermostat et du barostat. Elles déterminent la vitesse de réaction du système à une dérive des valeurs instantanées de la température et de la pression des valeurs macroscopiques imposées.

¹ $k_B = 1.381 \cdot 10^{-23} \text{ J/K}$.

3. ANALYSE SIMPLE DE SIMULATIONS

Dans la suite seront présentées quelques méthodes pour l'analyse d'une trajectoire générée par simulation MD. Il s'agit de méthodes classiques qui sont implémentées dans le code nMoldyn [18].

3.1 Fonction de corrélation des vitesses et son spectre de Fourier

Afin d'obtenir une première impression de la dynamique d'un système moléculaire, on calcule la fonction d'autocorrélation des vitesses (FACV) et son spectre de Fourier. La dernière quantité est souvent appelée densité d'états (*density of states* = DOS). Dans un système contenant plusieurs types d'atomes on calcule une moyenne sur les N atomes du système. Souvent on applique un schéma de pondération où chaque atome α obtient un poids w_α suivant sa masse ou une autre propriété physique. Avec la condition $\sum_{\alpha=1}^N w_\alpha = N$ on définit ainsi la FACV pour un système isotrope par

$$c_{vv}(t) = \frac{1}{3N} \sum_{\alpha=1}^N w_\alpha \langle \mathbf{v}_\alpha(0) \cdot \mathbf{v}_\alpha(t) \rangle \quad (3.1)$$

où $\mathbf{v}_\alpha = (v_{\alpha,x}, v_{\alpha,y}, v_{\alpha,z})$ contient les coordonnées cartésiennes de la vitesse de l'atome α .

Le symbole $\langle \dots \rangle$ indique ici une moyenne sur le temps. Pour deux variables dynamiques A et B on définit la *moyenne* de A et la *fonction de corrélation* de A et B par

$$\langle A(t_0) \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} d\tau A(\tau + t_0), \quad (3.2)$$

$$\langle A(t_1)B(t_0) \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} d\tau A(\tau + t_1)B(\tau + t_0). \quad (3.3)$$

Afin d'établir une relation avec la mécanique statistique, on suppose que la moyenne sur t_0 est équivalente à une moyenne sur un ensemble dans l'espace de phases. Ici chaque point Γ définit une condition initiale, $\mathbf{\Gamma} \equiv \mathbf{\Gamma}(t_0)$, pour une trajectoire $\mathbf{\Gamma}(t)$ qui évolue pendant un temps $t = t_1 - t_0$. La pondération de chacun de ces points est déterminée par une densité de probabilité comme $f_{NVE}(\mathbf{\Gamma})$ ou $f_{NpT}(\mathbf{\Gamma})$. Une fonction de corrélation entre deux variables dynamiques est définie par

$$\langle A(t_1)B(t_0) \rangle = \int_{\Gamma} d\Gamma(t_0) f(\mathbf{\Gamma}(t_0)) B(\mathbf{\Gamma}(t_0)) A(\mathbf{\Gamma}(t_1)) \quad (3.4)$$

Si le système étudié est dans l'équilibre, la moyenne $\langle A(t_0) \rangle$ ne dépend pas de t_0 et $\langle A(t_1)B(t_0) \rangle$ ne dépend que de la différence $t = t_1 - t_0$. On note que l'équivalence d'une moyenne sur le temps et une moyenne sur l'ensemble est en générale une hypothèse qui ne peut être prouvée que pour très peu de systèmes. Ce point est entre autres discuté dans la dernière monographie de W. Hoover [19].

Numériquement la FACV pour chaque atome peut être estimée par

$$\langle \mathbf{v}(0) \cdot \mathbf{v}(m) \rangle \approx \frac{1}{N_t - m} \sum_{k=0}^{N_t - m - 1} \mathbf{v}(k) \cdot \mathbf{v}(k + m), \quad (3.5)$$

partant d'une trajectoire discrète, $\mathbf{v}(n) \equiv \mathbf{v}(n\Delta t)$. Ici Δt est la distance temporelle entre deux configurations consécutives dans la trajectoire et N_t est le nombre de configurations dans une trajectoire. Le calcul direct des FACV des atomes individuels peut être considérablement augmentée en utilisant la transformation de Fourier rapide (*Fast Fourier Transform*=FFT) [18, 20, 21].

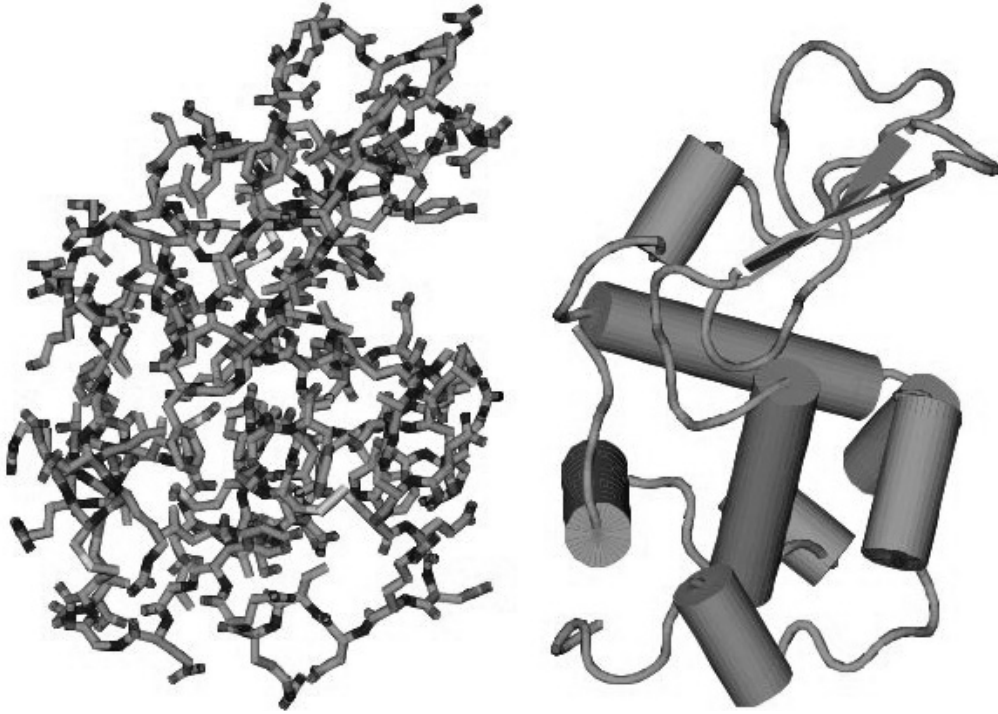


Figure 2. Le lysozyme représenté par les liaisons covalentes (à gauche) et sous forme d'un "cartoon" qui montre les hélices et les feuillets beta comme éléments de structure secondaire (à droite).

Le spectre de Fourier de la FACV est une quantité spectroscopique qui est accessible par diffusion inélastique de neutrons [22]. On écrit

$$g_{vv}(\omega) = \int_0^{\infty} dt \cos(\omega t) c_{vv}(t), \quad (3.6)$$

où on note que $c_{vv}(t)$ est symétrique pour un système dans l'équilibre. Numériquement on obtient $g_{vv}(\omega)$ par FFT de la FACV discrète, en appliquant préalablement une fenêtre pour le lissage nécessaire du spectre [23],

$$g_{vv}(n) = \Delta t \cdot \sum_{m=-(N_t-1)}^{N_t-1} \exp\left(-2\pi i \left(\frac{nm}{2N_t}\right)\right) w(m) c_{vv}(m). \quad (3.7)$$

La fenêtre $w(t)$ peut avoir différentes formes, en fonction des propriétés du spectre souhaitées [24]. On remarque que $g_{vv}(n) \equiv g_{vv}(n\Delta\omega)$, où $\Delta\omega = \pi/(N\Delta t)$. La partie gauche de la figure 3 montre la DOS séparément pour les atomes d'hydrogène et les atomes d'oxygène dans une boîte de simulation de 256 molécules d'eau en totale. La dynamique a été simulée à température et pression ambiante, en utilisant le potentiel SPCE [25]. La DOS des atomes d'hydrogène est dominée par une bande large autour de 15 THz qui est due aux librations des molécules d'eau (1 THz = 10^{12} Hz). On remarque que la position de l'oxygène coïncide pratiquement avec le centre de masse d'une molécule d'eau, et pour cette raison la bande de libration est absente dans la DOS de l'oxygène. La partie droite de la figure 3 montre la DOS pour une molécule de lysozyme en solution. La structure de cette protéine est montrée dans la fig. 2. Chaque atome a été pondéré suivant sa section efficace pour la diffusion incohérente de neutrons. On remarque que $\sigma_{inc} = 4\pi b_{inc}^2$ pour un atome fixe [22]. Pour cette raison on ne tient pratiquement compte

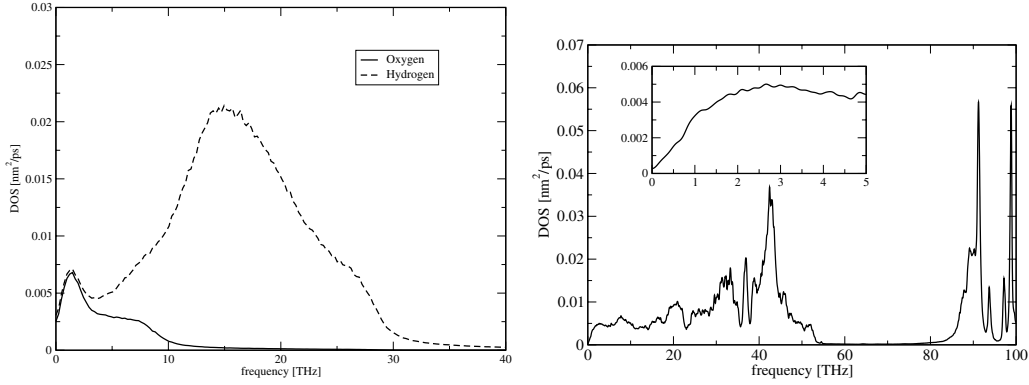


Figure 3. **A gauche:** DOS simulée pour les atomes d’hydrogène et les atomes d’oxygène dans l’eau à température et pression ambiantes. Le système simulé consiste de 256 molécules d’eau dans une boîte cubique. **A droite:** DOS moyennée sur tous les atomes d’une molécule de lysozyme en solution. La pondération de chaque atome est suivant la section efficace pour la diffusion incohérente de neutrons.

Table 1. Longueurs de diffusion de quelques éléments en fm ($10^{-15} m$).

Element	H	D	C	O	N	S
b_{coh}	-3.741	6.674	6.648	5.805	9.300	2.847
b_{inc}	25.217	4.022	0.285	0.000	2.241	0.188

que des atomes d’hydrogène (voir tableau 1). Les hautes fréquences entre 80 et 100 THz correspondent aux vibrations des liaisons impliquant un atome d’hydrogène. En descendant vers les basses fréquences on obtient d’abord, respectivement, les contributions des vibrations des liaisons entre des atomes lourds, les vibrations d’angles de liaison, et les mouvements de torsion. En descendant encore plus bas en fréquence un voit des mouvements collectifs impliquant un nombre croissant d’atomes. Cette partie du spectre est montré dans dans l’insert de la figure 3. On remarque la DOS est très similaire pour différentes protéines d’une taille comparable [26], et uniquement les mouvements de très basse fréquence sont spécifiques à une protéine [27].

3.2 Déplacement carré moyen

Une quantité dynamique qui donne accès aux amplitudes des mouvements d’une particule est le déplacement carré moyen (DCM). Dans un système contenant différents types d’atomes on définit

$$W(t) = \frac{1}{N} \sum_{\alpha=1}^N w_{\alpha} \langle [\mathbf{R}_{\alpha}(t) - \mathbf{R}_{\alpha}(0)]^2 \rangle \quad (3.8)$$

où w_{α} est le poids de l’atome α et $\mathbf{R}_{\alpha}(t)$ sa position en fonction du temps. Le DCM pour un atome individuel peut être estimé par

$$\langle [\mathbf{R}(m) - \mathbf{R}(0)]^2 \rangle \approx \frac{1}{N_t - m} \sum_{k=0}^{N_t - m - 1} [\mathbf{R}(k + m) - \mathbf{R}(k)]^2. \quad (3.9)$$

Comme pour les fonctions de corrélation, on peut utiliser la transformation de Fourier rapide, afin d’accélérer le calcul des DCM [21].

La partie gauche de figure 4 montre un DCM simulé pour l’eau, où chaque atome a été pondéré suivant sa section efficace pour la diffusion incohérente de neutrons. Le système simulé est le même qui

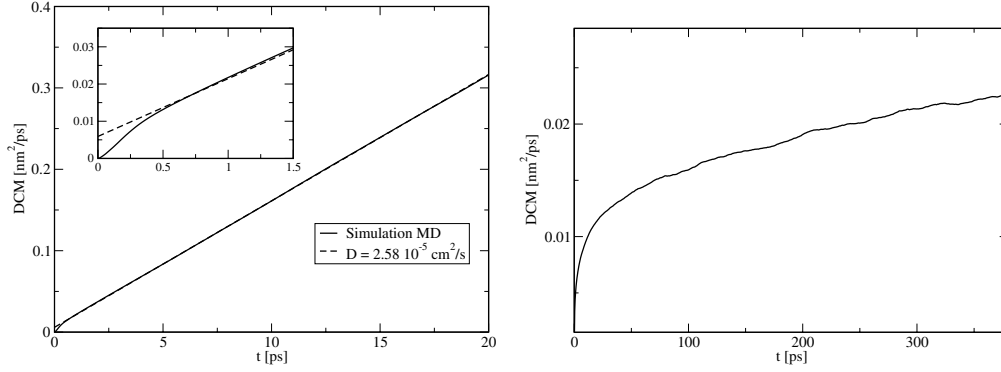


Figure 4. A gauche: DCM simulé pour l'eau à température et pression ambiantes. Le système simulé consiste de 256 molécules d'eau dans une boîte cubique. **A droite:** la même quantité pour une molécule de lysozyme en solution.

celui qui a été utilisé pour le calcul de la DOS montrée dans la figure 3. On voit que le DCM s'approche d'une droite pour $t > 0.5$ ps dont la pente définit le coefficient de diffusion:

$$W(t) \approx 6Dt, \quad t \rightarrow \infty \quad (3.10)$$

La croissance linéaire du DCM avec le temps est caractéristique pour la dynamique d'une particule brownienne [14, 28] qui diffuse sans l'influence d'une force externe. On peut établir une relation entre la constante de diffusion et la FACV, partant de l'identité [29]

$$W(t) = 6 \int_0^t d\tau (t - \tau) c_{vv}(\tau) \quad (3.11)$$

Avec (3.10) on trouve que

$$D = \int_0^\infty dt c_{vv}(t) = g_{vv}(0) \quad (3.12)$$

Le comportement du DCM est très différent si l'on considère une particule dont le mouvement est limité dans l'espace, comme ceci est le cas pour un atome dans une protéine dont les mouvements globaux sont bloqués. Une réalisation physique est une protéine dans une poudre hydratée. Dans le cas d'un mouvement confiné le DCM s'approche d'un plateau dans la limite $t \rightarrow \infty$. La partie droite de la figure 4 montre le DCM pour les mouvements internes d'une molécule de lysozyme en solution. Avant de stocker les configurations, la translation et la rotation globale de la protéine ont été soustraites par superposition optimale avec la configuration de départ [30]. Les poids des atomes sont les mêmes que dans la DOS montré dans la figure 3 (partie droite). La trajectoire simulée a une longueur d'une nanoseconde ($1 \text{ ns} = 10^{-9} \text{ s}$). On voit que le DCM n'atteint pas encore un plateau en 400 ps. Ce comportement indique qu'ils existent des processus de relaxation très lents dans les protéines qui ne peuvent pas être décrits à l'échelle de la ns.

On note ici qu'il existe une relation étroite entre le DCM et la DOS. Il suit de la relation (3.11) que

$$W(t) = \frac{12}{\pi} \int_0^\infty d\omega \frac{1 - \cos \omega t}{\omega^2} g_{vv}(\omega) \quad (3.13)$$

Cette relation montre bien que le DCM est essentiellement déterminé par les contributions de $g_{vv}(\omega)$ à basses fréquences [31].

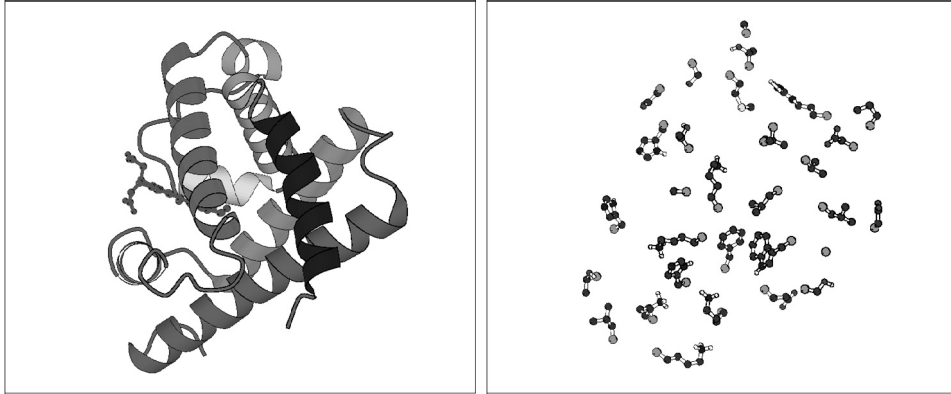


Figure 5. A gauche: squelette de la myoglobine. A droite: 31 chaînes latérales sélectionnées.

3.3 Fonctions de diffusion intermédiaires et facteurs de structure dynamiques

La quantité essentielle qui est obtenue par une expérience de diffusion de neutrons est le facteur de structure dynamique [22, 32]. Il peut être divisé en une partie cohérente et une partie incohérente,

$$\mathcal{S}(\mathbf{q}, \omega) = \mathcal{S}_{coh}(\mathbf{q}, \omega) + \mathcal{S}_{inc}(\mathbf{q}, \omega) \quad (3.14)$$

qui sont chacune la transformée de Fourier d'une fonction de corrélation dans le temps,

$$\mathcal{S}_{coh/inc}(\mathbf{q}, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dt \exp(-i\omega t) \mathcal{I}_{coh/inc}(\mathbf{q}, t) \quad (3.15)$$

Ici $\mathcal{I}_{coh}(\mathbf{q}, t)$ est la fonction intermédiaire de diffusion cohérente

$$\mathcal{I}_{coh}(\mathbf{q}, t) = \frac{1}{N} \sum_{\alpha, \beta} b_{\alpha, coh} b_{\beta, coh} \langle \exp(-i\mathbf{q} \cdot \mathbf{R}_{\alpha}(0)) \exp(i\mathbf{q} \cdot \mathbf{R}_{\beta}(t)) \rangle \quad (3.16)$$

et $\mathcal{I}_{inc}(\mathbf{q}, t)$ est la fonction intermédiaire de diffusion incohérente,

$$\mathcal{I}_{inc}(\mathbf{q}, t) = \frac{1}{N} \sum_{\alpha} b_{\alpha, inc}^2 \langle \exp(-i\mathbf{q} \cdot \mathbf{R}_{\alpha}(0)) \exp(i\mathbf{q} \cdot \mathbf{R}_{\alpha}(t)) \rangle \quad (3.17)$$

Les quantités $b_{\alpha, coh}$ et $b_{\alpha, inc}$ sont, respectivement, la longueur de diffusion cohérente et incohérente de l'atome α . Le tableau 1 donne quelques valeurs typiques. Souvent les fonctions intermédiaires de diffusion sont normalisées, tel que

$$\mathcal{I}_{inc}(\mathbf{q}, 0) = 1, \quad (3.18)$$

$$\lim_{q \rightarrow \infty} \mathcal{I}_{coh}(\mathbf{q}, t) = 1. \quad (3.19)$$

Ceci revient à redéfinir

$$b_{\alpha, inc} \longrightarrow \frac{b_{\alpha, inc}}{\sqrt{\frac{1}{N} \sum_{\alpha} b_{\alpha, inc}^2}} \quad (3.20)$$

$$b_{\alpha, coh} \longrightarrow \frac{b_{\alpha, coh}}{\sqrt{\frac{1}{N} \sum_{\alpha} b_{\alpha, coh}^2}}. \quad (3.21)$$

On note que cette définition est utilisée dans [18].

La diffusion incohérente donne accès aux corrélations dans les mouvements d'atomes individuels, tandis que la diffusion cohérente permet d'étudier des mouvements collectifs. La pondération respective des contributions cohérentes et incohérentes est déterminée par les longueurs de diffusion $b_{\alpha,inc}$ et $b_{\alpha,coh}$. Le tableau 1 montre que la diffusion incohérente de l'hydrogène domine tout autre processus de diffusion incohérente et cohérente. Pour cette raison le remplacement $H \rightarrow D$ est utilisé afin d'établir un contraste entre les composantes d'un système dont on souhaite étudier la dynamique et le reste. Ceci concerne l'approche expérimentale. L'avantage d'une approche combinée de simulation moléculaire et diffusion de neutrons est la facilité d'étudier par analyse d'une trajectoire simulée toute contribution d'un type de mouvement ou d'un groupe d'atomes au spectre total. Une étude de la diffusion quasiélastique de la myoglobine illustre cet aspect [33]. Cette protéine – comme d'autres – subit une transition dynamique autour d'une température de 200 K environ [34]. En dessous de cette température la dynamique d'une protéine est caractérisée par des vibrations de sa structure autour d'un minimum de l'énergie potentielle. En passant à des températures au dessus de 200 K on observe par diffusion de neutrons l'apparition d'une ligne quasiélastique qui indique la présence des mouvements diffusifs. Le but de l'étude [33] était de trouver le type de mouvement associé à ces mouvements diffusifs. Nous avons pu montrer que la diffusion quasiélastique d'une protéine est essentiellement due aux mouvements des chaînes latérales, vues comme corps rigides. La fig. 6 (partie gauche) montre un spectre quasiélastique d'une poudre hydratée de la myoglobine [34] et le spectre simulé correspondant. La comparaison n'est pas directe sur toute la gamme des fréquences, car des contributions inélastiques estimées ont été enlevées des spectres expérimentaux. Dans la partie quasiélastique du spectre (entre 0 et 0.2 meV environ) l'accord entre le spectre simulé et les données expérimentales est très bon et on ne voit aucune différence entre le spectre total simulé est la contribution des chaînes latérales rigides. On peut faire le même constat pour la diffusion élastique qui est représentée par l'EISF (voir section 3.5) qui est montré dans la partie droite de la figure 6.

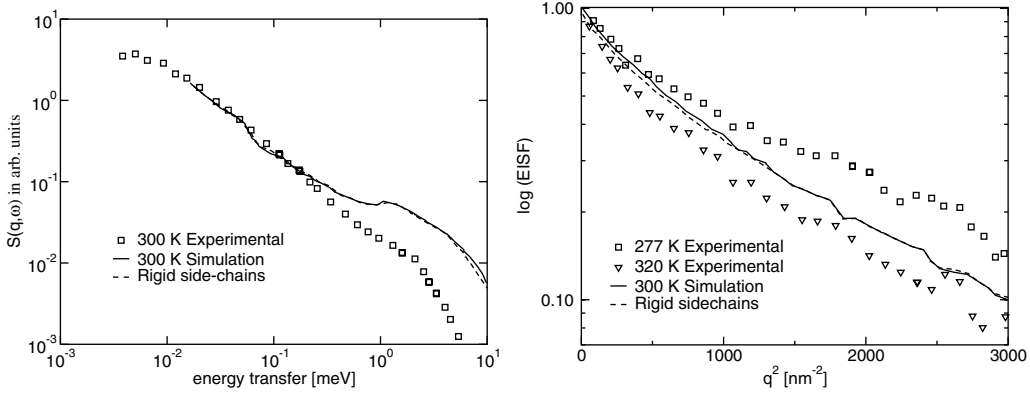


Figure 6. A gauche: spectre quasiélastique de neutrons de la myoglobine par expérience (carreaux) et simulation. La ligne droite montre le spectre simulé total et la ligne pointillée la contribution des mouvements du type corps rigide des chaînes latérales. Plus de détails sont donnés dans le texte. **A droite:** la même comparaison pour l'EISF.

3.4 Approximation gaussienne

La fonction de diffusion intermédiaire incohérente peut être présentée sous forme d'un développement en cumulants [29, 35]

$$\mathcal{I}_{inc}(\mathbf{q}, t) = \frac{1}{N} \sum_{\alpha} b_{\alpha,inc}^2 \exp(-q^2 \rho_{\alpha,1}(t) + q^4 \rho_{\alpha,2}(t) \mp \dots), \quad (3.22)$$

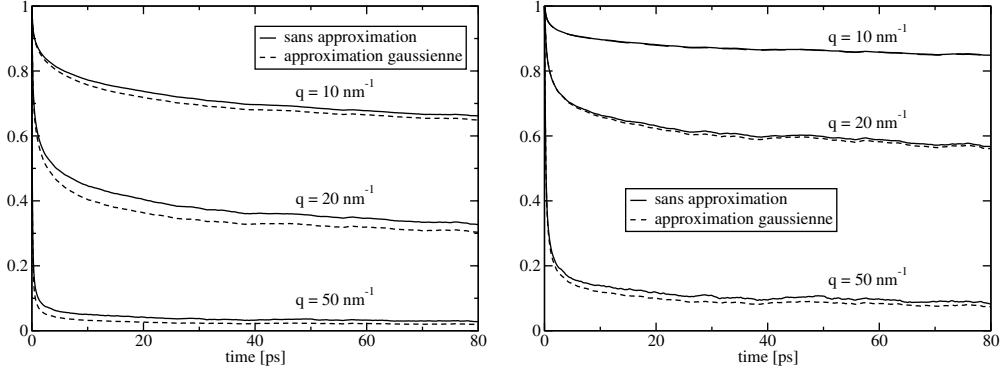


Figure 7. A gauche: approximation gaussienne pour la fonction intermédiaire de diffusion des atomes d'hydrogène dans le lysozyme. **A droite:** idem pour les atomes C_α du squelette.

où $q \equiv |\mathbf{q}|$. Les deux premiers cumulants $\rho_{\alpha,k}(t)$ sont donnés par

$$\rho_{\alpha,1}(t) = \frac{1}{2!} \langle d_\alpha^2(t; \mathbf{n}_q) \rangle, \quad (3.23)$$

$$\rho_{\alpha,2}(t) = \frac{1}{4!} [\langle d_\alpha^4(t; \mathbf{n}_q) \rangle - 3 \langle d_\alpha^2(t; \mathbf{n}_q) \rangle^2]. \quad (3.24)$$

où $d_\alpha(t; \mathbf{n}_q)$ est la projection du déplacement

$$\mathbf{d}_\alpha(t) = \mathbf{R}_\alpha(t) - \mathbf{R}_\alpha(0) \quad (3.25)$$

sur le vecteur d'unité $\mathbf{n}_q = \mathbf{q}/|\mathbf{q}|$,

$$d_\alpha(t; \mathbf{n}_q) \doteq \mathbf{n}_q \cdot \mathbf{d}_\alpha(t). \quad (3.26)$$

En négligeant tous les cumulants $\rho_{\alpha,k}(t)$ pour $k > 1$ on obtient l'approximation gaussienne de la fonction intermédiaire de diffusion,

$$\mathcal{I}_{inc}(\mathbf{q}, t) \approx \frac{1}{N} \sum_{\alpha} b_{\alpha,inc}^2 \exp(-q^2 \rho_{\alpha,1}(t)) \quad (3.27)$$

qui est entièrement définie par les DCM des atomes. Dans un système isotrope, où toutes les directions sont équivalentes, ceci devient

$$\mathcal{I}_{inc}(\mathbf{q}, t) \approx \frac{1}{N} \sum_{\alpha} b_{\alpha,inc}^2 \exp\left(-\frac{q^2}{6} W_\alpha(t)\right) \quad (3.28)$$

On remarque que la forme gaussienne de $\mathcal{I}_{inc}(\mathbf{q}, t)$ est exacte pour un certain nombre de modèles, comme le gaz parfait, l'oscillateur harmonique et la particule brownienne libre [29]. Les deux derniers modèles seront discutés dans la section suivante. La figure 7 montre l'approximation gaussienne pour la fonction intermédiaire incohérente des atomes d'hydrogène dans le lysozyme (à gauche) et la quantité correspondante pour les atomes C_α du squelette (à droite). Dans le dernier cas l'approximation est très bonne et pour les atomes d'hydrogène elle est toujours bonne. La comparaison confirme les résultats d'une étude de simulation de la protéine C-phycocyanine qui a montré que la dynamique de cette protéine au niveau des résidus est bien décrite par un modèle d'oscillateurs harmoniques couplés avec friction [36]. Pour ce modèle la forme gaussienne de la fonction intermédiaire de diffusion est exacte. Le fait que l'approximation gaussienne pour les atomes d'hydrogène n'est pas aussi bonne que celle pour les atomes C_α peut être expliqué par la distribution de ces atomes dans la protéine. Au contraire aux atomes C_α , les

atomes d'hydrogène sont distribués sur toute la protéine. Une fraction non-négligeable est située dans les extrémités des chaînes latérales et participe ainsi à des mouvements de rotation qui ne peuvent pas être décrits par l'approximation gaussienne.

3.5 EISF

Si les mouvements d'un système moléculaire sont limités dans l'espace, tous les cumulants dans (3.22) restent bornés dans la limite $t \rightarrow \infty$ et la fonction intermédiaire de diffusion incohérente tend vers une valeur non nulle,

$$\boxed{EISF(\mathbf{q}) = \lim_{t \rightarrow \infty} \mathcal{I}_{inc}(\mathbf{q}, t)} \quad (3.29)$$

Cette quantité est le facteur de structure incohérent élastique (EISF = *elastic incoherent structure factor*). Avec la définition

$$\mathcal{I}'_{inc}(\mathbf{q}, t) = \mathcal{I}_{inc}(\mathbf{q}, t) - EISF(\mathbf{q}) \quad (3.30)$$

le facteur de structure dynamique peut être écrit sous la forme

$$\boxed{\mathcal{S}_{inc}(\mathbf{q}, \omega) = EISF(\mathbf{q})\delta(\omega) + \mathcal{S}'_{inc}(\mathbf{q}, \omega)} \quad (3.31)$$

où $\mathcal{S}'_{inc}(\mathbf{q}, \omega)$ est la transformée de Fourier de $\mathcal{I}'_{inc}(\mathbf{q}, t)$. Ceci montre que l'EISF pondère la contribution élastique du facteur de structure dynamique, ce que lui donne son nom. Comme les positions $\mathbf{R}_\alpha(t)$ et $\mathbf{R}_\alpha(0)$ dans l'expression (3.17) deviennent non-corrélées pour $t \rightarrow \infty$, l'EISF est donné par la moyenne statique

$$EISF(\mathbf{q}) = \frac{1}{N} \sum_{\alpha} b_{\alpha,inc}^2 |\langle \exp(i\mathbf{q} \cdot \mathbf{R}_\alpha) \rangle|^2. \quad (3.32)$$

Dans l'approximation gaussienne l'EISF prend une forme particulièrement simple. Partant de l'expression (3.27) il suit de la définition de l'EISF que

$$EISF(\mathbf{q}) = \lim_{t \rightarrow \infty} \frac{1}{N} \sum_{\alpha} b_{\alpha,inc}^2 \exp(-q^2 \rho_{\alpha,1}(t)).$$

Avec la définition

$$R_q(t) = \mathbf{n}_q \cdot \mathbf{R}(t) \quad (3.33)$$

le DCM pour l'atome α peut être écrit sous la forme

$$\rho_{\alpha,1}(t) = 2\langle R_q^2 \rangle - 2\langle R_q(t)R_q(0) \rangle.$$

Ici on utilise que $\langle R_q^2(t) \rangle = \langle R_q^2(0) \rangle \equiv \langle R_q^2 \rangle$ dans l'équilibre. Dans la limite $t \rightarrow \infty$ la fonction de corrélation $\langle R_q(t)R_q(0) \rangle$ tend vers zéro et on obtient

$$\lim_{t \rightarrow \infty} \rho_{\alpha,1}(t) = 2\langle R_q^2 \rangle. \quad (3.34)$$

Par conséquent l'EISF prend la forme

$$EISF(\mathbf{q}) = \frac{1}{N} \sum_{\alpha} b_{\alpha,inc}^2 \exp(-q^2 \langle R_q^2 \rangle). \quad (3.35)$$

Pour un système isotrope on peut également écrire

$$\boxed{EISF(\mathbf{q}) = \frac{1}{N} \sum_{\alpha} b_{\alpha,inc}^2 \exp\left(-\frac{q^2}{3} \langle \mathbf{R}_\alpha^2 \rangle\right)} \quad (3.36)$$

3.6 Limite classique

Jusqu'à présent nous avons tacitement supposé que toutes les fonctions de corrélation peuvent être calculées suivant les lois de la mécanique classique. Cette approximation n'est pas toujours justifiée. *A priori* toute grandeur spectroscopique est formellement obtenue par une fonction de corrélation quantique [14, 22]. L'intégration sur le point initial $\Gamma(t_0)$ dans la moyenne (3.4) est remplacé par une somme (ou intégrale) sur les états quantiques initiales, qui sont pondérés par un facteur de Boltzmann $\propto \exp(-E_n/k_B T)$. Ici n numérote les états quantiques et E_n est l'énergie qui correspond à l'état no. n . Il suit de ce calcul que

$$\langle A(t)B(0) \rangle = \langle B(-t + i\beta\hbar)A(0) \rangle, \quad \beta := \frac{1}{k_B T}, \quad (3.37)$$

où $\hbar = h/2\pi$ et h est la constante de Planck ². Une conséquence de l'identité (3.37) est la "relation du bilan détaillé" qui est fondamentale pour l'interprétation des expériences de diffusion de neutrons [22, 32]

$$\boxed{S(\mathbf{q}, \omega) = \exp(\beta\hbar\omega)S(-\mathbf{q}, -\omega)} \quad (3.38)$$

Le remplacement des fonctions de corrélations quantiques par les analogues classiques revient à poser

$$\boxed{\hbar \rightarrow 0} \quad (3.39)$$

Comme il a été exposé dans [37], cette limite implique la négligence de tout effet quantique dans le système étudié et également la négligence de sa perturbation par la collision avec le neutron. Ce dernier effet, qui n'est pas d'origine quantique, peut être négligé si

$$\frac{\hbar^2 q^2}{2M} \ll k_B T, \quad (3.40)$$

où M est la masse (effective) de l'atome diffuseur. On note que $\hbar q = |\mathbf{p}_0 - \mathbf{p}|$ est le module du transfert de quantité du mouvement du neutron au système pendant le processus de diffusion.

On suppose maintenant que la condition (3.40) est vérifiée et on discute la validité de l'approximation classique pour le système étudié. L'oscillateur harmonique est un bon modèle qui permet de discuter cette question dans un cas simple. D'après les lois de la mécanique quantique l'énergie d'un oscillateur harmonique, dont le potentiel en une dimension est donné par $U(x) = \frac{1}{2}Kx^2$ ($K > 0$), peut prendre les valeurs $E_n = (n + \frac{1}{2})\hbar\omega_0$, où la pulsation ω_0 est donnée par $\omega_0 = \sqrt{K/M}$, M est la masse de l'oscillateur. En revanche, dans le système classique correspondant l'énergie peut prendre toute valeur $E > 0$. L'approximation classique est valable si les énergies E_n constituent un quasi-continuum à l'échelle de l'énergie thermique,

$$\boxed{\hbar\omega_0 \ll k_B T} \quad (3.41)$$

Afin d'estimer la validité de l'approximation classique pour la simulation MD d'un système moléculaire quelconque on peut approcher le potentiel par une fonction quadratique. Pour le potentiel Lennard-Jones on trouve par exemple

$$U_{LJ}(r) = 4\epsilon \left(\left[\frac{\sigma}{r} \right]^{12} - \left[\frac{\sigma}{r} \right]^6 \right) \approx -\epsilon + \frac{18 \cdot 2^{2/3} \epsilon (r - r_0)^2}{\sigma^2}, \quad (3.42)$$

où $r_0 = 2^{1/6}\sigma$ est la distance pour laquelle le potentiel prend sa valeur minimale, $U_{LJ,min} = -\epsilon$. A part une constante, l'approximation a donc la forme $U(x) = \frac{1}{2}Kx^2$, où $x = r - r_0$, et on peut en déduire une

² $h = 6.626 \cdot 10^{-34} \text{ Js}$

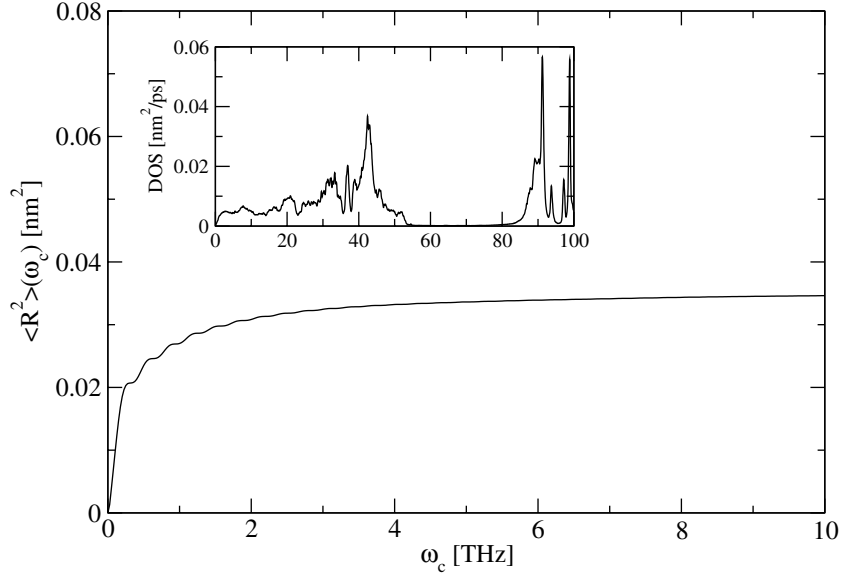


Figure 8. Contribution cumulative de $g_{vv}(\omega)$ à la fluctuation moyenne des atomes dans une molécule de lysozyme en solution. L'insert montre la DOS.

constante de force K . La pulsation correspondante devient

$$\omega_0 = \sqrt{\frac{18 \cdot 2^{2/3} \epsilon}{\mu \sigma^2}},$$

où $\mu = M/2$ est la masse réduite d'un atome dans une interaction entre deux atomes identiques de masse M . Pour le cas concret de l'argon liquide on trouve pour une température de $T = 94.4 \text{ K}$ que $\hbar \omega_0 \approx 0.4 k_B T$, si l'on utilise les paramètres ϵ et σ de la simulation de Rahman [38]. Ceci montre que l'approximation classique est raisonnable pour la simulation MD d'un liquide simple. Dans une protéine la situation n'est pas tout à fait la même, car il y a des "interactions liées" qui décrivent entre autres les vibrations de liaisons à haute fréquence pour lesquelles la condition (3.41) n'est pas vérifiée. On trouve au contraire $\hbar \omega_0 \gg k_B T$. En utilisant $\omega_0 = 2\pi \nu_0$, où ν_0 est la fréquence de la vibration de l'oscillateur, la condition (3.41) devient à température ambiante

$$\nu_0 \ll 6 \text{ THz}$$

La figure 3 montre que la description classique de $g_{vv}(\omega)$ n'est justifiée que pour une petite fraction des fréquences du spectre des mouvements d'une protéine. Ceci ne peut pour autant pas dire que les simulations MD des protéines sont inutiles, car ce sont des mouvements de basse fréquence qui contribuent le plus aux mouvements. Ce point est illustré par la figure 8 qui montre la contribution cumulative de $g_{vv}(\omega)$ à la fluctuation moyenne des atomes dans une molécule de lysozyme en solution. On note que

$$W(t) = \langle [\mathbf{R}(t) - \mathbf{R}(0)]^2 \rangle = 2(\langle \mathbf{R}^2 \rangle - \langle \mathbf{R}(t) \cdot \mathbf{R}(0) \rangle),$$

pour une particule dans un système dans l'équilibre thermodynamique. Comme $\lim_{t \rightarrow \infty} \langle \mathbf{R}(t) \cdot \mathbf{R}(0) \rangle = 0$ on obtient

$$\lim_{t \rightarrow \infty} W(t) = 2 \langle \mathbf{R}^2 \rangle \quad (3.43)$$

pour un mouvement confiné dans l'espace. Partant de cette relation on trouve avec (3.13)

$$\langle \mathbf{R}^2 \rangle(\omega_c) = \lim_{t \rightarrow \infty} \frac{6}{\pi} \int_0^{\omega_c} d\omega \frac{1 - \cos \omega t}{\omega^2} g_{vv}(\omega) \quad (3.44)$$

où ω_c définit une borne supérieure pour l'intégration sur les fréquences. On obtient donc la contribution de toutes les fréquences (pulsations) $\omega \leq \omega_c$ à la fluctuation $\langle \mathbf{R}^2 \rangle$. La figure 8 illustre bien que seules les basses fréquences, qui sont bien décrites par la dynamique classique, contribuent aux mouvements essentiels d'une protéine. On ne peut pourtant pas conclure que tous les mouvements de haute fréquence peuvent être "figés", car les termes correspondants du potentiel contribuent à la flexibilité de la protéine, qui est de son tour essentiel pour les mouvements de basse fréquence et grande amplitude [31].

4. RELAXATION MULTI-ÉCHELLES DANS UNE PROTÉINE

La dynamique interne des protéines est caractérisée par une multitude d'échelles de temps caractéristiques très différentes, qui vont de la femtoseconde à la seconde. Quant aux phénomènes de relaxation, on ne peut pas associer une échelle de temps caractéristique à la décroissance d'une fonction de corrélation, comme la FACV ou la fonction intermédiaire de diffusion. Depuis longtemps ce phénomène est observé dans la spectroscopie de relaxation des matériaux diélectriques et dans la spectroscopie de relaxation RMN [39–41]. Plusieurs modèles empiriques ont été développés afin de tenir compte du comportement non-exponentiel des fonctions de corrélation mesurées. Dans certains cas différents modèles apparaissent comme cas spéciaux d'un modèle pilote [42]. Bien connue est "l'exponentielle étirée"

$$\psi_{EE}(t) = \exp\left(-[t/\tau]^\beta\right), \quad 0 < \beta \leq 1, \quad (4.1)$$

qui a été proposé par Williams and Watt [40] et qui a été également utilisée pour l'analyse des spectres provenant de la spectroscopie de neutrons [43]. En 1995 Glöckle et Nonnenmacher ont proposé le modèle de la dynamique brownienne fractionnaire afin de décrire la cinétique de la liaison de l'oxygène à la myoglobine [44]. Une autre variante – le processus de Ornstein-Uhlenbeck fractionnaire – a été utilisée par Xie et ses collaborateurs afin de modéliser les spectres de fluorescence provenant d'une seule protéine [3] et le transfert d'électrons dans une seule protéine [2]. On note que la dynamique brownienne fractionnaire est un modèle qui a été développé par B. Mandelbrot et ses collaborateurs [45], et qui a trouvé de nombreuses applications, notamment en mathématiques des finances.

4.1 Équation de Langevin généralisé – fonction mémoire

Depuis les travaux de Robert Zwanzig sur la théorie de la dynamique des liquides on sait que toute variable dynamique $u(t) \equiv u(p(t), x(t))$ dans l'espace de phases (p, x) peut être décrite par une *équation de Langevin généralisée* [46]

$$\frac{d}{dt} u + \int_0^t d\tau \zeta(t - \tau) u(\tau) = f^+(t) \quad (4.2)$$

Le noyau $\zeta(t)$ est la *fonction mémoire* et $f^+(t)$ a la propriété

$$\langle u(0) f^+(t) \rangle = 0. \quad (4.3)$$

La relation (4.3) permet de dériver une équation du mouvement pour $c_{uu}(t) = \langle u(t)u(0) \rangle$:

$$\frac{d}{dt} c_{uu}(t) = - \int_0^t d\tau \zeta(t - \tau) c_{uu}(\tau) \quad (4.4)$$

L'équation (4.4) apparaît comme une généralisation de l'équation différentielle pour une exponentielle. Si l'on pose

$$\zeta(t) = \gamma\delta(t), \quad (4.5)$$

où $\delta(t)$ est la distribution de Dirac, l'équation (4.4) prend la forme d'une équation différentielle simple,

$$\frac{d}{dt} c_{uu}(t) = -\gamma c_{uu}(t). \quad (4.6)$$

On voit facilement que la solution est une exponentielle, $c_{uu}(t) = c_{uu}(0) \exp(-\gamma t)$. La forme (4.5) de la fonction mémoire signifie qu'on regarde une dynamique sans mémoire, ce qui définit le régime de la dynamique brownienne [14, 47].

4.2 Dynamique brownienne fractionnaire

Dans le cadre de la dynamique brownienne fractionnaire (DBF) l'évolution d'une fonction de corrélation est décrite par une généralisation de l'exponentielle étirée

$$\psi_{DBF}(t) = E_{\beta}(-[t/\tau]^{\beta}), \quad 0 < \beta \leq 1 \quad (4.7)$$

où $E_{\beta}(z)$ est la fonction de *Mittag-Leffler* [48],

$$E_{\beta}(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(1 + \beta k)}. \quad (4.8)$$

Ici $\Gamma(z)$ est la fonction de Gamma [48, 49] et $\phi_{DBF}(0) = 1$. La fonction mémoire a la forme [50]

$$\zeta_{DBF}(t) = \frac{\beta - 1}{\Gamma(\beta)\tau^2} \left(\frac{t}{\tau}\right)^{\beta-2}, \quad t > 0, \quad 0 < \beta < 1 \quad (4.9)$$

Il est important de noter que le point $t = 0$ est exclu ci-dessus. Une particularité de cette fonction mémoire est que son intégrale est nulle, bien qu'elle devienne singulière en $t \rightarrow 0$,

$$\int_0^{\infty} dt \zeta_{DBF}(t) = 0. \quad (4.10)$$

Ceci signifie que $\zeta(t)$ est effectivement une distribution au sens mathématique.

On note finalement que le spectre de Fourier de $\phi_{DBF}(t)$ a la forme [50]

$$S_{DBF}(\omega) = \frac{2\tau \sin(\beta\pi/2)}{|\omega\tau| (|\omega\tau|^{\beta} + 2 \cos(\beta\pi/2) + |\omega\tau|^{-\beta})}, \quad 0 < \beta \leq 1 \quad (4.11)$$

Dans la limite $\beta \rightarrow 1$ on s'approche d'une lorentzienne qui est la transformée de Fourier d'une exponentielle "normale":

$$S_{DB}(\omega) = \lim_{\beta \rightarrow 1} S_{DBF}(\omega) = \frac{2\tau}{1 + (\omega\tau)^2}. \quad (4.12)$$

Ici l'indice *DB* indique "dynamique brownienne".

4.3 Approche numérique

Dans la suite on verra comment la dynamique interne d'une protéine peut être analysée par une combinaison de simulation moléculaire et traitement numérique du signal qui permet notamment d'extraire la fonction mémoire [51]. Les méthodes décrites dans la suite sont intégrées dans le code nMoldyn [18].

On considère une quantité dynamique discrète $u(j) \equiv u(j\Delta t)$ quelconque et on fait l'hypothèse que l'équation du mouvement pour la variable dynamique $u(j)$ a la forme

$$u(j) = \sum_{n=1}^P a_n^{(P)} u(j-n) + \epsilon_P(j) \quad (4.13)$$

Il s'agit d'une équation du mouvement stochastique qui décrit la dynamique de la variable $u(t)$ en question par un modèle *autorégressif* (AR) de l'ordre P . La fonction $\epsilon_P(j)$ est un bruit blanc avec $\langle \epsilon_P(j) \rangle = 0$ et $\langle \epsilon_P^2(j) \rangle = \sigma_P^2$. Comme $\langle \epsilon_P(j) u(j-k) \rangle = 0$ pour $k > 0$ il suit que

$$\sum_{n=1}^P c_{uu}(|k-n|) a_n^{(P)} = c_{uu}(k), \quad (4.14)$$

où on suppose que $u(j)$ est un processus stochastique stationnaire. Les eqs. (4.14) constituent un système de P équations linéaires pour les P coefficients $a_1^{(P)}, \dots, a_P^{(P)}$. L'amplitude de $\epsilon_P(j)$ est donnée par

$$\sigma_P^2 = c_{uu}(0) - \sum_{n=1}^P a_n^{(P)} c_{uu}(n). \quad (4.15)$$

On note que (4.13) et (4.14) correspondent, respectivement, à l'équation de Langevin généralisée (4.2) et l'équation de la fonction mémoire (4.4).

Les zéros z_1, \dots, z_P du polynôme caractéristique

$$p(z) = z^P - \sum_{k=1}^P a_k^{(P)} z^{P-k} \quad (4.16)$$

jouent un rôle central pour le modèle AR, car ils permettent d'exprimer d'une façon simple la fonction d'autocorrélation $c_{uu}(n)$ et son spectre $S_{uu}(\omega)$. Ici on note uniquement les résultats essentiels. Tous les détails sont donnés dans la référence [51].

1. Dans le cadre du modèle AR $c_{uu}(n)$ a une forme *multiexponentielle*

$$c_{uu}(n) = \sum_{k=1}^P \beta_k z_k^{|n|} \quad (4.17)$$

où $\{z_k\}$ sont les zéros du polynôme (4.16) et les coefficients β_k sont déterminés par les $\{z_k\}$ et par σ_P ,

$$\beta_k = \frac{1}{a_P^{(P)}} \frac{-z_k^{P-1} \sigma_P^2}{\prod_{j=1, j \neq k}^P (z_k - z_j) \prod_{l=1}^P (z_k - z_l^{-1})}. \quad (4.18)$$

Un modèle AR stable est caractérisé par $|z_k| < 1$, $k = 1, \dots, P$. Afin de garantir cette stabilité on peut utiliser l'algorithme de Burg [23, 52, 53] au lieu de résoudre les équations (4.14). Il convient d'écrire les pôles sous la forme

$$z_k = \exp(-[i\omega_k + \eta_k]\Delta t), \quad (4.19)$$

où $\eta_k > 0$. Le modèle AR d'ordre P donne donc P fréquences, ω_k , et P constantes de relaxation, η_k . Comme les fonctions de corrélation en question sont réelles, il existe à chaque pôle complexe un pôle conjugué complexe. Toute paire de pôles conjugués complexes $\{z_k, z_k^*\}$ donne alors une contribution

$$z_k^{|n|} + z_k^{*|n|} = 2 \exp(-\eta_k |n| \Delta t) \cos(\omega_k n \Delta t) \quad (4.20)$$

à la fonction de corrélation en question, et cette contribution est pondérée par le coefficient β_k donné ci-dessus.

2. Le spectre $S_{uu}(\omega)$ a une forme rationnelle en $\exp[i\omega\Delta t]$ (*all-pole form*)

$$S_{uu}(\omega) = \frac{\sigma_P^2}{\left(1 - \sum_{k=1}^P a_k^{(P)} e^{-ik\omega\Delta t}\right) \left(1 - \sum_{l=1}^P a_l^{(P)} e^{il\omega\Delta t}\right)} \quad (4.21)$$

3. L'équation de la fonction mémoire discrétisée

$$\frac{c_{uu}(n+1) - c_{uu}(n)}{\Delta t} = - \sum_{k=0}^n \Delta t \zeta(n-k) c_{uu}(k). \quad (4.22)$$

peut être résolue par rapport à $\zeta(n)$ en utilisant le modèle AR pour $c_{uu}(n)$. Ici la transformée en z joue le même rôle que la transformation de Laplace pour les fonctions continues. Avec la définition

$$F_{>}(z) = \sum_{n=0}^{\infty} f(n) z^{-n} \quad (4.23)$$

pour la transformation en z unilatérale d'une fonction discrète $f(n)$ quelconque on trouve que

$$\Xi_{>}(z) = \frac{1}{\Delta t^2} \left(\frac{z}{C_{uu,>}(z)} + 1 - z \right), \quad (4.24)$$

avec $\psi(0) = 1$. Or, $C_{uu,>}(z)$ peut être obtenue dans le cadre du modèle autorégressif:

$$C_{uu,>}(z) = \sum_{j=1}^P \beta_j \frac{z}{z - z_j}, \quad |z| > |z_j|_{\max}. \quad (4.25)$$

Insertion de cette expression dans (4.24) permet d'obtenir $\zeta(n)$ par division polynomiale, en utilisant la définition (4.23) pour la transformée en z unilatérale [51].

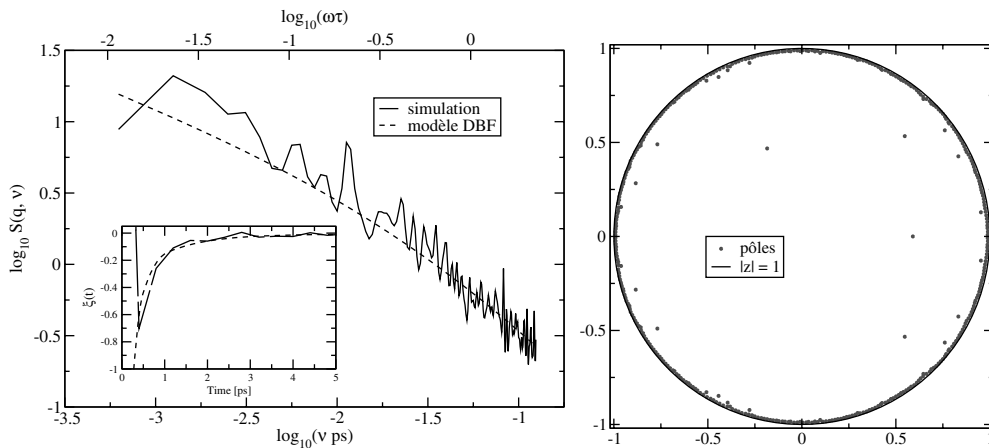


Figure 9. **A gauche:** le spectre $S_{coh,AR}(\mathbf{q}, \omega)$ (ligne continue) et modèle DBF ajusté (ligne pointillée). L'insert montre la fonction mémoire correspondante et le fit par le modèle DBF. **A droite:** pôles du modèle AR.

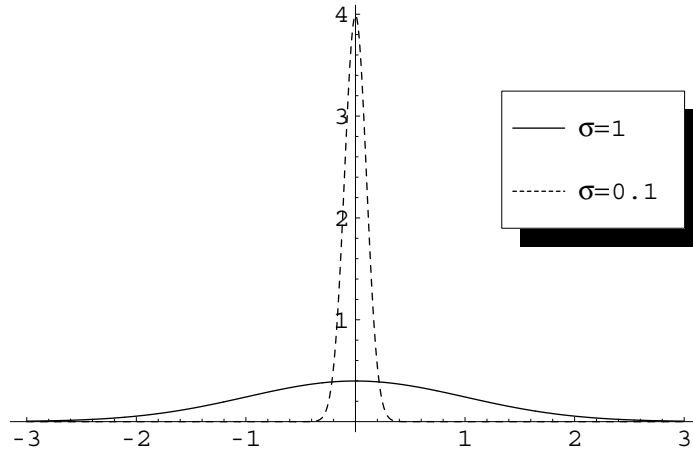


Figure 10. Une gaussienne normalisée pour deux largeurs différentes.

4.4 Dynamique du lysozyme

On considère maintenant une application de la méthode précédemment décrite afin de calculer le facteur de structure cohérent et la fonction mémoire associé. Dans ce cas la variable dynamique est

$$u(t) = \sum_{\alpha=1}^N b_{\alpha,coh} \exp(i\mathbf{q}\cdot\mathbf{R}_{\alpha}(t)) - \left\langle \sum_{j=1}^N b_{j,coh} \exp(i\mathbf{q}\cdot\mathbf{R}_j(t)) \right\rangle. \quad (4.26)$$

Apart une constante, qui joue le rôle de l'EISF pour la diffusion cohérente, $c_{uu}(t)$ est la fonction intermédiaire de diffusion cohérente. Elle est l'équivalent de la fonction $\mathcal{I}'_{inc}(\mathbf{q}, t)$ (voir l'éq. (3.30)) pour la diffusion cohérente. On a donc la correspondance

$$\mathcal{I}'_{coh,AR}(\mathbf{q}, t) \equiv c_{uu}(t), \quad \mathcal{S}'_{coh,AR}(\mathbf{q}, \omega) \equiv S_{uu}(\omega). \quad (4.27)$$

La figure 9 (partie gauche) montre $S_{uu}(q, \omega)$ pour $q = 10 \text{ nm}^{-1}$, la fonction mémoire associée (insert) et la localisation des pôles z_k dans le plan complexe (partie droite). On voit que tous les pôles sont situés à l'intérieur du cercle trigonométrique, $|z| = 1$, ce qui montre que le modèle auto-régressif ajusté est stable. Les lignes pointillées représentent un fit *simultanée* de la fonction mémoire et du spectre du modèle DBF aux fonctions correspondantes simulées. Non seulement le spectre, mais également la fonction mémoire sont bien représentés par le modèle DBF. On remarque dans ce contexte que l'exponentielle étirée ne fait pas partie des modèles auxquels on peut associer une fonction mémoire [50]. Elle ne vérifie pas la relation (4.4) qui représente une condition nécessaire pour la validité d'un modèle pour une fonction de corrélation. L'accord entre la fonction mémoire simulée et celle d'un *modèle* analytique donne un support supplémentaire pour la validité du modèle. Il montre aussi comment la simulation moléculaire peut contribuer au développement de modèles analytiques qui reflètent l'essentiel de la dynamique d'une protéine.

APPENDIX

A Distribution de Dirac

La distribution de Dirac est une distribution au sens mathématique. Elle est définie par les propriétés

$$\int_{-\infty}^{+\infty} dx \delta(x) = 1, \quad (A.1)$$

$$\int_{-\infty}^{+\infty} dx \delta(y-x) f(x) = f(y), \quad (\text{A.2})$$

$$\delta(x) = \delta(-x). \quad (\text{A.3})$$

On suppose que $f(x)$ est une fonction continue. La distribution de Dirac peut être représentée par une gaussienne normalisée dont la largeur tend vers zéro:

$$\delta(x) = \lim_{\sigma \rightarrow 0} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right). \quad (\text{A.4})$$

Références

- [1] H.P. Lu, L. Xun, and X.S. Xie Single molecule enzymatic dynamics. *Science*, 282:1877–1882, 1998.
- [2] H. Yang, G. Luo, P. Karnchanaphanurach, T.M. Louie, I. Rech, S. Cova, L. Xun, and X.S. Xie. Protein conformational dynamics probed by single-molecule electron transfer. *Science*, 302(5643):262–266, 2003.
- [3] H. Yang and X.S. Xie. Probing single-molecule dynamics photon by photon. *J. Chem. Phys.*, 117(24):10965–10979, 2002.
- [4] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, and M. Karplus. Charmm: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comp. Chem.*, 4:187–217, 1983.
- [5] A.D. MacKerell, D. Bashford, M. Bellott, R.L. Dunbrack, J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. J. McCarthy, L. Kuchnic, K. Kuczera, F.T.K. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguien, B. Prudhom, W.E. Reicher, B. Roux, M. Schlenkirch, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiokiewicz-Kuczera, D. Yin, and M. Karplus. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B*, 102:3586–3616, 1998.
- [6] BIOMOS b.v., Laboratory of Physical Chemistry, ETH Zurich. *The GROMOS96 users manual*.
- [7] W.D. Cornell, P. Cieplak, C.I. Bayly, I.R. I.R. Gould, K.M. Merz Jr, D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, and P.A. Kollman. A second generation force field for the simulation of proteins and nucleic acids. *J. Am. Chem. Soc.*, 117:5179, 1995.
- [8] L. Verlet. Computer “experiments” on classical fluids. Thermodynamic properties of Lennard-Jones molecules. *Phys. Rev.*, 159(1):98–103, 1967.
- [9] M.P. Allen and D.J. Tildesley. *Computer Simulation of Liquids*. Oxford University Press, Oxford, 1987.
- [10] J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis*. Springer, Berlin, New York, 1980.
- [11] D. Wolf, P. Keblinski, S.R. Philpot, and J. Eggebrecht. Exact method for the simulation of coulombic systems by spherically truncated, pairwise r⁻¹ summation. *J. Chem. Phys.*, 110(17):8254–8282, 1999.
- [12] T. Hill. *Statistical Mechanics: Principles and Selected Applications*. McGraw Hill, 1956.
- [13] R. Becker. *Theorie der Wärme*. Springer, 1978.
- [14] D.A. McQuarrie. *Statistical Mechanics*. Harper’s Chemistry Series. Harper Collins Publishers, 1976.
- [15] H.C. Andersen. Molecular dynamics at constant pressure and/or constant temperature. *J. Chem. Phys.*, 72(4):2384–2393, 1980.
- [16] S. Nosé. A unified formulation of the constant temperature molecular dynamics methods. *J. Chem. Phys.*, 81(1):511–519, 1984.

- [17] M. Ferrario. *Thermodynamic constraints*, volume 397 of *Computer simulation in chemical physics, NATO ASI series C (Mathematical and Physical Sciences)*, pages 153–171. Kluwer Academic Publishers, 1993.
- [18] T. Rog, K. Murzyn, K. Hinsén, and G.R. Kneller. nMoldyn: A Program Package for a Neutron Scattering Oriented Analysis of Molecular Dynamics Simulations. *J. Comp. Chem.*, 24(5):657–667, 2003.
- [19] W.G. Hoover. *Time Reversibility, Computer Simulation, and Chaos*, volume 13 of *Advanced Series in Nonlinear Dynamics*. World Scientific, 1999.
- [20] E.O. Brigham. *The Fast Fourier Transform*. Prentice Hall, 1974.
- [21] G.R. Kneller, V. Keiner, M. Kneller, and M. Schiller. Nmoldyn, a program package for the calculation and analysis of neutron scattering spectra from MD simulations. *Comp. Phys. Comm.*, 91:191–214, 1995. Full description in report ILL95KN02T, Institut Laue-Langevin, 156 X, F-38042 Grenoble Cedex, France.
- [22] S.W. Lovesey. *Theory of Neutron Scattering from Condensed Matter*, volume I. Clarendon Press, Oxford, 1984.
- [23] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw Hill, 3rd edition, 1991.
- [24] A. Papoulis. *Signal Analysis*. McGraw Hill, 1984.
- [25] H.J.C. Berendsen, J.R. Grigera, and T.P. Straatsma. The missing term in effective pair potentials. *J. Phys. Chem.*, 91:6269–6271, 1987.
- [26] K. Hinsén and G.R. Kneller. A simplified force field for describing vibrational protein dynamics over the whole frequency range. *J. Chem. Phys.*, 111(24):10766–10769, 1999.
- [27] K. Hinsén. Analysis of domain motions by approximate normal mode calculations. *Proteins (Structure, Function and Genetics)*, 33:417–429, 1998.
- [28] D. Chandler. *Introduction to Modern Statistical Mechanics*. Oxford University Press, New York, Oxford, 1987.
- [29] J.P. Boon and S. Yip. *Molecular Hydrodynamics*. McGraw Hill, 1980.
- [30] G.R. Kneller. Superposition of molecular structures using quaternions. *Mol. Sim.*, 7:113–119, 1991.
- [31] K. Hinsén and G.R. Kneller. The influence of geometrical constraints on the dynamics of polypeptide chains. *Phys. Rev. E*, 52(6):6868–6874, 1995.
- [32] M. Bée. *Quasielastic Neutron Scattering: Principles and Applications in Solid State Chemistry, Biology and Materials Science*. Adam Hilger, Bristol, 1988.
- [33] G.R. Kneller and J.C. Smith. Liquid-like side-chain motions in myoglobin. *J. Mol. Biol.*, 242:181–185, 1994.
- [34] W. Doster, S. Cusack, and W. Petry. Dynamical transition of myoglobin revealed by inelastic neutron scattering. *Nature*, 337:754–756, 1989.
- [35] A. Rahman, K.S. Singwi, and A. Sjolander. Theory of slow neutron scattering by liquids. I. *Phys. Rev.*, 126(3):986–996, 1962.
- [36] K. Hinsén, A.-J. Petrescu, S. Dellerue, M.C. Bellissent-Funel, and G.R. Kneller. Harmonicity in slow protein dynamics. *Chem. Phys.*, 261(1+2):25–38, 2000. Special Issue “Condensed Phase Structure and Dynamics: A combined neutron scattering and molecular modelling approach”.
- [37] G.R. Kneller. Neutron scattering from classical systems: Stationary phase approximation of the scattering law. *Mol. Phys.*, 83(1):63–87, 1994.
- [38] A. Rahman. Correlations in the motion of liquid argon. *Phys. Rev.*, 136(2A):405–411, 1964.
- [39] D.W. Davidson and R.H. Cole. Dielectric relaxation on glycerol, propylene glycol, and n-propanol. *J. Chem. Phys.*, 19(12):1484–1490, 1951.
- [40] G. Williams and D. Watts. Non-symmetrical dielectric relaxation behaviour arising from a simple empirical decay function. *Trans. Faraday Soc.*, 66:80–85, 1969.

- [41] C.P. Lindsey and G.D. Patterson. Detailed comparison of the williams-watt and coolde-davidson functions. *J. Chem. Phys.*, 73(7):3348–3357, 1980.
- [42] M.D. Zeidler. The tricomi time correlation function. *Ber. Bunsenges. Phys. Chem.*, 95(9):971–976, 1991.
- [43] I. Köper and M.-C. Bellissent-Funel. Hindered protein dynamics in the presence of a cryprotecting agent. *App. Phys. A.*, 74(1):1257–1259, 2002.
- [44] W.G. Glöckle and T.F. Nonnenmacher. A fractional calculus approach to self-similar protein dynamics. *Biophys. J.*, 68:46–53, 1995.
- [45] B.B. Mandelbrot and J.W. van Ness. Fractional Brownian Motions, Fractional Noises and Applications. *SIAM Review*, 10(4):422–437, 1968.
- [46] R. Zwanzig. *Statistical mechanics of irreversibility*, pages 106–141. Lectures in Theoretical Physics. Wiley-Interscience, New York, 1961.
- [47] N.G. van Kampen. *Stochastic Processes in Physics and Chemistry*. North Holland, Amsterdam, revised edition, 1992.
- [48] A. Erdélyi, W. Magnus, F. Oberhettinger, and F.G. Tricomi. *Higher Transcendental Functions*. McGraw Hill, 1955.
- [49] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions*. Dover Publications, New York, 1972.
- [50] G.R. Kneller and K. Hinsen. Fractional brownian dynamics in proteins. *J. Chem. Phys.*, 121(20):10278–10283, 2004.
- [51] G.R. Kneller and K. Hinsen. Computing memory functions from Molecular Dynamics simulations. *J. Chem. Phys.*, 115(24):11097–11105, 2001.
- [52] J. Burg. *Maximum entropy spectral analysis*. PhD thesis, Stanford University, Stanford (CA), USA, 5 1975.
- [53] J. Makhoul. Stable and efficient lattice methods for linear prediction. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-25(5):423–428, 1977.