

Université d'Orléans

**ÉCOLE DOCTORALE : SANTÉ, SCIENCES  
BIOLOGIQUES ET CHIMIE DU VIVANT**

Centre de Biophysique Moléculaire, CNRS Orléans Rue Charles  
Sadron, 45071 Orléans

**THÈSE** présentée par :

**Abir Nesrine Hassani**

soutenue le : **19 décembre 2023**

pour obtenir le grade de : **Docteur de l'Université d'Orléans**

Discipline/ Spécialité : **Biophysique**

**Modélisation de spectres quasiélastiques de neutrons  
provenant de protéines en solution**

**THÈSE DIRIGÉE PAR :**

**M. Gerald KNELLER**

Professeur à l'Université d'Orléans

**RAPPORTEURS :**

**M. Frank SCHREIBER**

Professeur à l'Université Eberhard Karl de  
Tübingen, Allemagne

**M. Daniel ABERGEL**

Directeur de recherche CNRS, ENS Paris

**JURY :**

**M. Yann VAILLS**

Professeur à l'Université d'Orléans, Président du  
jury

**M. Gerald KNELLER**

Professeur à l'Université d'Orléans

**M. Frank SCHREIBER**

Professeur à l'Université Eberhard Karl de  
Tübingen, Allemagne

**M. Daniel ABERGEL**

Directeur de recherche CNRS, ENS Paris

**M. Jeremy SMITH**

Professeur à l'Université du Tennessee, USA

**Mme. Judith PETERS**

Professeure à l'Université Grenoble Alpes

**M. Andreas STADLER**

Chercheur au Forschungszentrum Jülich et *Privat-  
dozent* RWTH Aachen, Allemagne, invité



# Modélisation de spectres quasiélastiques de neutrons provenant de protéines en solution

## Résumé :

Les protéines sont des systèmes moléculaires complexes dont la dynamique interne est caractérisée par un vaste spectre d'échelles de temps, allant de la picoseconde pour les vibrations des liaisons chimiques à la seconde et au-delà pour les grands changements de conformation. En utilisant un modèle "minimaliste" multi-échelles pour la dynamique de relaxation des protéines, nous montrons ici que même de petits changements dus à des contraintes externes, telles que la température, la modification du solvant ou la liaison d'un ligand, peuvent être élucidés par la diffusion quasi-élastique des neutrons (QENS). La relaxation est ici décrite par une fonction de Mittag-Leffler étirée, qui présente une décroissance lente en loi de puissance pour les temps longs. Un point technique important est l'estimation de le facteur de structure incohérente élastique (EISF) sur la base de sa contrepartie mesurée et des paramètres du modèle de la fonction de relaxation, qui sont le paramètre de forme  $\alpha$  et le paramètre d'échelle de temps  $\tau$ . La première étude concerne la protéine intrinsèquement désordonnée Myelin Basic Protein (MBP) en solution, qui est étudiée dans un solvant D<sub>2</sub>O pur et dans un mélange de D<sub>2</sub>O avec 30% de trifluoroéthanol deutéré à différentes températures, afin d'évaluer l'impact de la formation d'éléments de structure secondaire sur la dynamique interne. Cette analyse a notamment montré que EISF est proche de zéro, ce qui indique des fluctuations de position très importantes, compatibles avec la grande flexibilité interne de cette protéine. Pour confirmer ce résultat une deuxième étude a été réalisée afin de mettre au point une procédure d'ajustement dans laquelle cette quantité est éliminée en tant que paramètre d'ajustement. Les résultats confirment la disparition de l'EISF pour cette protéine. Pour montrer que ce modèle ne conduit pas généralement à la disparition de l'EISF pour les protéines en solution, une analyse de données QENS de la myoglobine a été réalisée, et le résultat montre que l'EISF est clairement non nul pour cette protéine globulaire. La troisième étude est consacré à la compréhension de la dynamique fonctionnelle de l'enzyme Phosphoglycérate kinase. Ici, le modèle permet de déterminer sans ambiguïté l'amplitude des fluctuations inter-domaines qui sont importantes pour sa fonction catalytique.

Mots clés : Diffusion quasiélastiques de neutrons, dynamique des protéines, protéine intrinsèquement désordonnée, analyse de données.

## Modeling quasielastic neutron scattering spectra from proteins in solution

### Abstract :

Proteins are complex molecular systems whose internal dynamics is characterized by a vast spectrum of time scales, ranging from sub-picoseconds for vibrations of chemical bonds to seconds and beyond for large conformational rearrangements. Using a "minimalistic" multi-time scale model for the relaxation dynamics of proteins, we show here that even small changes due to external stress, such as temperature, solvent modification or ligand binding, can be elucidated by quasi-elastic neutron scattering (QENS). The relaxation is here described by a stretched Mittag-Leffler function, which exhibits slow power law decay for long times. An important technical point is the estimation of the elastic incoherent structure factor (EISF) on the basis of its measured counterpart and the model parameters of the relaxation function, which are the  $q$ -dependent form parameter  $\alpha$  and the time scale parameter  $\tau$ . Our first study concerns the intrinsically disordered protein Myelin Basic Protein (MBP) in solution, which is studied in pure D<sub>2</sub>O-buffer and in a mixture of D<sub>2</sub>O-buffer with 30% deuterated Trifluoroethanol at different temperatures, in order to evaluate the impact of formation of secondary structure elements on the internal dynamics. In particular, this analysis showed that EISF is close to zero, indicating very large position fluctuations, which are compatible with the high internal flexibility of this protein. To confirm this result, a second study was carried out to develop a fitting procedure in which this quantity is eliminated as a fitting parameter. The results confirm vanishing EISF for this protein. To show that this model does not systematically lead to vanishing EISFs for proteins in solution, we have analyzed QENS data from Myoglobin, and the result shows that the EISF is clearly non-zero for this globular protein. The third study is devoted to understanding the functional dynamics of the enzyme Phosphoglycerate kinase. Here the model allows for determining unambiguously the amplitude of the inter-domain fluctuations which are important for its catalytic function.

Keywords : Quasi-elastic neutron scattering, protein dynamics, intrinsically disordered proteins, data analysis.



## Remerciements

*La présentation de ce modeste travail m'offre l'occasion d'exprimer ma profonde gratitude :*

*À mon directeur de thèse Gerald Kneller, je tiens à le remercier grandement, il m'a encadré tout au long de cette thèse et m'a fait partager ses brillantes intuitions, ses qualités humaines, ses compétences pratiques et sa contribution à la réalisation de ce travail, qu'il soit aussi remercié pour sa gentillesse, sa disponibilité permanente et pour les nombreux encouragements qu'il m'a prodigués.*

*À mon coencadrant Andreas Stadler, j'ai apprécié sa disponibilité, sa simplicité, sa rigueur et son souci du travail bien fait, ses nombreux conseils ne m'ont jamais fait défaut, je suis heureuse de lui exprimer ma respectueuse reconnaissance.*

*J'adresse mes remerciements à Frank Schreiber et Daniel Abergel de l'honneur qu'ils m'ont fait en acceptant d'être rapporteurs de cette thèse.*

*Je remercie Yann Vaills pour toutes les facilités accordées et d'avoir bien voulu présider le jury de thèse et aussi Jeremy Smith et Judith Peters d'avoir accepté d'en faire partie.*

*Merci à Konrad Hinsén pour son aide et ses précieux conseils.*

*Je remercie vivement la Région Centre-Val de Loire et Jülich Centre for Neutron Science (JCNS) pour leur soutien financier.*

*Un grand merci à mes parents qu'aucun langage ne saurait exprimer mon respect et ma considération pour leur soutien et encouragement afin que je puisse atteindre mes objectifs et à mon cher frère Moundhir pour son soutien moral et ses conseils précieux et surtout de m'avoir aidé et supporté dans les moments difficiles.*

*À toute ma famille pour leur indéfectible soutien et leur patience infinie, et particulièrement mon oncle Mokhtar Zater à qui je suis redevable, il a développé en moi le désir et la volonté d'un continuel progrès et m'a accordé son soutien dans les instants les plus difficiles.*

*Je ne peux clore cette liste sans remercier mes amis(ies) qui m'ont soutenue, encouragée et ont participé à ma réussite et partagé avec moi les moments qui émaillent un parcours de thèse et je suis comblée qu'ils en voient l'aboutissement.*

## Table des matières

Remerciements	5
Chapitre 1. Introduction	11
Chapitre 2. Diffusion de neutrons	15
1. Travaux pionniers	15
1.1. La découverte du neutron	15
1.2. Propriétés uniques	16
2. Théorie de la diffusion de neutrons	18
2.1. Expérience de diffusion	18
2.2. Facteur de structure dynamique cohérent et incohérent	23
2.3. Condition du bilan détaillée et la limite classique	24
2.4. Systèmes riches en hydrogène	25
2.5. Fonctions de Van Hove	26
2.6. La diffusion quasiélastique de neutrons (QENS)	29
2.7. Facteur de structure élastique incohérent (EISF)	29
Chapitre 3. Systèmes étudiés et motivation du travail	31
1. Relation structure-fonction et protéines intrinsèquement désordonnées	31
1.1. Relation structure-fonction	31
1.2. Protéines intrinsèquement désordonnées	32
2. Protéines étudiés dans la thèse	33
2.1. Protéine Basique de la Myeline (MBP)	34

2.2. Myoglobine (Mb)	35
2.3. Phosphoglycérate kinase (PGK)	35
3. Données QENS et spectromètres	36
3.1. Spectromètre IN16B	36
3.2. Spectromètre IN5.	37
Chapitre 4. Analyse de données	39
1. Concept général	39
1.1. Analyse en temps par transformée de Fourier	40
1.2. La soustraction des solvants	41
1.3. Symétrisation et normalisation des spectres QENS	42
1.4. Extraction et déconvolution de $F(q,t)$	43
2. Modèles pour la dynamique des protéines	45
2.1. Relaxation multi-échelle dans des systèmes complexes	45
2.2. Fonction intermédiaire de diffusion pour la dynamique interne	46
Chapitre 5. Dynamique d'une protéine intrinsèquement désordonnée	53
1. Diffusion dynamique de la lumière	53
2. Dichroïsme circulaire	54
3. Modèle utilisé pour l'analyse de données QENS de la MBP	55
4. Facteur de structure dynamique total	57
5. Résultats de l'analyse de données QENS de la MBP	57
Chapitre 6. EISF mesuré et réel	71
1. Elimination de la diffusion élastique comme paramètre du modèle	71
1.1. Diffusion Pseudo-élastique	72
1.2. La contribution du modèle pseudo-élastique	73
1.3. Calcul de $\xi_{ML}$	74
1.4. Test	78

Table des matières	9
2. Résultats de l'analyse de données QENS de la MBP	79
3. Résultats de l'analyse de données QENS de la myoglobine	82
3.1. Correction de la résolution de $\langle \hat{u}^2 \rangle$	85
Chapitre 7. Dynamique d'une enzyme en solution	89
1. Modèle utilisé pour l'analyse de données QENS de la PGK	90
2. Résultats de l'analyse de données QENS de la PGK	93
Chapitre 8. Conclusion	99
Chapitre 9. Publications	105
1. Multiscale relaxation dynamics and diffusion of myelin basic protein in solution studied by quasielastic neutron scattering	105
2. Quasi-analytical resolution-correction of elastic neutron scattering from proteins	118
3. Signature of functional enzyme dynamics in quasielastic neutron scattering spectra : The case of phosphoglycerate kinase	130
Bibliographie	137



## Chapitre 1

### **Introduction**

L'étude du lien entre la dynamique structurale des protéines et leur fonction est un aspect central de la biologie moléculaire moderne, notamment dans la perspective du développement de nouveaux médicaments et de nouvelles substances actives. Dans ce contexte, la dynamique des protéines est de plus en plus reconnue comme un facteur déterminant pour leur fonction biologique.

Le terme protéine est apparu en 1838 sous l'influence du chimiste néerlandais Gerhardus Johannes Mulder définir un groupe de substances d'origine végétale ou animale possédant des propriétés nutritives et dont l'analyse élémentaire montrait à peu près la même formule empirique en carbone, hydrogène, azote et oxygène selon une règle de proportionnalité. La vision des protéines dans les années 1920 était celle d'une sphère, pour les protéines dites globulaires, et des filaments pour les protéines fibreuses [1]. Le chimiste Anfinsen et ses collaborateurs ont proposé que la séquence d'une protéine contienne les informations nécessaires pour adopter une structure définie. Et d'après le postulat d'Anfinsen ou l'hypothèse thermodynamique, la structure tridimensionnelle de la protéine native dans son milieu physiologique normal est celle dans laquelle l'énergie libre de Gibbs de l'ensemble du système est la plus faible; c'est-à-dire que la conformation native est déterminée par la totalité des interactions interatomiques et donc par la séquence d'acides aminés, dans un environnement donné [2]. À partir des études cristallographiques de protéines, telles que l'hémoglobine et de nombreuses enzymes, Anfinsen et

al ont établi le paradigme structure-fonction [3]. Ce concept postule que la fonction est réalisée par la structure tridimensionnelle unique adoptée par une protéine, qui à son tour est déterminée par sa séquence d'acides aminés. La dynamique interne d'une protéine joue un rôle important pour sa fonction et aussi pour sa structure. Depuis les travaux pionniers de Frauenfelder et ses collaborateurs sur la cinétique de fixation du monoxyde de carbone (CO) par la myoglobine, la relation entre la dynamique et la fonction des protéines est devenue un paradigme en biologie moléculaire et en biophysique [4,5]. Diverses techniques ont été développées au cours des dernières décennies afin d'étudier la dynamique des protéines dans différentes conditions (par exemple Influence de la pression sur la dynamique du lysozyme en solution, par simulation numérique et diffusion de neutrons [6]) et à différentes échelles de temps. La myoglobine a été la première protéine dont la structure été résolue par diffraction de rayon X [7] et des études ont été effectuées par Frauenfelder et al par cette technique afin d'évaluer les amplitudes de mouvement atomiques pour faire apparaitre les fluctuations conformationnels de cette protéine [8]. De différentes techniques spectroscopiques ont été utilisés pour contribuer à bien comprendre la dynamique des protéines. On mentionne ici encore des travaux de Frauenfelder et ses collaborateurs [4,5], la spectroscopie par corrélation de fluorescence [9,10], la spectroscopie diélectrique [11], la résonance magnétique nucléaire (RMN) [12,13] et la diffusion de neutrons [14] qui couvrent les échelles de temps différentes . La dernière technique joue un rôle important, car il s'agit d'une technique spectroscopique particulièrement adaptée à l'étude de la dynamique des protéines qui ne permet non seulement de sonder les échelles de temps des mouvements atomiques (les échelles de temps accessibles se situent entre la picoseconde et la nanoseconde), mais aussi leurs

amplitudes à l'échelle nanométrique. La spectroscopie par diffusion de neutrons utilise les neutrons thermiques dont leurs énergies est comparable à des atomes de la matière condensée et correspond à une longueur d'onde qui est de l'ordre de grandeur des distances inter-atomiques et à des mouvements atomiques ; donc cela permet d'étudier les propriétés structurales et dynamiques des systèmes de la matière condensée au niveau atomique. A cause de la section efficace large des atome d'hydrogène elle sonde en particulier les mouvements des atomes d'hydrogène dans les systèmes riche en hydrogène, comme les protéines, dont près de 50% des atomes des macromolécules biologiques et sont répartis au sein de la structure [15].

Ce travail de thèse est motivé par les progrès de la recherche des trente dernières années au sujet de la relation structure-dynamique-fonction des protéines et par la nécessité de développer des modèles adaptés pour élucider cette relation par diffusion quasiélastique de neutrons. L'aspect central est la modélisation de spectres quasiélastiques de neutrons provenant des protéines en solution, en utilisant un nouveau type de modèle "minimaliste" qui prend en compte le caractère multi-échelle de la dynamique interne de protéines et qui en permet une interprétation physique.



## Chapitre 2

### Diffusion de neutrons

La diffusion de neutrons est l'une des méthodes expérimentales les plus puissantes et les plus polyvalentes pour étudier la structure et la dynamique des matériaux à l'échelle atomique et nanométrique, y compris le magnétisme. Dans les années 1950 et 1960, la diffusion de neutrons était un outil exotique de la physique de l'état solide et de la cristallographie chimique, mais aujourd'hui elle sert des communautés aussi diverses que la biologie, les sciences de la Terre, l'ingénierie et la chimie des polymères. Elle est en particulier fermement établie comme un complément inestimable des études de la structure de la matière par la diffraction des rayons X.

#### 1. Travaux pionniers

**1.1. La découverte du neutron.** Le neutron a été découvert à partir de trois séries d'expériences qui ont été menées dans trois pays différents. En 1930 en Allemagne, les spécialistes des rayons cosmiques W. Bothe et H. Becker ont observé que des éléments légers bombardés par des particules alpha émettaient des rayons "ultra pénétrants" qu'ils ont interprété comme des rayons gamma beaucoup plus énergétiques que ceux émis par des noyaux radioactifs [16]. Ensuite, en 1931, en France, Irène et Frédéric Joliot-Curie reprennent cette étude sur d'autres réactions pour comprendre la nature de ce rayonnement et découvrent qu'il a la propriété de mettre en mouvement des noyaux atomiques, confirmant donc tout en conservant la même interprétation [17]. Et finalement,

en 1932, en Angleterre, J. Chadwick confirme les résultats et va plus loin, par une analyse rigoureuse du recul des noyaux bombardés par ce rayonnement. Il peut affirmer que le rayonnement "ultra pénétrant" ne peut être un rayonnement gamma d'énergie très élevée, mais il s'agit d'une particule neutre et de masse voisine de celle du proton, qu'il nomme neutron [18] et en 1935, il a eu le prix Nobel à cause de cette découverte.

**1.2. Propriétés uniques.** Le neutron est une particule neutre présente dans le noyau de l'atome et qui n'interagit donc pas avec les champs électriques. Il a été utilisé pour la première fois par Clifford Shull en 1946 (prix Nobel de Physique 1994) comme outil pour des expériences de diffraction. Avec le proton, le neutron est un des constituants des noyaux atomiques, qui est constitué d'un quark up et de deux quarks down. Cette combinaison lui laisse une charge électrique nulle pour qu'il pénètre profondément dans la matière, une masse de  $m = 1.660 \times 10^{-24} \text{ g}$ , un spin de  $\frac{1}{2}\hbar$ , et un moment magnétique de  $\mu = -1.913 \mu_N$  (magnéton nucléaire). Un neutron libre n'est pas stable et se désintègre après un temps de vie moyen de 885.6 s en un proton, un électron et un anti-neutrino électronique. Les neutrons utilisés pour la recherche sont produits par un processus de fission dans un réacteur ou par un processus de spallation, qui a l'avantage de pouvoir être arrêté. Les expériences de diffusion de neutrons utilisent particulièrement des neutrons "thermiques", qui ont été ralentis par de multiples collisions avec les atomes ou molécules d'un modérateur (par ex. graphite,  $\text{H}_2\text{O}$ ,  $\text{D}_2\text{O}$ ) qui est tenu à une certaine température. Après le processus de modération, les neutrons ont donc une énergie typique de  $E_{\text{th}} = 3k_B T/2$ , où  $k_B$  est la constante de Boltzmann ( $1.381 \times 10^{-23} \text{ J/K}$ ) et  $T$  est la température en Kelvin. On parle ainsi de neutrons "froids", "thermiques"

et "chauds" pour une température du modérateur de 20 K, 300 K, 1473 K, respectivement. Après modération un neutron a donc une énergie cinétique de

$$E = \frac{|\mathbf{p}|^2}{2m_n} = \frac{3}{2}k_B T \quad (\text{II.1})$$

où  $\mathbf{p}$  est la quantité de mouvement et  $m_n$  la masse du neutron. Avec la dualité particule-onde la quantité de mouvement peut être reliée à la longueur d'onde du neutron,

$$\mathbf{p} = \hbar \mathbf{k} \quad (\text{II.2})$$

$$\mathbf{k} = \frac{2\pi}{\lambda} \mathbf{n}, \quad |\mathbf{n}| = 1 \quad (\text{II.3})$$

Ici  $\hbar = \frac{h}{2\pi}$  où  $h = 6.626176 \text{ Js}$  est la constante de Planck et  $\mathbf{n}$  est un vecteur unitaire.

À partir des relations (II.2) et (II.3), la longueur d'onde d'un neutron thermique est égale à

$$\lambda = \frac{h}{\sqrt{2mE}}. \quad (\text{II.4})$$

On peut écrire l'énergie du neutron en fonction de sa longueur d'onde  $\lambda$ , en utilisant les relations (II.1) et (II.4)

$$E = \frac{h^2}{2m\lambda^2}. \quad (\text{II.5})$$

L'énergie moyenne des neutrons thermiques peut être écrite sous la forme

$$\bar{E} = \frac{1}{2}m\bar{v}^2 = \frac{3}{2}k_B T \text{ où } \bar{v}^2 = \frac{3k_B T}{m}. \quad (\text{II.6})$$

L'énergie du neutron thermique, qui est de l'ordre de  $k_B T = 25 \text{ meV}$  à  $T = 300 \text{ K}$  et par définition comparable à celle des atomes de la matière condensée,

et sa longueur d'onde correspondante de  $\lambda = 1.78 \text{ \AA}$  montrent que la diffusion de neutrons thermiques peut fournir des informations précises sur les aspects structuraux et dynamiques du système étudié. C'est une caractéristique essentielle de la technique de diffusion de neutrons pour fournir des informations précises sur les aspects structuraux et dynamiques du système étudié puisque la diffusion de neutrons thermiques est sensible à la dynamique et à la structure de la matière condensée à l'échelle atomique.

## 2. Théorie de la diffusion de neutrons

**2.1. Expérience de diffusion.** Un faisceau de neutrons monochromatiques incident de vecteur d'onde initial  $k_0$  et d'énergie initiale  $E_0$  interagit avec l'échantillon. Les neutrons diffusés sont caractérisés par un vecteur d'onde final  $k$  et une énergie finale  $E$ . La quantité mesurée dans une expérience de diffusion de neutrons est la section efficace différentielle de diffusion qui est fonction de l'énergie et du transfert d'impulsion sur l'échantillon [15, 19]. Ces quantités sont désignées respectivement par  $\Delta E = E_0 - E$  et  $\Delta \mathbf{p} = \mathbf{p}_0 - \mathbf{p}$ . Les transferts d'énergie et de quantité de mouvement sont exprimés en unités de  $\hbar$

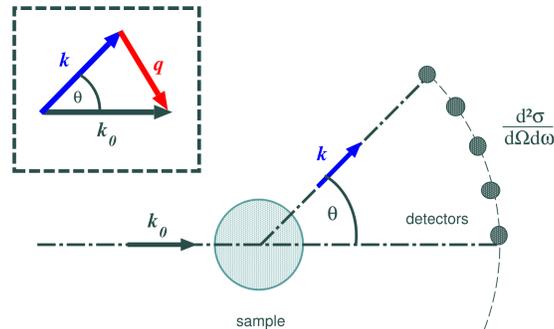


FIGURE II.1. Schéma d'une expérience de diffusion de neutrons [20].

$$\Delta E = \hbar\omega \quad (\text{II.7})$$

$$\Delta \mathbf{p} = \hbar(\mathbf{k}_0 - \mathbf{k}) = \hbar\mathbf{q} \quad (\text{II.8})$$

En utilisant les définitions de (II.7) et (II.8), la section efficace différentielle de diffusion peut être exprimée sous la forme suivante

$$\frac{d^2\sigma}{d\Omega d\omega} = \frac{|\mathbf{k}|}{|\mathbf{k}_0|} S(\mathbf{q}, \omega) \quad (\text{II.9})$$

où  $S(\mathbf{q}, \omega)$  est le facteur de structure dynamique,

$$S(\mathbf{q}, \omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dt e^{-i\omega t} F(\mathbf{q}, t). \quad (\text{II.10})$$

Les neutrons entrants sont caractérisés par une énergie  $E_0$  et une quantité de mouvement  $\hbar\mathbf{k}_0$ , et les neutrons diffusés par une énergie  $E$  et un momentum  $\hbar\mathbf{k}$ . Le facteur de structure dynamique  $S(\mathbf{q}, \omega)$  est la transformée de Fourier temporelle d'une fonction de corrélation quantique qui est la fonction intermédiaire de diffusion,  $F(\mathbf{q}, t)$ , elle contient les informations sur la dynamique structurelle du système considéré,

$$F(\mathbf{q}, t) = \frac{1}{N} \sum_{j,k} \Gamma_{jk} \langle e^{-i\mathbf{q}\cdot\hat{\mathbf{R}}_j(0)} e^{i\mathbf{q}\cdot\hat{\mathbf{R}}_k(t)} \rangle. \quad (\text{II.11})$$

Ici  $N$  est le nombre total d'atomes dans le système de diffusion et pour chaque paire  $\{j, k\}$  d'entre eux,  $\{\hat{\mathbf{R}}_j(t), \hat{\mathbf{R}}_k(t)\}$  désignent les opérateurs de position qui dépendent du temps associés. Le symbole  $\langle \dots \rangle$  représente une moyenne d'ensemble quantique et les facteurs de pondération  $\Gamma_{jk}$  sont de la forme

$$\Gamma_{jk} = \overline{b_j^* b_k} + \delta_{jk} \overline{|b_j - \overline{b_j}|^2}, \quad (\text{II.12})$$

où  $b_j$  et  $b_k$  sont les longueurs de diffusion (complexes) [15,19] qui caractérisent le (pseudo) potentiel d'interaction entre le neutron et l'échantillon,

$$\hat{V}(\hat{\mathbf{r}}) = \frac{2\pi\hbar}{m} b \delta(\hat{\mathbf{r}} - \hat{\mathbf{R}}). \quad (\text{II.13})$$

La moyenne  $\overline{\dots}$  sera expliquée ci-dessous. Les neutrons interagissent avec la matière par une interaction neutron-noyau via des forces nucléaires à courte portée de l'ordre de femtomètre, et aussi magnétiques (avec les moments magnétiques des électrons à travers des couplages dipolaires). Cette interaction peut être décrite par une fonction de Dirac selon Fermi et dans l'éq. (II.13),  $\hat{\mathbf{r}}$  et  $\hat{\mathbf{R}}$  sont respectivement, l'opérateur de position du neutron et du noyau de l'atome diffuseur. La longueur de diffusion  $b$  du noyau prend des valeurs dans la gamme du femtomètre et  $m$  désigne la masse effective du neutron dans l'interaction neutron-noyau.  $\hat{V}(\hat{\mathbf{r}})$  correspond à l'opérateur  $T$  qui est une quantité de la théorie de la diffusion, faisant référence au fait que  $\hat{V}(\hat{\mathbf{r}})$  implique une série de Born complète. Cet opérateur apparait dans la solution de l'équation de Lippmann-Schwinger [21] et pour cette raison  $\hat{V}(\hat{\mathbf{r}})$  est aussi appelé le pseudo-potential de Fermi. La section efficace de diffusion d'un atome fixe est donnée par

$$\sigma = 4\pi|b^2| \quad (\text{II.14})$$

et nous remarquons que la longueur de diffusion du noyau  $b$  change en passant d'une espèce chimique à l'autre, et aussi pour les différents isotopes de la même espèce, car l'interaction neutron-noyau dépend non seulement de la nature du noyau mais aussi du spin total du système ( $J \pm \frac{1}{2}\hbar$ ). Dans une expérience, l'échantillon est composé de différentes espèces atomiques, chacune d'elle pouvant être un mélange de plusieurs isotopes possédant leur propre spin nucléaire. Supposant que les distributions d'isotopes et du spin total pendant l'interaction sont totalement aléatoires et que les atomes  $j$  et  $k$  sont deux atomes différents, nous pouvons écrire la moyenne du produit des deux longueurs de diffusion de cette manière

$$\overline{b_j b_k} = \overline{b_j} \overline{b_k} \quad (\text{II.15})$$

car il n'y a pas de corrélation entre les deux longueurs de diffusion  $b_j$  et  $b_k$ . Dans le cas où  $j=k$ , on a

$$\overline{b_j b_k} = \overline{b_j^2} \quad (\text{II.16})$$

La moyenne  $\overline{\quad}$  ici porte sur tous les isotopes et toutes les projections pour une paire donnée  $(j, k)$  d'atomes, et en supposant que ces propriétés ne sont pas corrélées entre les atomes  $j$  et  $k$ , on peut écrire l'expression de  $b_j$  et  $b_k$  de la forme,

$$\overline{b_j b_k} = \underbrace{\overline{b_j}}_{b_{j,coh}} \underbrace{\overline{b_k}}_{b_{k,coh}} + \delta_{jk} \underbrace{(\overline{b_j^2} - \overline{b_j}^2)}_{b_{j,inc}^2} \quad (\text{II.17})$$

Ici la longueur de diffusion cohérente est

$$b_{j,coh} = \overline{b_j}, \quad (\text{II.18})$$

et la longueur de diffusion incohérente est

$$b_{j,inc} = \sqrt{\overline{b_j^2} - \overline{b_j}^2} \quad (\text{II.19})$$

La section efficace différentielle d'un atome  $j$ , est associée aux longueurs de diffusion cohérente  $b_{j,coh}$  et incohérente  $b_{j,inc}$  par cette relation

$$\sigma_{j,coh} = 4\pi b_{j,coh}^2, \quad (\text{II.20})$$

$$\sigma_{j,inc} = 4\pi b_{j,inc}^2, \quad (\text{II.21})$$

et la section efficace différentielle d'un atome  $j$  totale, peut être écrite de cette manière

$$\sigma_{j,tot} = 4\pi (b_{j,coh}^2 + b_{j,inc}^2). \quad (\text{II.22})$$

Le tableau (1) montre la différence entre les longueurs de diffusion et les sections efficaces cohérentes et incohérentes de quelques éléments chimiques [15]. Comparant les sections efficaces de chaque éléments, on remarque que la section efficace incohérente de l'atome d'hydrogène  $^1H$  est très grande et

Élément	A	Spin	$b_{coh}(fm)$	$b_{inc}(fm)$	$\sigma_{coh}(barn)$	$\sigma_{inc}(barn)$
H	1	$\frac{1}{2}$	-3.741	<b>25.217</b>	1.759	<b>79.90</b>
	2	1	6.674	4.033	5.597	2.04
C	12	0	6.653	0	5.563	0
O	16	0	5.805	0	4.235	0
Al	27	$\frac{5}{2}$	3.449	0.26	1.495	0.008
S	32	0	2.804	6.29	0.988	0
V	51	$\frac{7}{2}$	-0.402	6.419	0.020	5.178

TABLE 1. Longueurs de diffusion et sections efficaces de certains éléments chimiques [15].

plus importante à celle de tous les autres éléments (elle est beaucoup plus importante que son isotope le deutérium  $^2H$  qui est un diffuseur incohérent plus faible). Pour cette raison la diffusion de neutrons sonde les mouvements des atomes d'hydrogène dans les systèmes qui sont riches en hydrogène. En raison de la section de diffusion incohérente dominante de l'hydrogène, les échantillons de la matière molle (polymères, biomolécules), qui contiennent généralement 50% d'atomes d'hydrogène, diffusent de manière essentiellement incohérente, ce qui fait que la diffusion incohérente de neutrons est un outil privilégié pour les études des composés hydrogénés. Les atomes d'hydrogène dans une protéine sont distribués de manière homogène, et les expériences de diffusion de neutrons donnent donc une vue moyenne de la dynamique des protéines. La plupart des expériences de diffusion de neutrons sur les protéines ont été réalisées avec des protéines hydrogénées, soit dans des solutions de  $D_2O$  ou dans des poudres hydratées de  $D_2O$ .

Comme les rayons X interagissent avec le nuage électronique, leurs sections efficaces de diffusion augmentent avec le numéro atomique  $Z$ , au contraire aux sections efficaces de diffusion des neutrons qui ne présentent pas de variations systématiques avec  $Z$ . La section efficace la plus importante pour la diffusion de neutrons est même celle de l'élément le plus léger – l'hydrogène (voir fig. II.2). Il est donc clair qu'il sera difficile de déterminer les positions de l'hydrogène avec les rayons X en présence d'éléments lourds, et pour cette raison on utilise la diffraction de neutrons pour "voir" les positions des atomes d'hydrogène.

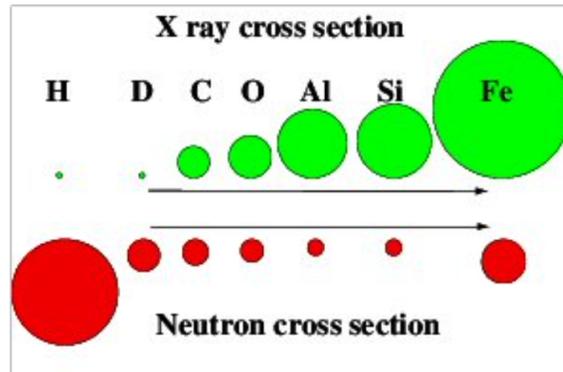


FIGURE II.2. Comparaison des sections efficaces de diffusion de neutrons et des rayons X [22].

**2.2. Facteur de structure dynamique cohérent et incohérent.** Il suit de la forme (II.12) des poids  $\Gamma_{jk}$  que la fonction intermédiaire de diffusion peut être divisée en deux parties dites cohérente et incohérente, respectivement,

$$F(\mathbf{q}, t) = F_{\text{coh}}(\mathbf{q}, t) + F_{\text{inc}}(\mathbf{q}, t), \quad (\text{II.23})$$

qui sont données par

$$F_{\text{coh}}(\mathbf{q}, t) = \frac{1}{N} \sum_{j,k} \bar{b}_j^* \bar{b}_k \langle e^{-i\mathbf{q}\cdot\hat{\mathbf{R}}_j(0)} e^{-i\mathbf{q}\cdot\hat{\mathbf{R}}_k(t)} \rangle, \quad (\text{II.24})$$

$$F_{\text{inc}}(\mathbf{q}, t) = \frac{1}{N} \sum_j (\bar{b}_j^2 - \bar{b}_j^2) \langle e^{-i\mathbf{q}\cdot\hat{\mathbf{R}}_j(0)} e^{-i\mathbf{q}\cdot\hat{\mathbf{R}}_j(t)} \rangle. \quad (\text{II.25})$$

La diffusion cohérente décrit la dynamique collective des atomes dans le système de diffusion, tandis que la diffusion incohérente décrit la dynamique d'atomes individuels, moyennée sur le système.

La décomposition de la fonction intermédiaire en une partie cohérente et une partie incohérente se traduit pour le le facteur de structure dynamique en

$$S(\mathbf{q}, \omega) = S_{\text{coh}}(\mathbf{q}, \omega) + S_{\text{inc}}(\mathbf{q}, \omega) \quad (\text{II.26})$$

avec

$$S(\mathbf{q}, \omega) = \underbrace{\frac{1}{2\pi} \int_{-\infty}^{+\infty} dt e^{-i\omega t} F_{\text{coh}}(\mathbf{q}, t)}_{S_{\text{coh}}(\mathbf{q}, \omega)} + \underbrace{\frac{1}{2\pi} \int_{-\infty}^{+\infty} dt e^{-i\omega t} F_{\text{inc}}(\mathbf{q}, t)}_{S_{\text{inc}}(\mathbf{q}, \omega)}. \quad (\text{II.27})$$

**2.3. Condition du bilan détaillée et la limite classique.** Comme  $F(\mathbf{q}, t)$  est une fonction de corrélation temporelle quantique, le facteur de structure dynamique n'est pas symétrique en  $\omega$ . La fonction intermédiaire de diffusion remplit les relations de symétrie d'une fonction de corrélation temporelle quantique,

$$F^*(\mathbf{q}, t) = F(\mathbf{q}, -t), \quad (\text{II.28})$$

$$F(\mathbf{q}, -t) = F(-\mathbf{q}, t + i\beta\hbar), \quad (\text{II.29})$$

où  $\beta = 1/k_B T$  est la température de Boltzmann inverse. Pour le facteur de structure dynamique l'équation II.29 devient la relation du bilan détaillée

$$S(\mathbf{q}, \omega) = e^{\beta\hbar\omega} S(-\mathbf{q}, -\omega). \quad (\text{II.30})$$

Nous rappelons que le transfert d'énergie de signe positif  $\hbar\omega > 0$  est un gain d'énergie de l'échantillon et donc une perte d'énergie par le neutron et inversement un transfert d'énergie  $\hbar\omega < 0$  est donc un gain d'énergie par le neutron. La relation de bilan détaillé exprime que la probabilité de perte d'énergie par le neutron est plus grande d'un facteur  $e^{\beta\hbar\omega}$  que le gain d'énergie.

**2.4. Systèmes riches en hydrogène.** La section efficace de diffusion incohérente de l'hydrogène est largement supérieure par rapport aux autres éléments chimiques (voir tableau (1)). Considérons un échantillon de protéine, c'est à-dire des systèmes riches en hydrogène, la diffusion est essentiellement incohérente et sonde les mouvements individuels des atomes d'hydrogène, on peut calculer approximativement

$$F(\mathbf{q}, t) \approx F_{H,\text{inc}}(\mathbf{q}, t) \equiv |b_{H,\text{inc}}|^2 F_H(\mathbf{q}, t). \quad (\text{II.31})$$

où  $b_{H,\text{inc}} = \sqrt{b_H^2 - \bar{b}_H^2}$  est la longueur de diffusion incohérente et

$$F_H(\mathbf{q}, t) = \frac{1}{N} \sum_{j \in H} \langle e^{-i\mathbf{q} \cdot \hat{\mathbf{R}}_j(0)} e^{i\mathbf{q} \cdot \hat{\mathbf{R}}_j(t)} \rangle \quad (\text{II.32})$$

où  $j$  passe sur tous les atomes d'hydrogène. Et on peut définir le facteur de structure dynamique correspondant

$$S_H(\mathbf{q}, \omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dt e^{-i\omega t} F_H(\mathbf{q}, t). \quad (\text{II.33})$$

Dans la figure II.3 qui montre un schéma d'un spectre de diffusion incohérente, on peut distinguer 3 régions :

- **Diffusion élastique** - aucun échange d'énergie avec le système  $\hbar\omega = 0$ .
- **Diffusion inélastique** - due aux mouvements vibrationnels. les pics inélastiques proviennent de la diffusion d'atomes qui vibrent de manière

périodique et à une fréquence fixe. Les neutrons diffusés inélastiquement peuvent perdre de l'énergie (diffusion Stokes) ou en gagner (diffusion antiStokes).

- **Diffusion quasi-élastique** - centré sur  $\omega = 0$  et décrit les mouvements stochastiques.

Le terme QENS désigne les processus inélastiques qui sont presque élastiques. La diffusion quasiélastique se situe dans la région de la distribution en énergie des neutrons qui correspond à de faibles transferts d'énergie avec les atomes, de l'ordre de  $+/- 2meV$ .

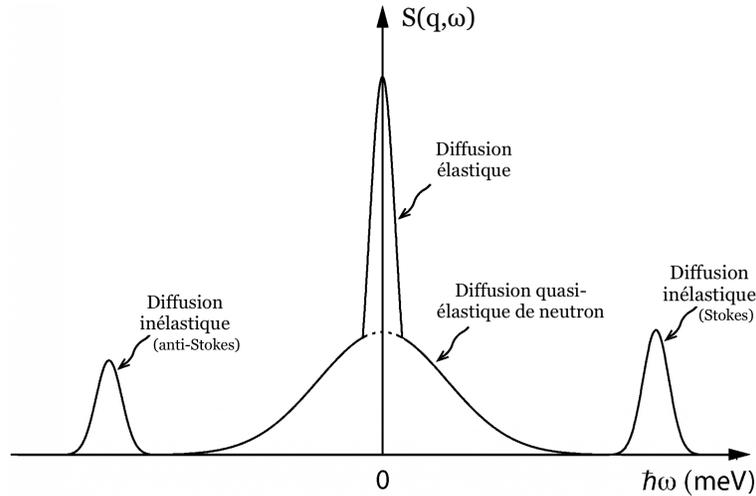


FIGURE II.3. Schéma d'un spectre de diffusion de neutrons, contenant les parties élastiques, inélastiques et quasi-élastiques [23].

**2.5. Fonctions de Van Hove.** L'idée de Van Hove était de passer d'une interprétation orientée vers les expériences de diffusion de neutrons dans l'espace  $(\mathbf{q}, \omega)$  à une description dans les variables de position et de temps  $(\mathbf{r}, t)$  en introduisant la transformée de Fourier spatiale de la fonction intermédiaire

de diffusion [24]

$$G(\mathbf{r}, t) = \frac{1}{N} \sum_{j,k} \frac{1}{(2\pi)^3} \int d^3 q e^{-i\mathbf{q}\cdot\mathbf{r}} \langle e^{-i\mathbf{q}\cdot\hat{\mathbf{R}}_j(0)} e^{-i\mathbf{q}\cdot\hat{\mathbf{R}}_k(t)} \rangle. \quad (\text{II.34})$$

Cette fonction dite de Van Hove est reliée au facteur de structure dynamique  $S(\mathbf{q}, \omega)$  par la double transformée de Fourier

$$S(\mathbf{q}, \omega) = \frac{1}{2\pi} \int \int d^3 q dt e^{i(\mathbf{q}\cdot\mathbf{r}-\omega t)} G(\mathbf{r}, t) \quad (\text{II.35})$$

et exprime la relation entre les descriptions  $(\mathbf{r}, t)$  et  $(\mathbf{q}, \omega)$  de la diffusion des neutrons. On omet ici les longueurs de diffusion afin de souligner la signification physique. Utilisant que

$$e^{i\mathbf{q}\cdot\hat{\mathbf{R}}_k(t)} = \int d^3 \mathbf{r} e^{i\mathbf{q}\cdot\mathbf{r}} \delta(\mathbf{r} - \hat{\mathbf{R}}_k(t)) \quad (\text{II.36})$$

et la définition de la distribution de Dirac,

$$\delta(\mathbf{r}) = \frac{1}{(2\pi)^3} \int d^3 \mathbf{q} e^{\pm i \mathbf{q}\cdot\mathbf{r}}, \quad (\text{II.37})$$

on obtient

$$G(\mathbf{r}, t) = \frac{1}{N} \sum_{j,k} \int d^3 \mathbf{r}' \langle \delta(\mathbf{r} - \mathbf{r}' + \hat{\mathbf{R}}_j(0)) \delta(\mathbf{r}' - \hat{\mathbf{R}}_k(t)) \rangle, \quad (\text{II.38})$$

sachant que les opérateurs  $\hat{\mathbf{R}}_j(0)$  et  $\hat{\mathbf{R}}_k(t)$  dans l'équation (II.34) ne commutent pas pour  $t \neq 0$ . La fonction de Van Hove est divisé à une partie "self" et une partie "distinct"

$$G(\mathbf{r}, t) = G_s(\mathbf{r}, t) + G_d(\mathbf{r}, t) \quad (\text{II.39})$$

où

$$G_s(\mathbf{r}, t) = \frac{1}{N} \sum_j \int d^3 \mathbf{r}' \langle \delta(\mathbf{r} - \mathbf{r}' + \hat{\mathbf{R}}_j(0)) \delta(\mathbf{r}' - \hat{\mathbf{R}}_j(t)) \rangle, \quad (\text{II.40})$$

$$G_d(\mathbf{r}, t) = \frac{1}{N} \sum_{j \neq k} \int d^3 \mathbf{r}' \langle \delta(\mathbf{r} - \mathbf{r}' + \hat{\mathbf{R}}_j(0)) \delta(\mathbf{r}' - \hat{\mathbf{R}}_k(t)) \rangle. \quad (\text{II.41})$$

On note que  $G(\mathbf{r}, t)$  est une généralisation de la fonction de corrélation densité-densité statique,  $g(\mathbf{r}) = G(\mathbf{r}, 0)$ , et décrit la diffusion cohérente par un système d'atomes identiques, tandis que  $G_s(\mathbf{r}, t)$  décrit la diffusion incohérente dans ce cas.

Pour  $t \neq 0$  la fonction de van Hove est une fonction complexe qui n'a pas d'interprétation physique directe. Pour cette raison, l'interprétation des expériences de diffusion de neutrons est souvent effectuée dans l'approximation classique, où

$$F_s^{(cl)}(\mathbf{q}, t) = \frac{1}{N} \sum_j \int \langle e^{i\mathbf{q} \cdot [\mathbf{R}_j(t) - \mathbf{R}_j(0)]} \rangle_{cl}, \quad (\text{II.42})$$

$$F_d^{(cl)}(\mathbf{q}, t) = \frac{1}{N} \sum_{j \neq k} \langle e^{i\mathbf{q} \cdot [\mathbf{R}_k(t) - \mathbf{R}_j(0)]} \rangle_{cl}. \quad (\text{II.43})$$

Les fonctions de Van Hove correspondantes sont

$$G_s^{(cl)}(\mathbf{r}, t) = \frac{1}{N} \sum_j \langle \delta(\mathbf{r} - [\mathbf{R}_j(t) - \mathbf{R}_j(0)]) \rangle_{cl}, \quad (\text{II.44})$$

$$G_d^{(cl)}(\mathbf{r}, t) = \frac{1}{N} \sum_{j \neq k} \langle \delta(\mathbf{r} - [\mathbf{R}_k(t) - \mathbf{R}_j(0)]) \rangle_{cl}, \quad (\text{II.45})$$

et la fonction totale de Van Hove est

$$G^{(cl)}(\mathbf{r}, t) = G_s^{(cl)}(\mathbf{r}, t) + G_d^{(cl)}(\mathbf{r}, t). \quad (\text{II.46})$$

Contrairement au cas quantique, les versions classiques de  $G_s(\mathbf{r}, t)$  et  $G_d(\mathbf{r}, t)$  ont une interprétation simple pour  $t \neq 0$ , car

- $\delta(\mathbf{r} - [\mathbf{R}_j(t) - \mathbf{R}_j(0)])$  est la densité de probabilité pour un déplacement  $r$  de l'atome  $j$  au temps  $t$ .
- $\delta(\mathbf{r} - [\mathbf{R}_k(t) - \mathbf{R}_j(0)])$  est la densité de probabilité de trouver l'atome  $k$  au temps  $t$  déplacé de  $\mathbf{r}$  par rapport à la position de l'atome  $k$  au temps  $t = 0$ .

On note ici que Van Hove a montré dans un article relativement peu cité que la partie imaginaire de la fonction de van Hove est due à l'impact du neutron [25].

**2.6. La diffusion quasiélastique de neutrons (QENS).** La diffusion quasiélastique de neutrons (QENS) est une technique spectroscopique à résolution spatiale particulièrement adaptée à l'étude de la dynamique des protéines, car elle sonde non seulement les échelles de temps de la dynamique interne et globale de ces molécules complexes, mais aussi les amplitudes des mouvements. On note ici que la diffusion quasiélastique n'est pas forcément incohérente; dans des échantillons complètement deutérés elle a aussi une contribution cohérente. En raison de la section efficace dominante pour la diffusion incohérente de l'hydrogène, les expériences QENS sur des systèmes riches en hydrogène sondent essentiellement les mouvements diffusifs "self" de ces atomes [14, 15, 26] et dans ce cas, on peut écrire la fonction de corrélation de Van Hove classique sous la forme

$$G_H^{(cl)}(\mathbf{r}, t) = \frac{1}{N_H} \sum_{j \in H} \langle \delta(\mathbf{r} - [\mathbf{R}_j(t) - \mathbf{R}_j(0)]) \rangle_{cl}. \quad (\text{II.47})$$

Cette fonction devient donc la densité de probabilité pour un déplacement  $r$  dans un temps  $t$ , moyennée sur tous les atomes d'hydrogène.

**2.7. Facteur de structure élastique incohérent (EISF).** La fonction intermédiaire de diffusion peut être décomposée en une composante statique et une composante dynamique,

$$F_H(\mathbf{q}, t) = F_H(\mathbf{q}, \infty) + \delta F_H(\mathbf{q}, t), \quad (\text{II.48})$$

définissant

$$\delta F_H(\mathbf{q}, t) = F_H(\mathbf{q}, t) - F_H(\mathbf{q}, \infty). \quad (\text{II.49})$$

La valeur plateau est appelée le facteur de structure élastique incohérent (EISF) et elle est donné par

$$F_H(\mathbf{q}, \infty) \equiv EISF(\mathbf{q}) = \frac{1}{N_H} \sum_{j \in H} \overline{|\langle e^{i\mathbf{q} \cdot \mathbf{R}_j} \rangle|^2}. \quad (\text{II.50})$$

Il est non-nul si les mouvements des atomes sont limités dans l'espace, comme pour la dynamique interne des protéines qui sont étudiés sous forme de poudres hydratées. Avec ceci le facteur de structure dynamique peut s'écrire donc sous la forme

$$S_H(\mathbf{q}, \omega) = EISF(\mathbf{q}) \delta(\omega) + \delta S_H(\mathbf{q}, \omega) \quad (\text{II.51})$$

La composante  $\delta S_H(\mathbf{q}, \omega)$  contient le spectre quasi-élastique, qui est centré sur  $\omega = 0$  et décrit les mouvements stochastiques, et le spectre inélastique, qui est dû aux mouvements vibratoires. Le symbole  $\delta(\omega)$  désigne une distribution de Dirac et représente la ligne élastique idéale de spectre de diffusion de largeur nulle et d'intégrale finie.

L'EISF est une quantité importante car elle donne une information sur les amplitudes des mouvements atomiques et une première idée des caractéristiques des processus dynamiques dans le système de diffusion. La définition II.35 permet d'écrire que  $\int_{-\infty}^{+\infty} d\omega S_H(\mathbf{q}, \omega) = F_H(\mathbf{q}, 0) = 1$ . En utilisant II.51, nous obtenons donc

$$EISF(\mathbf{q}) + \int_{-\infty}^{+\infty} d\omega \delta S_H(\mathbf{q}, \omega) = 1, \quad (\text{II.52})$$

ce qui montre que tout processus dynamique donnant une contribution à  $\delta S_H(\mathbf{q}, \omega)$  conduit à une chute de l'EISF.

## Systèmes étudiés et motivation du travail

### 1. Relation structure-fonction et protéines intrinsèquement désordonnées

**1.1. Relation structure-fonction.** Les protéines sont des bio(macro) molécules constituées de chaînes d'acides aminés de longueurs très variables. Elles font partie des molécules les plus importantes des organismes vivants car elles accomplissent une grande variété de tâches. Chaque protéine possède une séquence plus au moins unique qui conduit le plus souvent à une structure tridimensionnelle bien définie permettant d'effectuer une tâche spécifique. La fonction biologique d'une protéine dépend en général de son arrangement spatial et structurel qui peut être décrit selon différents niveaux :

- **La structure primaire** d'une protéine qui est la séquence linéaire d'acides aminés et constitue le niveau de description le plus bas.
- **La structure secondaire** décrit l'arrangement spatial relatif de la chaîne d'acides aminés et est stabilisée par des liaisons d'hydrogène. Les hélices  $\alpha$ , les feuillets  $\beta$  sont les éléments de structure secondaire les plus courants qui décrivent la structure tridimensionnelle des segments locaux.
- **La structure tertiaire** définit l'arrangement des structures secondaires et constitue le plus haut niveau d'organisation structurale. On remarque que des résidus distants en séquence peuvent être proches dans la structure tertiaire.

- **La structure quaternaire** des protéines multimères sont constituées de plusieurs sous-unités tertiaires liées stoechiométriquement par des liaisons non covalentes.

Ces éléments structurels des protéines déterminent leur fonctionnalité, de sorte qu'une mutation de la séquence d'acides aminés peut entraîner une perte de fonction. Lorsqu'une protéine perd sa structure tridimensionnelle, elle se dénature et perd son activité. Dans certains cas, comme la protéine kinase et la ribonucléase, il est possible de revenir à l'état natif [27,28]. Grâce à des travaux de Anfinsen en 1961 sur le repliement de la ribonucléase [28], en montrant que lorsque cette protéine est exposée à des conditions dénaturantes, elle peut retrouver sa structure native lorsque la protéine est dialysée contre un tampon physiologique. La structure tridimensionnelle est essentielle pour la fonction [29], ce qui conduit à un paradigme structure-fonction. Ce dernier postule que la fonction d'une protéine dépend de sa structure et que sa connaissance détaillée renseigne sur sa fonction.

**1.2. Protéines intrinsèquement désordonnées.** Au cours des dernières années, un nombre croissant d'études indiquent que l'absence d'une structure 3D précise et stable pour une protéine donnée ne signifie pas son absence d'activité dans des conditions physiologiques. Dans des travaux récents, on a pu observer un changement de paradigme dans l'étude de la fonction des protéines, dans la mesure où ce n'est pas leur structure mais leur dynamique qui est considérée comme fondamentale. Les protéines dites partiellement repliées ou intrinsèquement désordonnées (IDP = intrinsically disordered proteins) dont la structure n'est pas ou que partiellement connue [30] peuvent bien avoir une fonction biologique. La découverte des protéines intrinsèquement désordonnées (IDP), qui ont une fonction bien définie sans avoir une

structure bien définie, a marqué un changement de paradigme en biologie structurale [31–34]. Depuis les années 1990, la littérature a connu une augmentation considérable des études ciblant les IDPs. Le terme "désordonné" signifie l'absence totale ou partielle d'éléments de structure secondaires, tels que les hélices  $\alpha$  ou des feuillets  $\beta$ , qui sont caractéristiques pour les protéines qui adoptent une structure bien précise. Cela signifie qu'elles sont très flexibles car les IDPs manquent de structures tertiaires et secondaires. Une IDP est une protéine avec peu ou pas de structure secondaire dans des conditions physiologiques, qui existe comme un ensemble de conformères dynamiques. Elle contient typiquement une ou plusieurs régions d'au moins 40 acides aminés consécutifs qui ne présentent aucune ressemblance avec les éléments de structure secondaire bien connus, tels que les hélices  $\alpha$  ou les feuillets  $\beta$ . Environ 30% des protéines dans les cellules eucaryotes ont des régions désordonnées [35] et elles interviennent dans les fonctions aussi importantes que la signalisation ou la régulation intracellulaire et elles peuvent aussi adopter une structure bien précise quand elles rencontrent une molécule cible. Ces protéines ont une charge plus élevée que les protéines globulaires et présentent des interactions électrostatiques avec les ligands [36]. Ceci est dû à la composition caractéristique en acides aminés des IDPs.

## 2. Protéines étudiés dans la thèse

La majorité des études par diffusion de neutrons réalisées sur des molécules sont réalisées sur des poudres hydratées. Dans notre travail, on voulait rester plus proche des conditions physiologiques en étudiant des protéines en solution. Notre étude concerne essentiellement trois types de protéines qui sont décrites par la suite :

- une protéine intrinsèquement désordonnée (Protéine Basique de la Myeline),
- une protéine globulaire (Myoglobine) qui a environ la même taille que la Protéine Basique de la Myeline
- et une enzyme contenant deux domaines globulaires (Phosphoglycérate kinase).

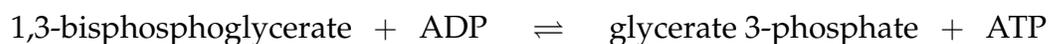
Le but est d'utiliser les modèles "minimalistes" mentionnés auparavant qui prennent en compte le caractère multi-échelle de la dynamique interne des protéines avec un minimum de paramètres, afin de dévoiler les changements dynamiques d'une protéine dus à une modification des conditions externes. Le défi est que la dynamique des modes de relaxation très lents se chevauche avec la diffusion élastique [37] et également avec la diffusion de molécules entières, ce qui doit être pris en compte puisque les protéines sont étudiées en solution.

**2.1. Protéine Basique de la Myeline (MBP).** L'un des membres important de la famille des IDPs est la protéine basique de la myéline (MBP), qui est le constituant de base de la gaine de myéline des nerfs et la deuxième protéine la plus abondante dans le système nerveux central. Du point de vue médical, la MBP est une molécule qui joue un rôle important dans la compréhension des maladies du système nerveux humain, car diverses pathologies neurologiques, comme la sclérose en plaques, sont dues à un dysfonctionnement ou à un mauvais repliement de cette protéine. En biophysique, la MBP est une protéine intéressante pour l'étude des propriétés physiques de protéines non- ou partiellement repliées en solution. Dans sa forme prédominante, la MBP contient 170 résidus et a un poids moléculaire de 18.5 kDa. En solution aqueuse, la protéine est intrinsèquement désordonnée, et la présente étude se concentre sur l'effet des mélanges D<sub>2</sub>O avec 30 % de trifluoroéthanol (TFE)

deutééré sur sa dynamique globale et interne. Le TFE est utilisé depuis de nombreuses années comme un agent connu pour induire la formation des hélices  $\alpha$  dans les IDPs, mais son action n'est pas encore complètement comprise.

**2.2. Myoglobine (Mb).** La myoglobine (Mb) est une protéine qui contient du fer (il s'agit d'une métalloprotéine) et qui est présente dans les muscles. Elle peut lier de manière réversible à l'oxygène et l'organisme l'utilise comme protéine de stockage de l'oxygène plutôt que de transporteur, comme l'hémoglobine. Elle est capable de lier et de libérer l'oxygène en fonction de la concentration d'oxygène dans la cellule. Sa fonction première est donc de fournir de l'oxygène aux cellules musculaires (myocyte). Notre choix était basé sur cette protéine car cette dernière est une protéine globulaire pliée de manière compacte, contrairement à la MBP, mais elle a environ le même poids moléculaire que la MBP qui est de 17.2 kDa et elle contient 154 acides aminés, mais avec une structure tridimensionnelle bien définie contenant huit  $\alpha$ -hélices comme éléments de structure secondaires.

**2.3. Phosphoglycérate kinase (PGK).** La phosphoglycérate kinase (PGK) est une enzyme monomérique qui est fondamentale pour le métabolisme de tous les organismes vivants. Elle catalyse d'une manière réversible l'une des deux réactions productrices d'ATP de la voie glycolytique, en convertissant le 1,3-bisphosphoglycérate en 3-phosphoglycérate, et participe également à la néoglucogénèse en catalysant la réaction inverse pour produire du 1,3- bisphosphoglycérate et de l'ADP [38].



La PGK est composée de deux domaines avec un site de liaison du 1,3-bisphosphoglycérate sur le domaine N et de l'ADP sur le domaine C. Ces deux domaines sont reliés par une région charnière bien conservée où les réactions catalytiques ont lieu. Cette enzyme a un poids d'environ 45 kDa et contient 417 acides aminés. L'idée était de voir l'influence de la liaison ATP sur la dynamique interne de la protéine en comparant l'enzyme phosphoglycérate kinase en présence et en l'absence de substrats.

### 3. Données QENS et spectromètres

Le travail présent concerne l'analyse de données de diffusion quasi-élastique de neutrons (QENS) qui ont été mesurées sur deux spectromètres différents à l'institut Laue- Langevin, Grenoble, France.

**3.1. Spectromètre IN16B.** est un spectromètre à rétrodiffusion "backscattering (BS)" de neutrons doté d'un nouveau "mode BATS" (Backscattering And Time of flight Spectrometer) qui étend considérablement la gamme de transfert d'énergie accessible [39,40]. Il s'agit d'un spectromètre à énergie incidente variable et à énergie finale fixe. Il est considéré comme le principal spectromètre de rétrodiffusion à haute résolution pour la diffusion incohérente quasi-élastique et inélastique des neutrons.

Instruments	Résolution énergétique	Gamme de Q	Diamètre du disque d	Fréquence de rotation f
IN16B	1.5-8 $\mu eV$	0.2-1.8 $\text{\AA}^{-1}$	750 [mm]	19000 [rpm]
IN5	100 $\mu eV$	0.2-11.8 $\text{\AA}^{-1}$	750 [mm]	17000 [rpm]

TABLE 1. Comparaison entre les deux instruments IN16B (Backscattering et TOF) et IN5 (TOF).

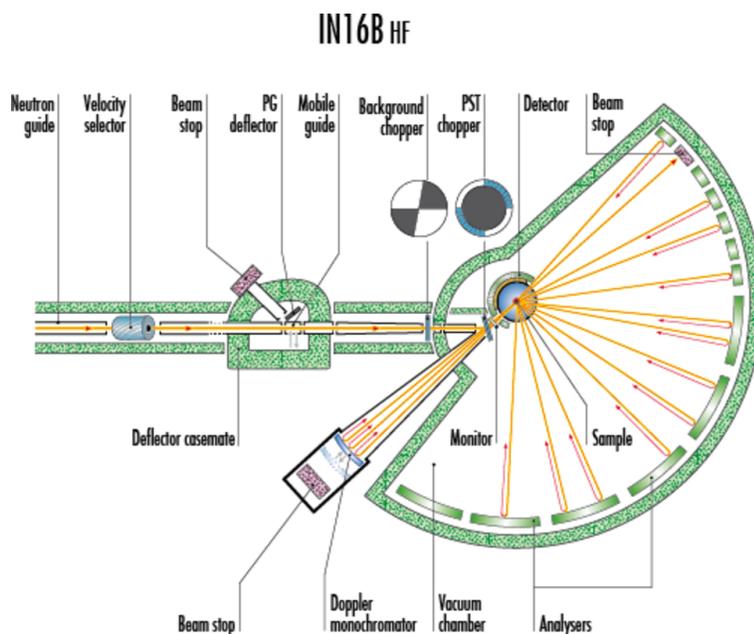


FIGURE III.1. Représentation schématique du spectromètre IN16B à l'ILL, Grenoble.

**3.2. Spectromètre IN5.** Il s'agit d'un spectromètre à temps de vol (Time-of-flight (ToF)) à géométrie directe de haute précision qui est utilisé pour étudier les processus de transfert à basse énergie (pourtant plus élevée que pour IN16B) en fonction du transfert de quantité de mouvement. Cet instrument est généralement utilisé pour des mesures dans la région des petits transferts d'énergie et de quantité de mouvement [41]. Il offre une flexibilité en matière de sélection de la longueur d'onde et de la vitesse du chopper, ce qui permet une optimisation considérable de la gamme d'énergie, de la résolution énergétique, de la gamme de transfert de quantité de mouvement.

Pour la protéine basique de la myéline et l'enzyme phosphoglycérate kinase, les données ont été collectées sur le spectromètre IN16B. Les données de la MBP ont été recueillies pour trois températures différentes, 283 K, 303 K, and 323 K, avec une résolution énergétique de  $3.5 \mu\text{eV}$  (FWHM), une gamme

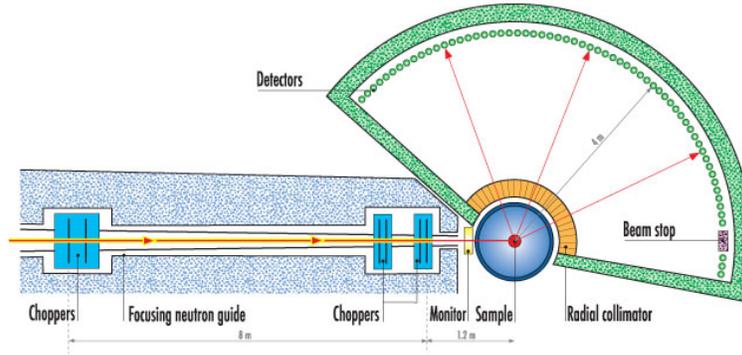


FIGURE III.2. Représentation schématique du spectromètre IN5 à l'ILL, Grenoble.

de transfert d'énergie de  $\pm 150 \mu\text{eV}$  et une gamme de transfert de quantité de mouvement de  $0.8 \text{ \AA}^{-1} < q < 1.8 \text{ \AA}^{-1}$  (en unités de  $\hbar$ ). Tous les échantillons et solvants ont été mesurés dans des porte-échantillons annulaires en aluminium avec une épaisseur d'interstice de 0.3 mm. Les données QENS des échantillons ont été normalisées par la référence du vanadium, et la réduction des données a été effectuée avec le logiciel MANTID [42].

Les données de la Myoglobine ont été collectées sur le spectromètre IN5, il s'agit des données qui ont été déjà analysées et publiées précédemment [43], dont l'idée était de faire une étude comparative de la dynamique de la myoglobine dans différents états structuraux en fonction de la température en utilisant la spectroscopie neutronique à temps de vol et à rétrodiffusion à haute résolution pour voir la dynamique la plus lente de la picoseconde à la nanoseconde.

## Chapitre 4

### Analyse de données

#### 1. Concept général

Les protéines sont des systèmes moléculaires complexes, avec un grand nombre de degrés de liberté dont la dynamique de relaxation est multi-exponentielle, ce qui est une caractéristique des systèmes complexes en général. Donc au lieu d'utiliser des modèles classiques pour différents types de mouvements atomiques, qui affichent une relaxation exponentielle pour quelques degrés de liberté particuliers, l'approche utilisée ici est "model-free", tenant compte de la dynamique de relaxation multi-échelle des protéines avec un minimum de paramètres et une distribution continue d'échelles de relaxation. Cette approche est liée au concept des "paysages d'énergies" (*energy landscapes*) proposé par Frauenfelder et ses collaborateurs dans les années 1970, qui ont utilisé la photolyse flash au laser pour étudier la cinétique fortement non-exponentielle de la re-fixation du monoxyde de carbone à la myoglobine [44, 45]. La dynamique de relaxation multi-échelle des protéines est ici considérée comme un processus de sauts entre les minima d'un paysage d'énergie libre, où chaque minimum correspond à un "sous-état conformationnel" de la protéine [46]. Le chapitre présent aborde l'analyse de données QENS provenant de protéines en solution (la Protéine Basique de la Myéline, la Myoglobine et l'enzyme Phosphoglycérate Kinase). L'analyse de données utilise les

concepts de l'analyse asymptotique, qui permet de lier la dynamique de relaxation aux temps longs à la forme du spectre QENS aux petites fréquences, et le concept des *energy landscapes* dans le contexte de la diffusion de neutrons [37]. Elle prend spécifiquement en considération l'aspect d'autosimilarité des fonctions de relaxation qui se manifeste par une décroissance asymptotique lente avec une loi de puissance,  $\phi(t) \sim t^{-\alpha}$ , avec  $0 < \alpha \leq 1$  [5, 46–48], qui a beaucoup en commun avec la dynamique des verres [49].

**1.1. Analyse en temps par transformée de Fourier.** L'analyse de données des spectres QENS a été menée dans le domaine du temps, c'est-à-dire sur la fonction intermédiaire de diffusion, en éliminant l'effet de la résolution instrumentale représentée par un spectre Vanadium. Donc l'analyse de données est effectuée sur la fonction intermédiaire de diffusion  $F(\mathbf{q}, t)$ , qui est obtenue par transformée de Fourier discrète à partir du facteur de structure dynamique mesuré,  $S(\mathbf{q}, \omega)$ ,

$$F(\mathbf{q}, t) = \int_{-\infty}^{+\infty} d\omega e^{i\omega t} S(\mathbf{q}, \omega), \quad (\text{IV.53})$$

où l'indice "H" est omis. En introduisant de la densité transformée de Fourier de l'atome j,

$$\hat{\rho}_j(\mathbf{q}, t) = e^{i\mathbf{q} \cdot \hat{\mathbf{R}}_j(t)}, \quad (\text{IV.54})$$

$$\delta \hat{\rho}_j(\mathbf{q}, t) = e^{i\mathbf{q} \cdot \hat{\mathbf{R}}_j(t)} - \langle e^{i\mathbf{q} \cdot \hat{\mathbf{R}}_j(t)} \rangle, \quad (\text{IV.55})$$

la fonction intermédiaire de diffusion pour la dynamique interne peut être exprimée sous la forme générique

$$F(\mathbf{q}, t) = \frac{1}{N} \sum_{j \in H} \langle \hat{\rho}_j^\dagger(\mathbf{q}, 0) \hat{\rho}_j(\mathbf{q}, t) \rangle, \quad (\text{IV.56})$$

$$= EISF(\mathbf{q}) + (1 - EISF(\mathbf{q}))\phi(\mathbf{q}, t), \quad (\text{IV.57})$$

où  $EISF(\mathbf{q})$  est le facteur de structure incohérent élastique défini dans le premier chapitre,

$$EISF(q) = \frac{1}{N} \sum_{j \in H} |\langle \hat{\rho}_j(\mathbf{q}) \rangle|^2, \quad (\text{IV.58})$$

et  $\phi(\mathbf{q}, t)$  est une fonction de relaxation,

$$\phi(\mathbf{q}, t) = \left( \frac{\sum_{j \in H} \langle \delta \hat{\rho}_j^\dagger(\mathbf{q}, 0) \delta \hat{\rho}_j(\mathbf{q}, t) \rangle}{\sum_{j \in H} \langle \delta \hat{\rho}_j^\dagger(\mathbf{q}, 0) \delta \hat{\rho}_j(\mathbf{q}, 0) \rangle} \right), \quad (\text{IV.59})$$

qui vérifie  $\phi(\mathbf{q}, t) = 1$ .

Les protéines en solution sont des systèmes isotropes et le facteur de structure dynamique total observé peut être exprimé comme une moyenne angulaire isotrope sur toutes les directions de  $\mathbf{q}$ ,

$$\mathcal{S}_{\text{tot}}(q, t) = \overline{\mathcal{S}_{\text{tot}}(\mathbf{q}, t)} = \frac{1}{4\pi} \int \int \sin^2(\theta) d\theta d\phi \mathcal{S}_{\text{tot}}(q\mathbf{n}(\theta, \phi), t). \quad (\text{IV.60})$$

Ici,  $\mathbf{n}_q = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)$  est le vecteur unitaire pointant dans la direction de  $\mathbf{q}$  et  $q \equiv |\mathbf{q}|$ .

**1.2. La soustraction des solvants.** Les données expérimentales des solvants  $D_2O$  ont été soustraites des données des solutions de protéines selon la méthode suivante

$$\mathcal{S}_{\text{protein}}(q, \omega) \approx \mathcal{S}_{\text{tot}}(q, \omega) - (1 - \phi) \mathcal{S}_{\text{buffer}}(q, \omega), \quad (\text{IV.61})$$

en considérant que le volume spécifique partiel de la protéine est  $\bar{v} = 0.73$  ml/g. Pour une concentration de  $c = 55$  mg/ml, cela conduit à une fraction de volume pour la protéine de  $\phi = c\bar{v} = 0.04015$  [43].

**1.3. Symétrisation et normalisation des spectres QENS.** Les relations de symétrie II.28 et II.29 montrent que la fonction intermédiaire de diffusion devient une fonction réelle symétrique dans le temps si l'on considère la limite classique  $\hbar \rightarrow 0$  et si l'on peut supposer que les fonctions de diffusion sont invariantes par rapport au changement  $\mathbf{q} \rightarrow -\mathbf{q}$ . Pour l'analyse de données, nous utilisons le facteur de structure dynamique symétrisé

$$S^{(+)}(q, \omega) = \frac{e^{-\beta\hbar\omega/2} S(q, \omega)}{\int_{-\infty}^{+\infty} d\omega e^{-\beta\hbar\omega/2} S(q, \omega)}, \quad (\text{IV.62})$$

qui est une fonction paire en  $\omega$  et qui est normalisée de telle sorte que

$$\int_{-\infty}^{+\infty} d\omega S^{(+)}(q, \omega) = F^{(+)}(q, 0) = 1. \quad (\text{IV.63})$$

La fonction intermédiaire de diffusion correspondante est une fonction symétrique dans le temps,

$$F^{(+)}(q, t) = \frac{F(q, t + i\beta\hbar/2)}{F(q, i\beta\hbar/2)}, \quad (\text{IV.64})$$

Nous notons que

$$F^{(+)}(q, t) \approx F^{(cl)}(q, t) \quad (\text{IV.65})$$

correspondent à l'approximation semi-classique de Schofield, [50] où "cl" indique la fonction de diffusion classique correspondante. Avec ce raisonnement, les modèles classiques peuvent être utilisés pour l'interprétation des spectres QENS. Pour pouvoir travailler avec des modèles de relaxation classiques, nous considérons maintenant pour la dynamique interne la forme générique symétrisée et réelle en temps

$$F_{int}^{(+)}(q, t) = EISF(q) + (1 - EISF(q))\phi^{(+)}(q, t) \quad (\text{IV.66})$$

de la fonction intermédiaire de diffusion, où la fonction de relaxation est définie comme suit

$$\phi^{(+)}(t) \equiv \frac{\phi(t + i\beta\hbar/2)}{\phi(i\beta\hbar/2)}, \quad (\text{IV.67})$$

afin d'assurer sa normalisation correcte.

**1.4. Extraction et déconvolution de  $F(q,t)$ .** Comme il est indiqué précédemment, nos analyses portent sur des mesures QENS collectées sur deux spectromètres différents IN16B et IN5 à l'ILL. La résolution instrumentale est obtenue à partir des spectres QENS du vanadium. Le facteur de structure dynamique expérimental, désigné par un indice "m" ("mesuré"), peut être écrit comme la convolution du facteur de structure dynamique "idéal" et de la résolution instrumentale plus un bruit,

$$S_m^{(+)}(q, \omega) = S^{(+)}(q, \omega) * \tilde{R}(q, \omega) + \tilde{N}(q, \omega), \quad (\text{IV.68})$$

où \* représente un produit de convolution

$$S^{(+)}(q, \omega) * \tilde{R}(q, \omega) = \int_{-\infty}^{+\infty} d\omega' R(q, \omega' - \omega) S^{(+)}(q, \omega'), \quad (\text{IV.69})$$

et  $\tilde{N}(q, \omega)$  et  $\tilde{R}(q, \omega)$  représentent respectivement le bruit et la résolution instrumentale qui est considérée comme symétrique en  $\omega$ . Dans notre étude,  $\tilde{R}(q, \omega) = S_{\text{vanadium}}^{(+)}(q, \omega)$  est le facteur de structure dynamique symétrisé d'un échantillon de vanadium aux dimensions identiques à celles de l'échantillon d'intérêt, et le bruit est essentiellement inconnu. Ici, on utilise le fait que le vanadium peut être considéré comme un diffuseur isotrope et élastique idéal. La transformée de Fourier inverse de l'équation (IV.68) est la suivante

$$F_m^{(+)}(q, t) = F^{(+)}(q, t)R(q, t) + N(q, t), \quad (\text{IV.70})$$

en notant que le produit de convolution de  $S^{(+)}(q, \omega)$  et  $\tilde{R}(q, \omega)$  en  $\omega$  devient un produit normal de  $F^{(+)}(q, t)$  et  $R(q, t)$  dans le domaine du temps. La résolution de l'équation (IV.70) par rapport à la fonction intermédiaire de diffusion symétrique déconvoluée,  $F^{(+)}(q, t)$ , est donc la suivante

$$F^{(+)}(q, t) = \frac{F_m^{(+)}(q, t)}{R(q, t)} - \frac{N(q, t)}{R(q, t)}. \quad (\text{IV.71})$$

Le point important est ici que le terme  $N(q, t)/R(q, t)$  devient important si  $R(q, t)$  devient petit, ce qui limite l'échelle de temps accessible expérimentalement. Notons que, le seuil de  $R(q, t)$  est atteint pour

$$t_{max} \approx \frac{\hbar}{2\Delta\omega}, \quad (\text{IV.72})$$

où  $\hbar$  est la constante réduite de Planck et  $\Delta\omega$  est la résolution (FWHM) en énergie de l'instrument. La forme du bruit est inconnue ici, mais sa contribution devient clairement visible lorsque  $t$  augmente, c'est pour ça qu'on divise le bruit, pour  $t$  croissant, par des petites intensités de la résolution instrumentale, on voit que l'effet du bruit peut être négligé sur un domaine de temps d'environ 0 – 600 ps. Pour le spectromètre IN16B, sur lequel les données QENS analysées dans notre première étude ont été enregistrées, nous avons  $\Delta\omega = 3.5 \mu\text{eV}$ , ce qui correspond à  $t_{max} \approx 600$  ps. La figure IV.1 illustre la déconvolution pour le spectre QENS de la MBP dans le solvant D<sub>2</sub>O pour T=283 K à  $q = 1.057/\text{Å}$ . Dans la partie gauche, nous montrons le facteur de structure dynamique symétrisé et normalisé  $S_m^{(+)}(\mathbf{q}, \omega)$  (ligne bleue) ainsi que le spectre correspondant du vanadium  $\tilde{R}(\mathbf{q}, \omega)$ , qui définit la résolution instrumentale (ligne jaune). La partie droite montre les fonctions intermédiaires de diffusion correspondantes  $F_m^{(+)}(q, t)$  (points bleus) et  $R(q, t)$  (points jaunes), qui ont été obtenues par la transformée de Fourier discrète à partir de tableaux échantillonnés de manière équidistante représentant  $S_m^{(+)}(q, \omega)$  et  $\tilde{R}(q, \omega)$ . La fonction intermédiaire de

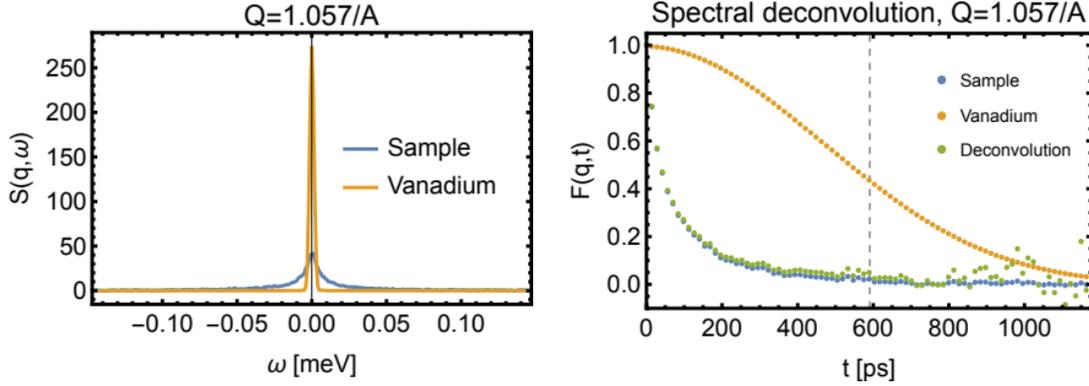


FIGURE IV.1. **Partie gauche** : Facteur de structure dynamique symétrisé et normalisé,  $S_m^{(+)}(q, \omega)$ , de la MBP dans le solvant  $D_2O$  à  $q = 1.057 \text{ \AA}^{-1}$  et  $T = 283 \text{ K}$  (ligne bleue) avec la fonction de résolution correspondante  $\tilde{R}(q, \omega)$  (ligne jaune) obtenue à partir d'une analyse au vanadium. **Partie droite** : Fonction intermédiaire de diffusion correspondante  $F_m^{(+)}(q, t)$ , dans le domaine temporel (points jaunes) et fonction intermédiaire de diffusion déconvoluée  $F^{(+)}(q, t)$  (points verts). La ligne verticale en pointillés indique le domaine du temps accessible correspondant à la résolution de l'instrument (Eq. (IV.72)).

diffusion déconvoluée par la résolution  $F^{(+)}(q, t)$ , est donnée par les points verts, et la ligne verticale en pointillés indique le seuil de  $t_{max} \approx 600 \text{ ps}$  correspondant à la résolution instrumentale. Tous les points des données remplissant  $t \leq t_{max}$  ont été utilisés pour l'ajustement du modèle à décrire dans la suite.

## 2. Modèles pour la dynamique des protéines

**2.1. Relaxation multi-échelle dans des systèmes complexes.** La dynamique interne des protéines et des systèmes moléculaires complexes est généralement caractérisée par une relaxation fortement non exponentielle des fonctions de corrélation temporelle pertinentes, qui présentent une décroissance

asymptotique en loi de puissance. La décroissance en loi de puissance reflète l'autosimilarité, c'est-à-dire l'invariance des formes des fonctions de corrélation en cas de changement de l'échelle de temps, ce qui est également typique pour les verres [9, 10, 46–49, 51–53]. Dans ce contexte, notre étude vise l'utilisation d'une approche qui décrit la dynamique multi-échelle de protéines avec des modèles "minimalistes" qui prend en compte le caractère multi-échelle de la dynamique interne des protéines. Ce modèle décrit un spectre QENS avec seulement trois paramètres : un paramètre de forme et un paramètre d'échelle pour décrire la dynamique de relaxation d'une protéine, ainsi que l'amplitude de diffusion élastique (elastic incoherent structure factor, EISF) qui donne une information sur l'amplitude des mouvements atomiques. L'idée est donc de décrire la dynamique d'un système complexe comme une protéine, telle qu'elle est observée par QENS, par des modèles qui décrivent leur dynamique essentielle sur la base de concepts physiques avec peu de paramètres.

**2.2. Fonction intermédiaire de diffusion pour la dynamique interne.** Le but de ce travail est d'ajuster les fonctions intermédiaires extraites déconvoluées de la résolution instrumentale. Pour pouvoir travailler avec des modèles de relaxation classiques, nous considérons la forme générique symétrisée dans le temps [54]

$$F^{(+)}(\mathbf{q}, t) = F(\mathbf{q}, \infty) + (F(\mathbf{q}, 0) - F(\mathbf{q}, \infty))\phi_{\text{ML}}^{(+)}(\mathbf{q}, t) \quad (\text{IV.73})$$

notant que  $F(\mathbf{q}, 0) = 1$ , la fonction intermédiaire de diffusion des neutrons est liée par la transformée de Fourier à la quantité mesurée dans la plupart des expériences de diffusion de neutrons, et qui se traduit par la forme générique correspondante

$$S^{(+)}(\mathbf{q}, \omega) = F(\mathbf{q}, \infty)\delta(\omega) + (F(\mathbf{q}, 0) - F(\mathbf{q}, \infty))\tilde{\phi}_{\text{ML}}(\mathbf{q}, \omega) \quad (\text{IV.74})$$

du facteur de structure dynamique.

Dans le contexte de nos études,  $F(\mathbf{q}, 0) = 1$ ,  $F(\mathbf{q}, \infty) \equiv \lim_{t \rightarrow \infty} F^{(+)}(t)$  est la valeur plateau qui correspond à l'EISF. La fonction de relaxation  $\phi_{\text{ML}}$  est décrite par une fonction Mittag-Leffler (ML) "étirée" [55], qui présente une décroissance asymptotique en loi de puissance et une auto-similarité,

$$\phi_{\text{ML}}^{(+)}(t) = E_{\alpha}(-(|t|/\tau_R)^{\alpha}), \quad 0 < \alpha < 1, \quad (\text{IV.75})$$

où la dépendance  $\mathbf{q}$  est omise pour une meilleure lisibilité. Ici  $E_{\alpha}(\cdot)$  est la fonction de Mittag-Leffler (ML), qui porte le nom du mathématicien suédois Gosta Mittag-Leffler (1846-1927) [55]. Cette fonction est définie par la série de Taylor,

$$E_{\alpha}(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(1 + \alpha n)} \quad (\alpha > 0), \quad (\text{IV.76})$$

pour tous les arguments dans le système complexe. Pour  $0 < \alpha < 1$ , la fonction de relaxation ML (IV.75) décroît de façon monotone, avec une forme de loi de puissance asymptotique,

$$\phi_{\text{ML}}(t) \stackrel{t \gg \tau_R}{\sim} \frac{(t/\tau_R)^{-\alpha}}{\Gamma(1 - \alpha)}, \quad (\text{IV.77})$$

et elle peut être considérée comme une généralisation de la fonction exponentielle, à laquelle elle se retrouve pour  $\alpha=1$ ,

$$\lim_{\alpha \rightarrow 1} \phi_{\text{ML}}(t) = \exp(-|t|/\tau). \quad (\text{IV.78})$$

Comme la transformée de Laplace de la fonction de relaxation ML a une forme analytique simple [55]

$$\hat{\phi}_{\text{ML}}(s) = \frac{1}{s(1 + (s\tau_R)^{-\alpha})}, \quad (\text{IV.79})$$

ses spectres de Fourier et de taux de relaxation peuvent être facilement obtenus par le biais de

$$\phi^{(+)}(\omega) = \lim_{\epsilon \rightarrow 0} \frac{1}{\pi} \Re \left\{ \hat{\phi}^{(+)}(i\omega + \epsilon) \right\}, \quad (\text{IV.80})$$

ce qui conduit à une “lorentzienne généralisée” [56]

$$\tilde{\phi}_{\text{ML}}(\omega) = \frac{1}{\pi} \frac{\sin\left(\frac{\pi\alpha}{2}\right)}{|\omega| \left( (|\omega|\tau_R)^{-\alpha} + (|\omega|\tau_R)^\alpha + 2 \cos\left(\frac{\pi\alpha}{2}\right) \right)}. \quad (\text{IV.81})$$

Ici  $\int_{-\infty}^{+\infty} d\omega \tilde{\phi}_{\text{ML}}(\omega) = 1$  et la lorentzienne standard est retrouvée pour  $\alpha \rightarrow 1$ . On note que  $\tilde{\phi}_{\text{ML}}(\omega) \sim |\omega|^{\alpha-1}$  pour  $\omega \rightarrow 0$ , et devient donc singulière pour  $0 < \alpha < 1$ .

La transformée de Laplace (IV.79) de la fonction de relaxation ML peut également être utilisée pour obtenir une forme analytique de son spectre de taux de relaxation. Pour une fonction de relaxation arbitraire à décroissance monotone,  $\phi(t)$ , la fonction est définie par la relation suivante

$$\phi(t) = \int_0^\infty d\lambda p(\lambda) \exp(-\lambda|t|), \quad (\text{IV.82})$$

où  $p(\lambda) \geq 0$  et  $\int_0^\infty d\lambda p(\lambda) = 1$ . En utilisant

$$p(\lambda) = \lim_{\epsilon \rightarrow 0} \frac{1}{\pi} \Im \left\{ \hat{\phi}^{(+)}(-\lambda - i\epsilon) \right\}, \quad (\text{IV.83})$$

on trouve pour la fonction de relaxation ML [56]

$$p_{\text{ML}}(\lambda) = \frac{\sin(\pi\alpha)}{\pi\lambda \left( (\lambda\tau_R)^{-\alpha} + (\lambda\tau_R)^\alpha + 2 \cos(\pi\alpha) \right)}. \quad (\text{IV.84})$$

La décroissance lente en loi de puissance (IV.75) de  $\phi_{\text{ML}}(t)$  conduit ici à une forme de loi de puissance du spectre de taux de relaxation pour des petits arguments,  $p_{\text{ML}}(\lambda) \stackrel{\lambda \rightarrow 0}{\sim} \lambda^{\alpha-1}$ .

Notant que la fonction génératrice de moments (IV.75) montre que les moments et les cumulants de  $p(\lambda)$  divergent, mais on peut définir la médiane,

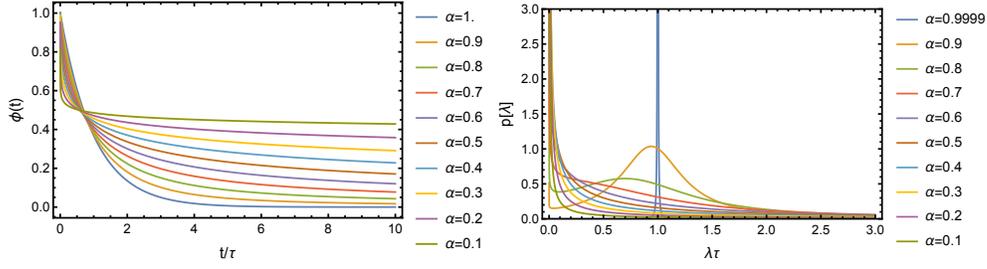


FIGURE IV.2. La fonction Mittag-Leffler “étirée” et le spectre de taux de relaxation.

$\lambda_{1/2}$ , de  $p(\lambda)$ , qui est définie par

$$\int_0^{\lambda_{1/2}} d\lambda p(\lambda) = \frac{1}{2}$$

et pour l’expression (IV.84), on obtient

$$\lambda_{1/2} = 1. \quad (\text{IV.85})$$

Il est aussi intéressant de calculer les spectres de temps de relaxation équivalents. Le temps de relaxation est défini comme l’inverse du taux de relaxation,

$$\tau = \frac{1}{\lambda}, \quad (\text{IV.86})$$

En utilisant que  $p(\lambda)d\lambda \stackrel{!}{=} P(\tau)d\tau$ , ou explicitement

$$p(\lambda)d\lambda = p(\lambda(\tau)) \left| \frac{d\lambda(\tau)}{d\tau} \right| d\tau.$$

On trouve

$$P(\tau) = p(\lambda(\tau)) \frac{1}{\tau^2}. \quad (\text{IV.87})$$

Dans cette partie, nous avons introduit des modèles pour les échelles de relaxation temporelle sans dimension, afin de conserver une notation compacte.

Pour une échelle de temps concrète  $\tau_s$ , nous avons donc

$$P(\tau) \rightarrow \frac{1}{\tau_s} P\left(\frac{\tau}{\tau_s}\right). \quad (\text{IV.88})$$

La fonction de relaxation ML a la propriété intéressante que son spectre de temps de relaxation a exactement la même forme que son spectre de taux de relaxation. En utilisant l'équation Eq. (IV.87), on trouve en effet,

$$P(\tau) = \frac{\sin(\pi\alpha)}{\pi\tau((\lambda\tau)^{-\alpha} + (\lambda\tau)^\alpha + 2\cos(\pi\alpha))}. \quad (\text{IV.89})$$

La dynamique des protéines peut être considérée comme un processus de diffusion dans un "potentiel rugueux", avec une large distribution de hauteurs de barrière. Pour une hauteur de barrière fixe, une étude a été réalisée par Zwanzig [57], et le résultat peut être utilisé pour convertir le spectre de taux de relaxation  $p(\lambda)$  en un spectre de barrière énergétique  $p(\epsilon)$ , en supposant que le potentiel est harmonique. Nous avons utilisé le modèle de Zwanzig [57] qui relie le coefficient de diffusion  $D_0$  d'une particule dans un potentiel lisse à la constante de diffusion  $D$  dans un potentiel rugueux avec une hauteur de barrières  $\Delta E$ .

$$D = D_0 \exp(-(\Delta E/k_B T)^2) \quad (\text{IV.90})$$

qui se traduit donc pour un potentiel harmonique par

$$\lambda = \lambda_0 \exp(-(\Delta E/k_B T)^2) \quad (\text{IV.91})$$

pour la constante de relaxation. Cette relation peut être utilisée pour convertir une distribution de taux de relaxation  $p(\lambda)$  en une distribution de hauteurs de barrière,

$$\epsilon = \frac{\Delta E}{k_B T} \quad (\text{IV.92})$$

la distribution des barrières d'énergie correspondant à la fonction de relaxation ML est la suivante [58]

$$P_{\text{ML}}(\epsilon) = \frac{2\epsilon \sin(\pi\alpha)}{\pi (e^{-\alpha\epsilon^2} + e^{\alpha\epsilon^2} + 2\cos(\pi\alpha))}. \quad (\text{IV.93})$$

L'interprétation physique du potentiel rugueux sera donnée d'une manière

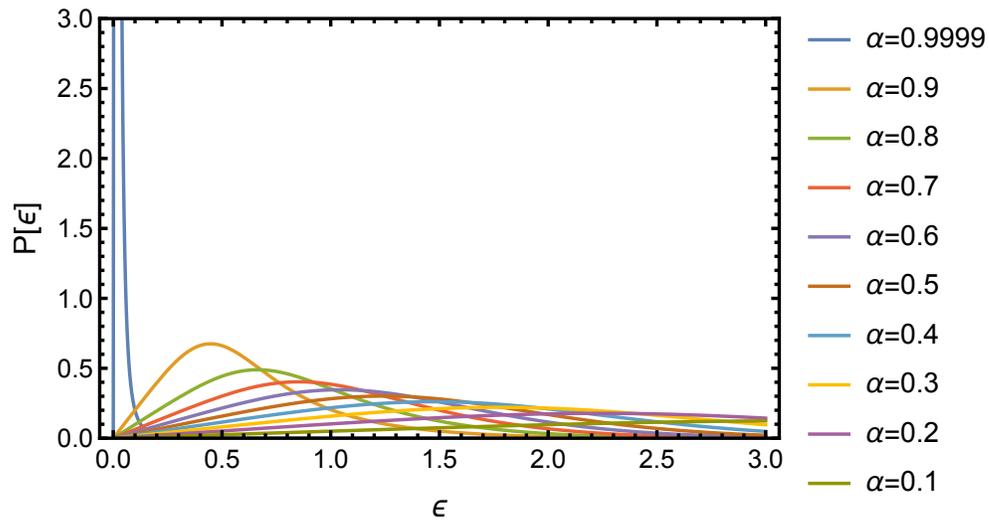


FIGURE IV.3. Spectre de barrière d'énergie pour différentes valeurs de  $\alpha$ .

plus détaillée dans le prochain chapitre 5.



### **Dynamique d'une protéine intrinsèquement désordonnée**

Cette partie présente une analyse des spectres quasiélastiques de neutrons (QENS) provenant de la protéine basique de la myéline (MBP) dans deux solvants,  $D_2O$  et en  $D_2O$  avec 30% de Trifluoroéthanol (TFE) deutéré, et à trois températures, 283 K, 303 K et 323 K. L'objectif de la présente étude est de voir l'effet du TFE sur le repliement global et la dynamique de cette protéine intrinsèquement désordonnée afin d'évaluer l'impact de la formation d'éléments de structure secondaire sur la dynamique interne [59]. Comme indiqué ci-dessus, la MBP est un composant principal de la gaine de myéline dans le système nerveux, en  $D_2O$  la molécule est intrinsèquement désordonnée et en  $D_2O$  + 30% TFE partiellement repliée. En plus aux expériences de diffusion de neutrons, des expériences complémentaires ont été réalisées comme la diffusion dynamique de la lumière (DLS) et la spectroscopie de dichroïsme circulaire utilisant le rayonnement synchrotron (SRCD) afin d'estimer, respectivement, le coefficient de diffusion global de la MBP et le contenu en structure secondaire de la molécule pour les deux solvants en fonction de la température.

#### **1. Diffusion dynamique de la lumière**

La diffusion dynamique de la lumière est également connue sous le nom de spectroscopie de corrélation de photons. Cette technique est l'une des méthodes les plus utilisées pour déterminer la taille des particules. Dans notre étude, des expériences de diffusion dynamique de la lumière (DLS) avec la

protéine basique de la myéline (MBP) ont été réalisées pour obtenir des estimations fiables du coefficient de diffusion translationnelle de la MBP. Les spectres DLS ont été enregistrés pour une gamme de température de 283 K à 323 K et une concentration de protéines de 4 mg/ml pour  $D_2O$  pur et de 7 mg/ml pour  $D_2O + 30\%$  TFE. Les spectres DLS ont été réalisés sur un instrument Zetasizer Nano ZS (Malvern Instruments, Malvern, Royaume-Uni) et les fonctions d'autocorrélation ont été analysées par l'algorithme CONTIN [60]. Les expériences DLS sont sensibles à la coagulation de la MBP et les concentrations choisies reflètent un compromis entre une concentration suffisante pour de bonnes statistiques de comptage et l'augmentation générale de la coagulation avec la concentration.

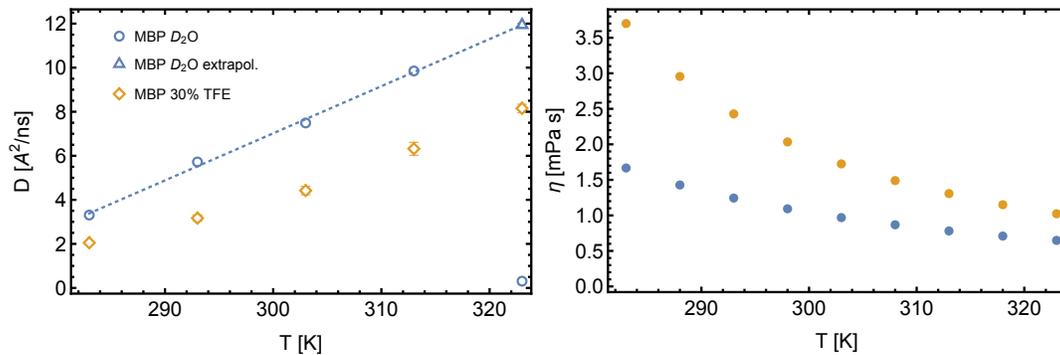


FIGURE V.1. Coefficients de diffusion (partie gauche) et viscosités (partie droite) obtenus par DLS de la MPB dans les deux solvants et à différentes températures.

## 2. Dichroïsme circulaire

Le dichroïsme circulaire est une technique analytique essentielle et rapide en biologie structurale pour déterminer le contenu de la structure secondaire des protéines, des polypeptides et des structures peptidiques. Les spectres de

dichroïsme circulaire par rayonnement synchrotron (SRCD) ont été enregistrés pour estimer la structure secondaire de la protéine basique de la myéline (MBP). Les mesures ont été effectuées sur la ligne de faisceau DISCO de la source nationale française de rayonnement synchrotron SOLEIL à Gif-sur-Yvette, en sélectionnant la région spectrale de l'UV lointain (150 – 261 nm). Pour la MBP dans le solvant D<sub>2</sub>O, nous avons utilisé une cuvette avec une longueur optique de 0.0012 cm et une concentration de 44 mg/mL et pour la MBP dans le solvant D<sub>2</sub>O contenant 30% de TFE deutéré, nous avons utilisé une cuvette avec une longueur optique de 0.002 cm et une concentration de 20 mg/mL. Pour l'analyse des spectres SRCD, nous avons utilisé BestSel [61] pour estimer le contenu de la structure secondaire de la MBP dans les deux solvants et à différentes températures.

### 3. Modèle utilisé pour l'analyse de données QENS de la MBP

L'analyse des spectres QENS des protéines intrinsèquement désordonnées présente la difficulté de déterminer la dynamique globale et la dynamique interne. Formellement, la trajectoire pour la diffusion d'une macromolécule flexible est définie comme une accumulation de déplacements infinitésimaux de corps rigides reliant par superposition optimale des images successives dans une trajectoire de la dynamique moléculaire. Les calculs et les applications dans la référence [62] suggèrent que les mouvements segmentaires internes contribuent essentiellement à une diffusion rotationnelle effective et une observation similaire a été faite récemment [63]. Pour les considérations suivantes, nous faisons maintenant l'hypothèse que la fonction intermédiaire de diffusion peut être factorisée sous la forme

$$F^{(+)}(q, t) \approx e^{-Dq^2|t|} F_{\text{int}}^{(+)}(q, t), \quad (\text{V.94})$$

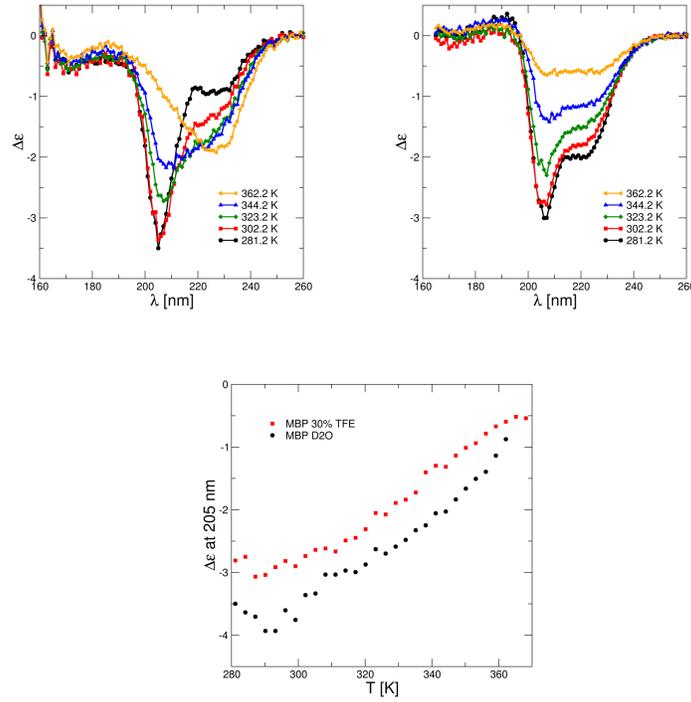


FIGURE V.2. Spectres SRCD pour la MBP dans le solvant D<sub>2</sub>O (partie gauche) et le solvant D<sub>2</sub>O+ 30% TFE (partie droite), et la valeur minimale pour les deux solvants à λ=205 nm en fonction de la température (panneau en bas).

où  $D$  est considéré comme un coefficient de diffusion translationnel effectif. On suppose alors que la diffusion de la molécule entière n'est pas corrélée avec sa dynamique interne. La fonction intermédiaire de diffusion pour la dynamique interne de la MBP a la forme générique suivante

$$F_{\text{int}}^{(+)}(q, t) = EISF(q) + (1 - EISF(q))\phi^{(+)}(q, t). \quad (\text{V.95})$$

#### 4. Facteur de structure dynamique total

En utilisant la forme générale de la fonction intermédiaire de diffusion totale, nous écrivons la version symétrique sous la forme suivante

$$F^{(+)}(t) = e^{-\epsilon|t|} (EISF + (1 - EISF)\phi_{ML}(|t|)), \quad (\text{V.96})$$

en omettant la dépendance  $q$  dans les formules, et en définissant

$$\epsilon := Dq^2. \quad (\text{V.97})$$

Le facteur de structure dynamique peut alors être exprimé en terme de transformée de Laplace de la fonction intermédiaire de diffusion,

$$S^{(+)}(\omega) = \frac{1}{\pi} \Re \left\{ \hat{F}^{(+)}(i\omega + \epsilon) \right\}, \quad (\text{V.98})$$

ce qui mène à

$$S^{(+)}(\omega) = EISF \frac{1}{\pi} \frac{\epsilon}{\omega^2 + \epsilon^2} + (1 - EISF) \tilde{\phi}_{ML}^{(\epsilon)}(|\omega|), \quad (\text{V.99})$$

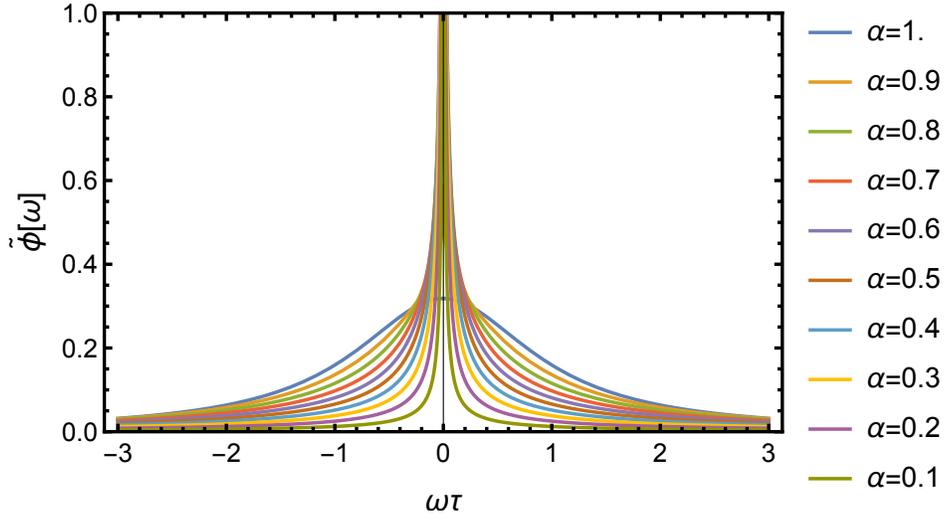
où  $\tilde{\phi}_{ML}^{(\epsilon)}(\omega)$  est la transformée de Fourier de la fonction de relaxation [12],

$$\tilde{\phi}_{ML}^{(\epsilon)}(\omega) = \frac{\epsilon (\omega^2 + \epsilon^2)^{\alpha/2} + \omega \sin(\alpha \arg(\epsilon + i|\omega|)) + \epsilon \cos(\alpha \arg(\epsilon + i|\omega|))}{(\omega^2 + \epsilon^2) \left( (\omega^2 + \epsilon^2)^\alpha + 1 \right) (\omega^2 + \epsilon^2)^{-\alpha/2} + 2 \cos(\alpha \arg(\epsilon + i|\omega|))}. \quad (\text{V.100})$$

Il convient de noter que la fonction lorentzienne pondérée par l'EISF dans l'équation Eq. (V.99) tend vers une distribution de Dirac dans la limite  $\epsilon \rightarrow 0$ .

#### 5. Résultats de l'analyse de données QENS de la MBP

Les résultats ont été obtenus en ajustant le modèle (V.94) aux fonctions intermédiaires de diffusion expérimentales déconvoluées en fonction de la résolution. Pour faire apparaître explicitement les paramètres et les fonctions du

FIGURE V.3. Spectre de Fourier pour différentes valeurs de  $\alpha$ .

modèle, nous écrivons

$$F^{(+)}(q, t) = e^{-Dq^2t} (EISF(q) + (1 - EISF(q))E_\alpha(-(t/\tau)^\alpha)), \quad t \geq 0, \quad (\text{V.101})$$

où  $D$  est un coefficient de diffusion,  $\alpha \equiv \alpha(q)$  définit la forme de la fonction de relaxation,  $\tau \equiv \tau(q)$  définit l'échelle de temps de la fonction de relaxation et  $EISF(q)$  détermine l'intensité élastique. Il est possible d'ajuster les quatre paramètres en même temps, mais ça conduit à des ajustements avec des barres d'erreur importantes, en particulier pour le coefficient de diffusion  $D$ . C'est pour cette raison que nous avons décidé de réaliser des expériences de la diffusion dynamique de la lumière (DLS) pour obtenir le coefficient de diffusion et l'injecter dans nos ajustements, et à titre de comparaison, on a également fait un ajustement à quatre paramètres avec les coefficients de diffusion moyennés en  $q$ ,  $\overline{D_{\text{QENS}}(q)}$ , provenant de l'ajustement à quatre paramètres des spectres QENS. Les deux ajustements (trois et quatre paramètres) donnent des valeurs très similaires pour les trois paramètres.

Temp [K]	$D_{\text{DLS (D}_2\text{O)}} [\text{\AA}^2/\text{ns}]$	$D_{\text{QENS (D}_2\text{O)}} [\text{\AA}^2/\text{ns}]$	$D_{\text{DLS (TFE)}} [\text{\AA}^2/\text{ns}]$	$D_{\text{QENS (TFE)}} [\text{\AA}^2/\text{ns}]$
283	$3.31 \pm 0.053$	$6.04 \pm 0.46$	$2.05 \pm 0.049$	$5.59 \pm 0.73$
293	$5.72 \pm 0.057$	/	$3.17 \pm 0.145$	/
303	$7.49 \pm 0.090$	$6.15 \pm 0.80$	$4.42 \pm 0.19$	$5.91 \pm 0.70$
313	$9.85 \pm 0.049$	/	$6,32 \pm 0.292$	/
323	$0.31 \pm 0.013$ 11.94 (extrapol.)	$6.43 \pm 0.83$	$8.15 \pm 0.19$	$6.29 \pm 0.36$

TABLE 1. Les coefficients de diffusion obtenus par DLS et QENS pour le MBP dans les deux solvants et à différentes températures.

Le tableau 1 présente les deux coefficients de diffusion et les barres d'erreur de  $\overline{D_{\text{QENS}}(q)}$  et  $\overline{D_{\text{DLS}}}$ . La très faible valeur de  $D_{\text{DLS}}$  pour le solvant  $\text{D}_2\text{O}$  à 323 K indique la coagulation de la MBP et nous montrons également pour la même température la valeur extrapolée à partir de températures plus basses pour une solution non coagulée. La figure (V.5) montre les ajustements à trois paramètres de la fonction intermédiaire de diffusion déconvoluée et symétrisée  $F^{(+)}(q, t)$ . Tous les ajustements ont été effectués pour les deux solvants, et les trois températures différentes, 283 K, 303 K et 323 K. Les points bleus correspondent ici au  $\text{D}_2\text{O}$ , les points oranges au  $\text{D}_2\text{O}+30\%\text{TFE}$ , et les lignes continues correspondantes aux ajustements.

Commençons notre discussion par les résultats des constantes de diffusion dans les deux solvants et pour les trois températures, qui ont été obtenus par DLS. Nous rappelons que le MBP dans le solvant  $\text{D}_2\text{O}$  coagule à 323 K et que

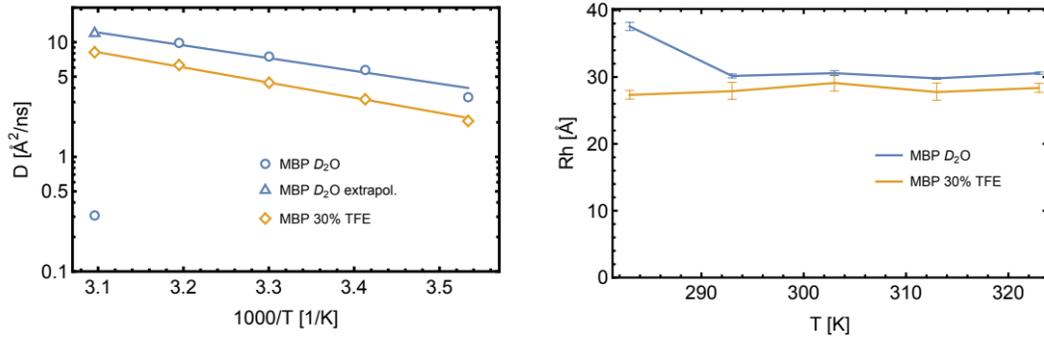


FIGURE V.4. **Partie gauche** : Tracé d'Arrhenius du coefficient de diffusion de la MBP dans le solvant  $D_2O$  (bleus) et le solvant  $D_2O+30\%TFE$  (jaunes). Les lignes continues représentent les lois d'Arrhenius ajustées correspondantes. **Partie droite** : Rayons hydrodynamiques effectifs de la MBP dans les deux solvants en fonction de la température [59].

la constante de diffusion correspondante a été estimée par un ajustement linéaire à partir des températures inférieures (Fig (V.1)). Notons que les expériences DLS ont été réalisées à des concentrations plus faibles que les expériences QENS (4 mg/ml pour  $D_2O$  et 7 mg/ml pour  $D_2O + 30\% TFE$ ), qui sont donc certainement affectées par la coagulation. Ce phénomène n'affecte pas nos résultats concernant la dynamique interne sondée par le spectromètre IN16B puisque même en fixant  $D = 0$ , c'est-à-dire en négligeant complètement la diffusion globale dans les ajustements des paramètres, les valeurs de ces derniers ne sont que légèrement modifiées. La partie gauche de la figure V.4 montre un tracé d'Arrhenius des constantes de diffusion, où pour le solvant  $D_2O$  à 323 K, le point triangulaire extrapolé a été utilisé à la place du point mesuré. Les points bleus se réfèrent au solvant  $D_2O$  et les points jaunes foncés au solvant  $D_2O + 30\%TFE$ . Les lignes continues correspondantes indiquent les

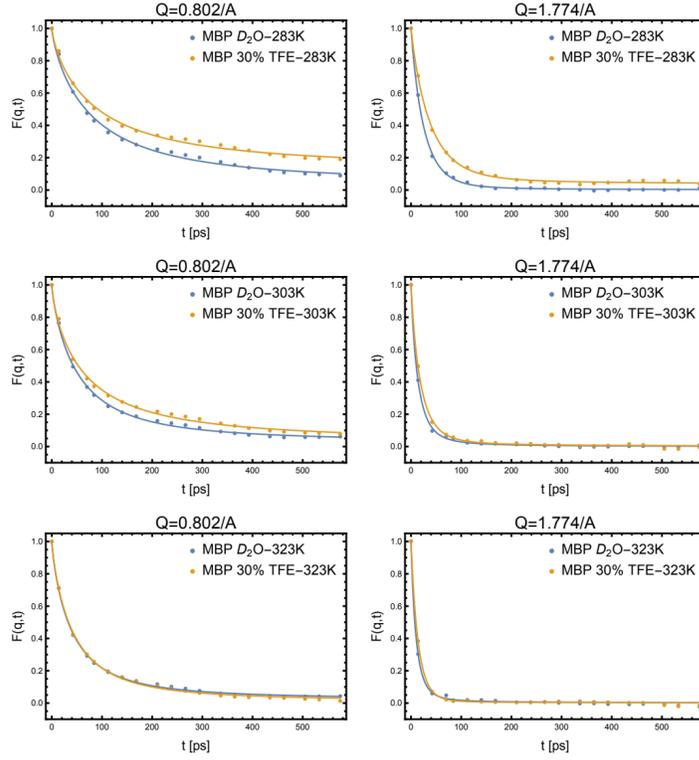


FIGURE V.5. Les fonctions intermédiaires de diffusion déconvoluées  $F^{(+)}(q, t)$  en comparant les deux solvants pour les valeurs minimales et maximales de  $q$ . La couleur bleue correspond au  $D_2O$  et la couleur jaune  $D_2O+30\%TFE$  [59].

ajustements d'une loi d'Arrhenius [64, 65]

$$D(T) = D_0 e^{-\frac{\Delta G}{k_B T}}, \quad (V.102)$$

où l'énergie d'activation est  $\Delta G = 5.10$  kcal/mol pour le  $D_2O$  and  $\Delta G = 6.05$  kcal/mol pour le  $D_2O+30\%TFE$ . Nous notons dans ce contexte que l'énergie d'activation pour le solvant  $D_2O$  est proche des valeurs qui ont été trouvées pour d'autres IDP [65] et correspond à l'énergie d'activation pour l'auto-diffusion dans l'eau liquide [66, 67]. La partie droite de la figure V.4 montre une estimation du rayon hydrodynamique effectif,  $R_h$ , du MBP en fonction de

la température et du solvant, en utilisant la relation de Stokes-Einstein

$$D = \frac{k_B T}{6\pi\eta R_h}, \quad (\text{V.103})$$

avec les coefficients de diffusion de la DLS et les valeurs expérimentales de la viscosité dynamique qui sont présentées dans la partie droite de la figure V.1. Les résultats montrent que le  $R_h$  est essentiellement constant avec l'augmentation de la température et entre 283 K et 303 K, il diminue même pour la MBP dans le D<sub>2</sub>O pur, un tel comportement a en fait été rapporté pour les IDPs [68, 69]. Il faut noter qu'aucune transition de repliement n'est observée avec l'augmentation de la température, ce qui est une caractéristique des protéines globulaires repliées. Ce résultat est en accord avec le résultat obtenu par la spectroscopie CD selon laquelle le contenu de la structure secondaire est essentiellement constant sur toute la gamme de température, étant légèrement plus important pour le solvant qui est enrichi en TFE ce qui est montré dans la figure V.6. On notant que le TFE est connu pour induire des hélices  $\alpha$  et on peut observer qu'à 283 K, mais cet effet disparaît progressivement avec l'augmentation de la température. L'effet du TFE consiste donc à déplacer des feuillets  $\beta$  en hélices  $\alpha$  et vice-versa, en gardant le contenu total de la structure secondaire constant.

Passant maintenant à la discussions des paramètres d'ajustement concernant la dynamique interne de la MBP, en comparant à nouveau les trois températures différentes, 283 K, 303 K et 323 K, et les deux solvants D<sub>2</sub>O et D<sub>2</sub>O+30%TFE qui sont montrés dans la figure V.7. Commençons par le premier paramètre qui est le paramètre d'échelle de temps  $\tau$  où il varie considérablement avec la température et  $q$ , la règle générale étant qu'il diminue avec la température et  $q$ . Cela peut se comprendre puisque baisser la température entraîne un ralentissement de la dynamique et que des valeurs plus élevées

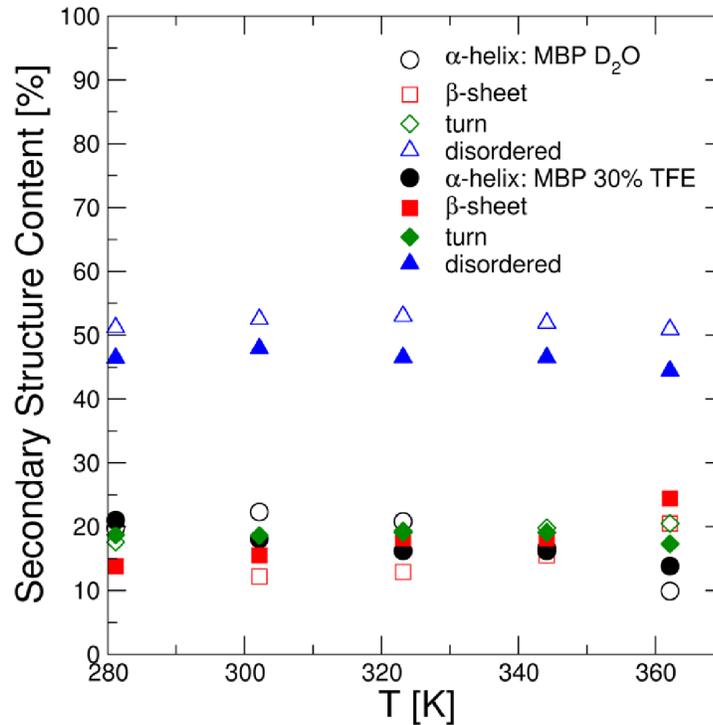


FIGURE V.6. Contenu de la structure secondaire de la MBP dans les deux solvants D<sub>2</sub>O (symboles creux) et D<sub>2</sub>O+30%TFE (symboles pleins) [59].

de  $q$  indiquent des mouvements de plus en plus localisés et de plus en plus rapides. Contrairement au paramètre d'échelle  $\tau$ , le paramètre de forme de la fonction de relaxation  $\alpha$ , varie en revanche faiblement avec la température et augmente avec  $q$  jusqu'à des valeurs proches de 1, ce qui indique une relaxation proche de l'exponentielle pour les mouvements localisés. Le résultat particulièrement intéressant est que la dynamique de relaxation du MBP dans les deux solvants devient similaire avec l'augmentation de la température et est presque identique à 323 K. Cette observation est cohérente avec le fait que

les rayons hydrodynamiques ainsi que le contenu et la composition des hélices  $\alpha$  et des feuillets  $\beta$  sont comparables. On constate également que la coagulation de la MBP, trouvée par DLS pour le  $D_2O$  à 323 K et certainement présente dans les échantillons QENS de plus forte concentration, n'a aucune d'influence sur la dynamique interne de la protéine. En ce qui concerne le paramètre EISF, nous observons qu'il est généralement proche de zéro pour la gamme  $q$  accessible, sauf pour le solvant le  $D_2O+30\%TFE$  à 283 K et pour de petites valeurs de  $q$ , où la protéine est plus compacte, ce qui indique qu'à basse température, le solvant  $D_2O+30\%TFE$  gêne un peu plus les mouvements de grande amplitude dus à une grande flexibilité interne, même dans une structure compacte par rapport au  $D_2O$ . Nous notons dans ce contexte que les EISF ne sont pas systématiquement faibles pour ce modèle et nous nous référons à des travaux sur une analyse de données QENS à partir de poudres hydratées d'acétylcholinestérase humaine [58,70] qui est une enzyme avec une structure bien définie. Dans ce cas, l'EISF s'est avéré clairement non nul, comme on peut s'y attendre pour une telle protéine qui est plus compacte et pour un tel échantillon, où les mouvements globaux de la protéine sont empêchés. Pour compléter la discussion sur les paramètres d'ajustement, on montre dans la figure V.8 les spectres de taux de relaxation  $p(\lambda)$  qui sont définis par l'équation (IV.84). Pour les deux solvants et toutes les températures, nous observons que le pic du spectre de taux de relaxation se déplace vers des valeurs plus élevées avec l'augmentation de  $q$ , et que sa largeur diminue. Ceci reflète le fait que le paramètre de forme  $\alpha$  s'approche de 1 avec l'augmentation de  $q$ , en notant que  $\alpha = 1$  correspond à une relaxation exponentielle et à un spectre de taux de relaxation parfaitement monodispersé de la forme  $p(\lambda) = \delta(\lambda - 1/\tau)$ .

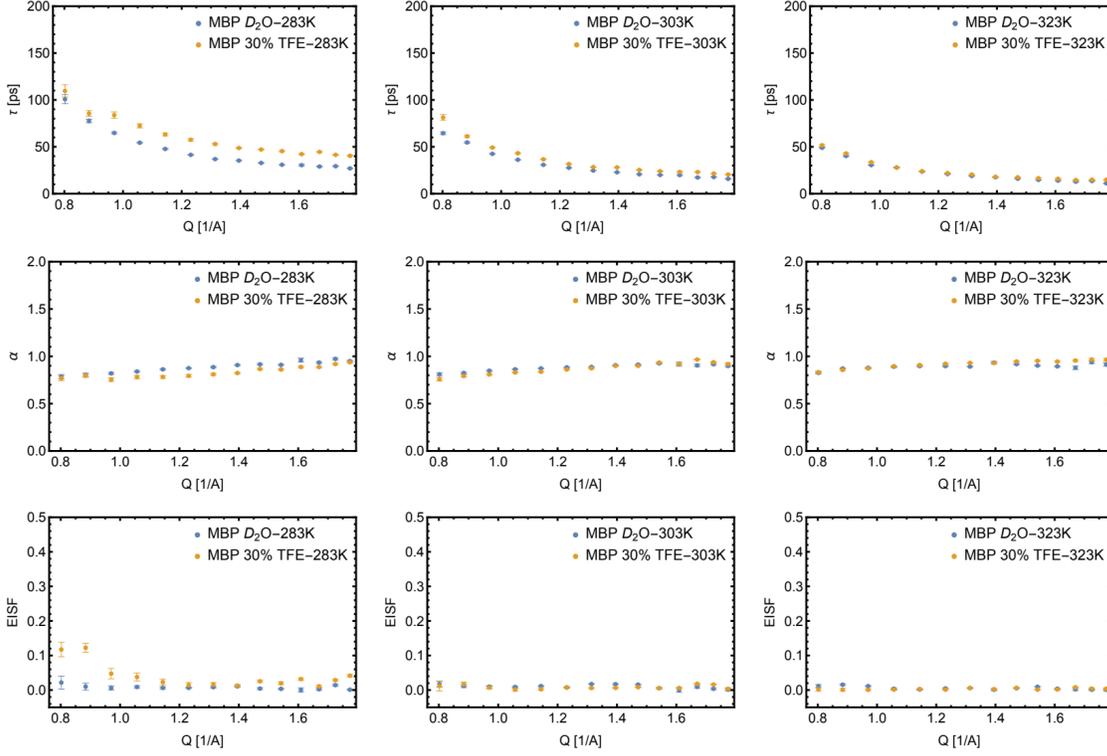


FIGURE V.7. De haut en bas, les paramètres du modèle  $\tau$ ,  $\alpha$ , et  $EISF$  pour le MBP à différentes températures en comparant les deux solvants,  $D_2O$  et  $D_2O+30\%TFE$  deutéré [59].

Nous pouvons maintenant donner une interprétation physique du paramètre de forme  $\alpha$ , en nous référant au concept de paysages énergétiques protéiques proposé par Frauenfelder et ses collaborateurs [46]. Dans ce contexte, nous considérons la variable

$$\chi(t) = a_{\mathbf{q}}(t) - \langle a_{\mathbf{q}}(t) \rangle_{cl}, \quad \text{où} \quad a_{\mathbf{q}}(t) = e^{i\mathbf{q} \cdot \mathbf{r}_a(t)}, \quad (\text{V.104})$$

où  $\mathbf{r}_a$  désigne la position d'un atome "représentatif" et  $\langle \dots \rangle_{cl}$  un moyenne sur un ensemble classique, et nous supposons que la dynamique de  $\chi$  peut être modélisée comme un processus fractionnaire d'Ornstein-Uhlenbeck (fOU) pour la partie réelle et imaginaire de  $\chi$ , qui sera détaillé dans ce chapitre. Ce

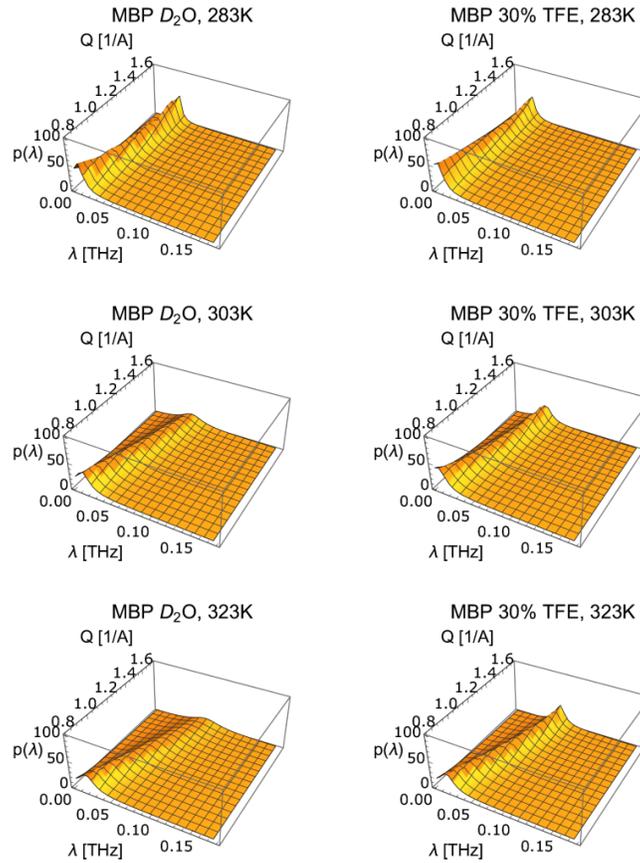


FIGURE V.8. De haut en bas, spectres de taux de relaxation  $p(\lambda)$  pour  $T = 283K$ ,  $T = 303K$ , and  $T = 323K$ , respectivement, et pour les deux solvants  $D_2O$  (partie gauche) et  $D_2O+30\%TFE$  (partie droite) [59].

dernier peut être interprété comme un processus de diffusion dans un potentiel parabolique "rugueux" ou "paysage énergétique", qui tend à ramener  $a(t)$  à sa valeur moyenne,  $\langle a \rangle$ . Comme  $\chi$  est une fonction complexe, les parties réelle et imaginaire doivent être considérées comme des variables indépendantes et des modèles stochastiques peuvent être utilisés pour l'évolution temporelle du

vecteur réel à deux composantes

$$\boldsymbol{\chi} \equiv \begin{pmatrix} \Re\{a_{\mathbf{q}} - \langle a_{\mathbf{q}} \rangle_{\text{cl}}\} \\ \Im\{a_{\mathbf{q}} - \langle a_{\mathbf{q}} \rangle_{\text{cl}}\} \end{pmatrix}, \quad (\text{V.105})$$

en supposant que le  $\langle a_{\mathbf{q}} \rangle_{\text{cl}}$  n'est pas strictement nul. En termes de variable vectorielle  $\boldsymbol{\chi}$ , la fonction intermédiaire de diffusion peut être exprimée comme suit

$$F(\mathbf{q}, t) = |\langle a_{\mathbf{q}} \rangle_{\text{cl}}|^2 + \langle \boldsymbol{\chi}^T(0) \cdot \boldsymbol{\chi}(t) \rangle_{\text{cl}} \quad (\text{V.106})$$

La forme explicite de la fonction d'autocorrélation de  $\boldsymbol{\chi}$  est la suivante

$$\langle \boldsymbol{\chi}^T(0) \cdot \boldsymbol{\chi}(t) \rangle_{\text{cl}} = \int \int d^2\boldsymbol{\chi}_0 d^2\boldsymbol{\chi}_t \boldsymbol{\chi}_0^T \cdot \boldsymbol{\chi}_t p(\boldsymbol{\chi}, t | \boldsymbol{\chi}_0, 0) p_{\text{eq}}(\boldsymbol{\chi}_0) \quad (\text{V.107})$$

Comme indiqué ci-dessus, nous avons choisi ici le processus fractionnaire d'Ornstein-Uhlenbeck [71, 72] qui décrit une diffusion anormale, non markovienne, dans un potentiel de biais harmonique,

$$V(\boldsymbol{\chi}) = \frac{K_{\boldsymbol{\chi}}}{2} |\boldsymbol{\chi}|^2 \quad (K_{\boldsymbol{\chi}} > 0),$$

et tend à rétablir la configuration d'équilibre correspondant à  $\langle \boldsymbol{\chi} \rangle = \mathbf{0}$ . La constante de force  $K_{\boldsymbol{\chi}}$  a la dimension d'une énergie et l'équation fractionnaire de Fokker-Planck

$$\partial_t p = {}_0\partial_t^{1-\alpha} \left\{ \eta_{\boldsymbol{\chi}}^{(\alpha)} \frac{\partial}{\partial \boldsymbol{\chi}} \cdot \{\boldsymbol{\chi} p\} + D_{\boldsymbol{\chi}}^{(\alpha)} \frac{\partial}{\partial \boldsymbol{\chi}} \cdot \frac{\partial p}{\partial \boldsymbol{\chi}} \right\}. \quad (\text{V.108})$$

Ici  $0 < \alpha < 1$  et  ${}_0\partial_t^{1-\alpha}$  désigne une dérivée fractionnaire de l'ordre  $1 - \alpha$  par rapport à  $t$ . Ici,  $\eta_{\boldsymbol{\chi}}^{(\alpha)}$  désigne une constante de relaxation fractionnaire, avec une dimension physique de  $1/s^\alpha$ .

$$D_{\boldsymbol{\chi}}^{(\alpha)} = \eta_{\boldsymbol{\chi}}^{(\alpha)} k_B T / K_{\boldsymbol{\chi}} = \eta_{\boldsymbol{\chi}}^{(\alpha)} \langle |\boldsymbol{\chi}|^2 \rangle_{\text{cl}}$$

est un coefficient de diffusion fractionnaire. Pour  $0 < \alpha < 1$ , la dérivée fractionnaire de Riemann-Liouville

$${}_0\partial_t^{1-\alpha} f(t) \equiv \frac{d}{dt} \int_0^t dt' \frac{(t-t')^{\alpha-1}}{\Gamma(\alpha)} f(t')$$

représente les effets mémoire à long terme et la fonction d'autocorrélation normalisée (ACF) de  $\chi$  a la forme d'une fonction de Mittag-Leffler (ML) étirée [55]

$$\phi(t) \equiv \langle \chi^T(0) \cdot \chi(t) \rangle_{\text{cl}} / \langle |\chi|^2 \rangle_{\text{cl}} = E_\alpha(-\eta_\alpha t^\alpha), \quad \eta_\alpha \equiv \eta_\alpha(\mathbf{q}), \quad (\text{V.109})$$

qui interpole entre une exponentielle étirée à court terme et une loi de puissance inverse à long terme et converge vers une simple exponentielle (relaxation de Debye) pour  $\alpha \rightarrow 1$ .

La partie gauche de la figure V.9 représente une rugosité du potentiel parabolique en fonction de  $\Re\{\chi\}$  et  $\Im\{\chi\}$ . Si la variable  $\chi$  diffusait dans un potentiel parabolique lisse, alors sa fonction d'autocorrélation temporelle serait une fonction exponentielle  $\langle \chi^*(0)\chi(t) \rangle \propto \exp(-t/\tau)$ , et la diffusion dans le potentiel rugueux conduit à des effets mémoire à long terme, qui sont décrits par une fonction de relaxation ML, c'est-à-dire  $\langle \chi^*(0)\chi(t) \rangle \propto E_\alpha(-(t/\tau)^\alpha)$ . Comme déjà discuté, on peut trouver une distribution des barrières d'énergie dans le potentiel parabolique "rugueux" en utilisant le modèle de diffusion proposé par R. Zwanzig [57] qui décrit la diffusion d'une particule dans un potentiel harmonique rugueux avec une barrière d'énergie fixe, en introduisant l'hauteur de barrière d'énergie sans dimension  $\epsilon = \frac{\Delta E}{k_B T}$ , et en utilisant la distribution des barrières énergétiques correspondant à la fonction de relaxation ML donnée ci-dessus IV.93, ce qui est montré dans la partie droite de la figure V.9, cette figure montre que  $\lim_{\alpha \rightarrow 1} P_{\text{ML}}(\epsilon) = \delta(\epsilon)$ , c'est-à-dire qu'elle devient de plus en plus la distribution des barrières énergétiques pour un potentiel lisse avec une hauteur de barrière nulle, alors que pour  $\alpha \rightarrow 0$  la distribution  $P_{\text{ML}}(\epsilon)$  devient

très large, incluant aussi des barrières d'une hauteur presque infinie. Dans ce contexte, les paramètres d'ajustement présentés dans la figure V.7 montrent que les mouvements localisés sont caractérisés par un potentiel plus lisse que les mouvements globaux ( $\alpha$  augmente avec  $q$ ).

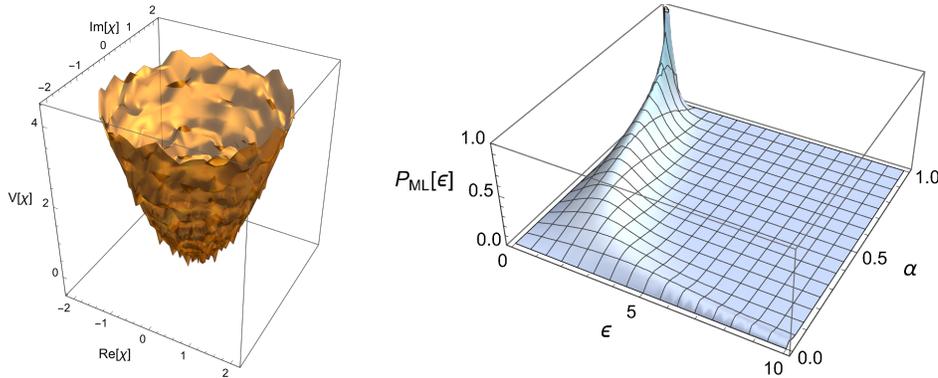


FIGURE V.9. **Partie gauche** : Potentiel harmonique pour les parties réelle et imaginaire de  $\delta\chi(\mathbf{q}, t)$ . **Partie droite** : Distribution de la barrière énergétique  $P(\epsilon, \alpha)$  obtenu par Eq. (IV.93) [59].

Pour conclure et résumer cette partie dont nous avons présenté une analyse des spectres de diffusion quasi-élastique de neutrons à haute résolution de la MBP, cette étude montre que les modèles multi-échelles "minimalistes" pour la dynamique de relaxation des protéines, combinés aux résultats de la diffusion dynamique de la lumière, donnent des indications significatives sur la dynamique des protéines intrinsèquement désordonnées et permettent en particulier de détecter et d'interpréter les changements dans la dynamique de relaxation en fonction du solvant, de la température et aussi de la résolution. L'analyse QENS a montré qu'un solvant enrichi en trifluoroéthanol à 283 K rend le paysage énergétique plus rugueux, ce qui entraîne une dynamique de relaxation plus lente avec un spectre de taux de relaxation plus dispersé, et que cet effet disparaît continuellement avec l'augmentation de la température.

Suite à cette observation, les expériences complémentaires de diffusion dynamique de la lumière et les expériences SRCD montrent que le TFE réduit à 283 K le rayon d'hydratation de la MBP et augmente le contenu en hélices  $\alpha$  au détriment des feuillets  $\beta$ . Ces effets disparaissent avec l'augmentation de la température, où les rayons hydrodynamiques sont comparables et constants. Indépendamment de l'ajout ou non de trifluoroéthanol au  $D_2O$ , il y a maintenant une transition de dépliage avec l'augmentation de la température, qui se manifesterait par une augmentation soudaine du rayon d'hydratation et qui est typique des protéines repliées.

## Chapitre 6

### **EISF mesuré et réel**

#### **1. Elimination de la diffusion élastique comme paramètre du modèle**

Dans l'interprétation classique de la diffusion de neutrons, la diffusion élastique par les protéines reflète les amplitudes de mouvement résultant de leur dynamique interne et elle est observable si la diffusion globale des molécules entières est bloquée ou ne peut être résolue par le spectromètre considéré. En raison d'une résolution instrumentale limitée, l'amplitude de diffusion élastique mesurée contient toujours des contaminations provenant de la diffusion quasi-élastique de neutrons, dans ce contexte, un modèle a été supposé pour extraire la contrepartie corrigée de la résolution des spectres expérimentaux correspondants. Nous avons décrit une méthode quasi-analytique, en supposant que la fonction intermédiaire de diffusion relaxe avec la fonction de Mittag-Leffler "étirée" IV.75, vers l'amplitude élastique et que la fonction de résolution instrumentale a une forme gaussienne. La fonction correspondante peut être intégrée dans une procédure d'ajustement et permet d'éliminer l'intensité élastique en tant que paramètre d'ajustement. Un point important est que la méthode peut être utilisée pour les spectromètres pour lesquels l'impact de la diffusion globale est à la limite d'être détectable, comme dans les expériences QENS de la MBP décrites auparavant. Dans cette partie, nous allons illustrer cette méthode pour l'analyse de ces expériences, confirmant les résultats discutés dans le chapitre précédent [59] ainsi que pour la myoglobine

en solution. Notre choix était basé sur cette protéine car la myoglobine a à peu près le même poids que la MBP, mais contrairement à cette dernière, il s'agit d'une protéine globulaire avec une structure bien définie.

**1.1. Diffusion Pseudo-élastique.** On considère un facteur de structure dynamique mesuré, qui est toujours élargi en raison de la résolution instrumentale finie. En définissant  $\tilde{R}(\omega)$  comme la fonction de résolution instrumentale et en omettant pour simplifier la dépendance en  $q$  des quantités pertinentes, le facteur de structure dynamique mesuré est donné par la convolution en fréquence du facteur de structure dynamique réel et de la fonction de résolution,

$$\tilde{F}_m(\omega) = (\tilde{R} * \tilde{F})(\omega) \equiv \int_{-\infty}^{+\infty} d\omega' \tilde{R}(\omega - \omega') \tilde{F}(\omega'). \quad (\text{VI.110})$$

avec (IV.74) on obtient dans un premier temps

$$\tilde{F}_m(\omega) = F(\infty)\tilde{R}(\omega) + (F(0) - F(\infty))(\tilde{R} * \tilde{\phi})(\omega). \quad (\text{VI.111})$$

On définit maintenant l'intensité élastique mesurée par l'intégrale

$$F_m(\infty) \equiv \int_{-\epsilon}^{+\epsilon} d\omega \tilde{F}_m(\omega), \quad (\text{VI.112})$$

où  $\epsilon > 0$  est défini de telle sorte que

$$\int_{-\epsilon}^{+\epsilon} d\omega \tilde{R}(\omega) < 1, \quad (\text{VI.113})$$

et le facteur de structure statique total qui est mesuré par

$$F_m(0) \equiv \int_{\omega_{\min}}^{\omega_{\max}} d\omega \tilde{F}_m(\omega), \quad (\text{VI.114})$$

où  $[\omega_{\min}, \omega_{\max}]$  est la gamme dynamique de l'instrument. Il découle donc de la forme générique VI.111 du facteur de structure dynamique mesuré et élargi à la résolution que

$$F_m(\infty) \approx F(\infty) + (F_m(0) - F(\infty))\xi, \quad (\text{VI.115})$$

où  $\xi$  est la contribution pseudo-élastique due à la résolution instrumentale finie,

$$\xi = \int_{-\epsilon}^{+\epsilon} d\omega (\tilde{R} * \tilde{\phi})(\omega). \quad (\text{VI.116})$$

En supposant que cette contribution puisse être calculée de manière fiable sur la base de modèles appropriés pour la fonction de relaxation et la résolution instrumentale, l'intensité élastique "réelle" peut être estimée comme suit

$$F(\infty) \approx \frac{F_m(\infty) - \xi F_m(0)}{1 - \xi}. \quad (\text{VI.117})$$

**1.2. La contribution du modèle pseudo-élastique.** L'estimation de la valeur plateau d'une fonction à partir de données expérimentales avec un intervalle de temps fini limité par un instrument est un peu délicate et il est souhaitable d'avoir un contrôle de cohérence dans les données expérimentales utilisées. Ceci peut être réalisé si la contribution pseudo-élastique  $\xi$ , peut être efficacement corrigée pour la fonction de relaxation du modèle donné, de telle sorte que l'estimation VI.117 puisse être intégrée dans la procédure d'ajustement. Dans ce contexte, nous avons supposé que la fonction de résolution est bien représentée par une fonction gaussienne,

$$\tilde{R}(\omega) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{\omega^2}{2\sigma^2}} \longleftrightarrow R(t) = e^{-\frac{\sigma^2 t^2}{2}}, \quad (\text{VI.118})$$

où  $\sigma$  est approximativement la largeur à mi-maximum (HWHM) de l'instrument considéré. En travaillant avec des spectres de diffusion de neutrons symétrisés et la fonction de relaxation modèle  $\phi_{\text{ML}}(t)$ , la contribution pseudo-élastique devient alors une fonction des paramètres  $\alpha$ ,  $\tau$ , et  $\sigma$ ,

$$\xi_{\text{ML}}(\tau, \alpha, \sigma) = \int_{-\epsilon}^{+\epsilon} d\omega (\tilde{R} * \tilde{\phi}_{\text{ML}})(\omega), \quad (\text{VI.119})$$

sachant que

$$F(\infty; \tau, \alpha, \sigma) \approx \frac{F_m(\infty) - \xi_{\text{ML}}(\tau, \alpha, \sigma) F_m(0)}{1 - \xi_{\text{ML}}(\tau, \alpha, \sigma)}, \quad (\text{VI.120})$$

remplace le paramètre d'ajustement  $F(\infty)$ . Pour une fonction gaussienne, nous avons

$$\int_{-3\sigma}^{+3\sigma} d\omega \tilde{R}(\omega) \approx 0.9978, \quad (\text{VI.121})$$

de sorte que  $\epsilon = 3\sigma$  est un bon choix.

**1.3. Calcul de  $\xi_{\text{ML}}$ .** Afin d'obtenir une formule quasi-analytique pour la contribution pseudo-élastique  $\xi_{\text{ML}}$  qui est définie dans l'Eq. (VI.119), nous introduisons

$$\tilde{W}(\omega) = \Theta\left(1 - \frac{|\omega|}{\epsilon}\right) \longleftrightarrow W(t) = \frac{2 \sin(\epsilon t)}{t}, \quad (\text{VI.122})$$

où  $\epsilon > 0$  est la fréquence de coupure. Avec le théorème de Parseval et le théorème de convolution de la transformée de Fourier,

$$(\tilde{R} * \tilde{\phi}_{\text{ML}})(\omega) \longleftrightarrow R(t)\phi_{\text{ML}}(t), \quad (\text{VI.123})$$

on écrit dans un premier temps

$$\xi_{\text{ML}} = \int_{-\infty}^{+\infty} d\omega \tilde{W}(\omega)(\tilde{R} * \tilde{\phi}_{\text{ML}})(\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dt W(t)R(t)\phi_{\text{ML}}(t). \quad (\text{VI.124})$$

On définit la fonction auxiliaire

$$H(t) \equiv W(t)R(t)\phi_{\text{ML}}(t), \quad (\text{VI.125})$$

et en notant que toutes les fonctions sont paires en temps, donc

$$\xi_{\text{ML}} = \frac{1}{\pi} \int_0^{\infty} dt e^{-st} H(t) \Big|_{s=0}. \quad (\text{VI.126})$$

La contribution pseudo-élastique  $\xi_{\text{ML}}$  peut donc être écrite comme la transformée de Laplace d'un produit de trois fonctions, évaluées à  $s = 0$ . Nous utilisons maintenant que pour une paire de fonctions,  $f(t)$  et  $g(t)$ ,

$$\int_0^{\infty} dt e^{-st} f(t)g(t) = \frac{1}{2\pi i} \oint_C ds' \hat{f}(s-s')\hat{g}(s'), \quad (\text{VI.127})$$

où  $C$  entoure toutes les singularités de l'intégrale, et on peut utiliser cette formule en deux étapes :

$$(1) \quad \hat{\phi}_{ML}^{(R)}(s) = \frac{1}{2\pi i} \oint_C ds' \hat{\phi}_{ML}(s-s') \hat{R}(s'), \quad (\text{VI.128})$$

$$(2) \quad \xi_{ML} = \frac{1}{\pi} \left\{ \frac{1}{2\pi i} \oint_C ds' \hat{\phi}_{ML}^{(R)}(-s') \hat{W}(s') \right\}. \quad (\text{VI.129})$$

Une forme analytique est connue pour la transformée de Laplace de la fonction de relaxation du modèle  $\hat{\phi}_{ML}(s)$ , pour la transformée de Laplace de la fenêtre temporelle de résolution du modèle,

$$\hat{R}(s) = \frac{1}{\sigma} \sqrt{\frac{\pi}{2}} e^{\frac{s^2}{2\sigma^2}} \operatorname{erfc} \left( \frac{s}{\sqrt{2}\sigma} \right), \quad (\text{VI.130})$$

et pour la transformée de Laplace de  $W(t)$  on a

$$\hat{W}(s) = 2 \operatorname{arccot} \left( \frac{s}{\epsilon} \right). \quad (\text{VI.131})$$

Notant qu'il est impossible de trouver une forme analytique des intégrales de contour (VI.128) et (VI.129), mais une bonne approximation quasi-analytique peut être obtenue en remplaçant  $\hat{R}(s)$  and  $\hat{W}(s)$  par des approximations de Padé [73],

$$\hat{R}(s) \approx \frac{1}{\sigma} \frac{P(s/\sigma)}{Q(s/\sigma)}, \quad (\text{VI.132})$$

$$\hat{W}(s) \approx \frac{P'(s/\epsilon)}{Q'(s/\epsilon)}, \quad (\text{VI.133})$$

où  $P(\cdot), Q(\cdot), P'(\cdot), Q'(\cdot)$  sont des polynômes. La méthode a été récemment utilisée pour calculer une bonne approximation de la transformée de Fourier élargie à la résolution  $\tilde{\phi}_{ML}^{(R)}(\omega)$  [70]. En introduisant des variables d'intégration à échelle appropriée, les intégrales de contour (VI.128) et (VI.129) peuvent alors

être évaluées par le théorème des résidus de l'analyse complexe. On trouve comme résultat

$$\xi_{\text{ML}} \approx \frac{1}{\pi} \sum_{j=1}^m \sum_{k=1}^{m'} \chi \lambda c_j d_k \hat{\Phi}_{\text{ML}}(-\chi(u_j + \lambda v_k)), \quad (\text{VI.134})$$

où  $\hat{\Phi}_{\text{ML}}(\cdot)$  est la version sans échelle de la transformée de Laplace la fonction de relaxation de  $\hat{\Phi}_{\text{ML}}(\cdot)$ ,

$$\hat{\Phi}_{\text{ML}}(u) = \frac{1}{u(1 + u^{-\alpha})}, \quad (\text{VI.135})$$

avec  $\chi$  et  $\lambda$  sont les paramètres d'échelle sans dimension

$$\chi = \sigma\tau \quad \text{and} \quad \lambda = \frac{\epsilon}{\sigma}, \quad (\text{VI.136})$$

et  $\{u_j\}$  et  $\{v_k\}$  sont les racines des polynômes  $Q(u)$  et  $Q'(v)$ , respectivement.

Les coefficients

$$c_j = \frac{P(u_j)}{\prod_{k=1, k \neq j}^m (u_j - u_k)}, \quad (\text{VI.137})$$

$$d_k = \frac{P'(v_k)}{\prod_{l=1, l \neq k}^{m'} (v_k - v_l)}, \quad (\text{VI.138})$$

sont les résidus des expressions sans dimension  $P(u)/Q(u)$  et  $P'(v)/Q'(v)$ , évaluées aux racines respectives des polynômes du dénominateur. dénominateur.

On a donc  $Q(u_j) = 0$  et  $Q'(v_k) = 0$ . L'expression finale (VI.134) pour  $\xi_{\text{ML}}$  est donc la superposition linéaire de termes simples de la forme (VI.135). Le codage de  $\xi_{\text{ML}}$  en tant que fonction compilée conduit à des temps d'exécution suffisamment courts pour permettre l'intégration de cette fonction dans une procédure d'ajustement.

**Dérivation de l'expression (VI.134)** Nous partons de l'équation eq. (VI.128) et introduisons l'expression (VI.132) pour l'approximation de

Padé de la fenêtre temporelle instrumentale de la transformée de Laplace. En outre, nous utilisons que

$$\hat{\phi}_{\text{ML}}(s) = \tau \hat{\Phi}_{\text{ML}}(s\tau) \quad (\text{VI.139})$$

pour obtenir dans un premier temps

$$\begin{aligned} \hat{\phi}_{\text{ML}}^{(R)}(s) &= \frac{1}{2\pi i} \oint_C \frac{ds'}{\sigma} \hat{\phi}_{\text{ML}}(s-s') \frac{P(s'/\sigma)}{Q(s'/\sigma)} \\ &\stackrel{s'/\sigma \rightarrow u}{=} \frac{1}{2\pi i} \oint_C du \tau \hat{\Phi}_{\text{ML}}(\tau(s-\sigma u)) \frac{P(u)}{Q(u)} \\ &= \frac{1}{\sigma} \left\{ \frac{1}{2\pi i} \oint_C du \chi \hat{\Phi}_{\text{ML}} \left( \chi \left( \frac{s}{\sigma} - u \right) \right) \frac{P(u)}{Q(u)} \right\}. \end{aligned}$$

En écrivant  $Q(u) = \prod_{k=1}^m (u - u_k)$  et en définissant les coefficients

$$c_j = \frac{P(u_j)}{\prod_{k=1, k \neq j}^m (u_j - u_k)}$$

À partir du théorème des résidus de l'analyse complexe on a

$$\hat{\phi}_{\text{ML}}^{(R)}(s) = \frac{1}{\sigma} \sum_{j=1}^m \chi c_j \hat{\Phi} \left( \chi \left( \frac{s}{\sigma} - u_j \right) \right). \quad (\text{VI.140})$$

Cette expression est maintenant insérée dans l'intégrale de contour (VI.129), et des changements appropriés des variables d'intégration conduisent à

$$\begin{aligned} \xi_{ML} &= \frac{1}{\pi} \left\{ \frac{1}{2\pi i} \oint_C ds' \left\{ \frac{1}{\sigma} \sum_{j=1}^m \chi c_j \hat{\Phi} \left( \chi \left( -\frac{s'}{\sigma} - u_j \right) \right) \right\} \frac{P'(s'/\epsilon)}{Q'(s'/\epsilon)} \right\} \\ &\stackrel{s'/\sigma \rightarrow u}{=} \frac{1}{\pi} \left\{ \frac{1}{2\pi i} \oint_C du \left\{ \sum_{j=1}^m \chi c_j \hat{\Phi} (\chi(-u - u_j)) \right\} \frac{P'(u\sigma/\epsilon)}{Q'(u\sigma/\epsilon)} \right\} \\ &\stackrel{\epsilon/\sigma \rightarrow \lambda, u/\lambda \rightarrow v}{=} \frac{1}{\pi} \left\{ \frac{1}{2\pi i} \oint_C dv \left\{ \sum_{j=1}^m \lambda \chi c_j \hat{\Phi} (\chi(-\lambda v - u_j)) \right\} \frac{P'(v)}{Q'(v)} \right\}. \end{aligned}$$

En écrivant  $Q'(v) = \prod_{l=1}^{m'} (v - v_l)$  et en définissant les coefficients

$$d_k = \frac{P(u_k)}{\prod_{l=1, l \neq k}^{m'} (v_k - v_l)},$$

nous obtenons à nouveau le résultat final à partir du théorème des résidus

$$\xi_{\text{ML}} = \frac{1}{\pi} \sum_{j=1}^m \sum_{k=1}^{m'} \lambda \chi c_j d_k \hat{\Phi}(-\chi(u_j + \lambda v_k)), \quad (\text{VI.141})$$

où  $\chi \equiv \sigma\tau$  et  $\lambda \equiv \epsilon/\sigma$ .

**1.4. Test.** Pour obtenir une image systématique de la contribution pseudoélastique en fonction de  $\tau$  et  $\alpha$ , nous pouvons la calculer, selon l'équation VI.134 pour  $\alpha = k \times 0.1$ ,  $k = 0, \dots, 10$ . Dans ce cas, nous avons défini  $\sigma$  comme étant la résolution (HWHM) du spectromètre IN16B  $\sigma = 1.75 \mu\text{eV}$ , et nous avons fixé  $\epsilon = 3\sigma$ . Nous faisons alors varier  $\chi \in [\sigma\tau_{\min}, \sigma\tau_{\max}]$ , où  $\tau_{\min} = 0.1 \text{ ps}$  et  $\tau_{\max} = 10^4 \text{ ps}$  (lignes continues). Pour comparer, nous avons calculé  $\xi_{\text{ML}}$  par intégration numérique de  $\phi_{\text{ML}}^{(R)}(\omega)$ , en choisissant les mêmes valeurs pour  $\alpha$ ,  $\sigma$ , et  $\epsilon$  et en fixant  $\tau$  aux valeurs discrètes  $\tau = 10^j \text{ ps}$ ,  $j = -1, \dots, 4$  (les points). On peut définir les matrices  $\xi_{\text{ML}} \equiv (\xi_{\text{ML}}[j, k])$  et  $\xi_{\text{ML}}^{(\text{i.n.})} \equiv (\xi_{\text{ML}}^{(\text{i.n.})}[j, k])$ , où "i.n." signifie "intégration numérique", nous trouvons que

$$\|\xi_{\text{ML}}^{(\text{i.n.})} - \xi_{\text{ML}}\| = 3.45 \times 10^{-6},$$

où  $\|\dots\|$  est définie comme la valeur singulière maximale de la matrice. Les résultats des calculs sont présentés dans la figure VI.1. On observe que pour les "bonnes résolutions", où  $\sigma\tau < 1$ , les contributions pseudoélastiques augmentent avec la diminution de  $\alpha$ , et que l'inverse est vrai pour les "mauvaises résolutions", où  $\sigma\tau > 1$ . Tous les calculs ont été effectués avec des approximations de Padé d'ordre  $m = 8$  pour les polynômes du dénominateur  $Q(u)$  et  $Q'(u)$  et d'ordre  $n = 8$  pour les polynômes du numérateur correspondants,  $P(u)$  et  $P'(u)$ , en choisissant  $s = 1$  comme point de référence. En utilisant l'approximation de Padé pour la fonction de résolution par le biais de

$$\tilde{R}_{\text{Padé}}(\omega) \equiv \frac{1}{\pi} \Re \left\{ \frac{P(i\omega)}{Q(i\omega)} \right\}$$

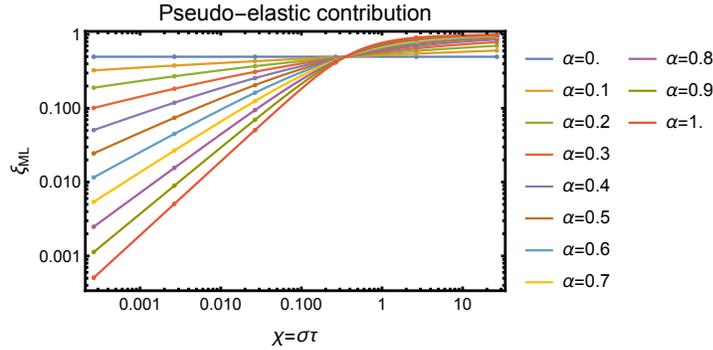


FIGURE VI.1. Contribution pseudo-élastique  $\xi_{ML}$  en fonction de  $\chi \equiv \sigma\tau$  and  $\alpha$  [54].

on trouve que

$$|\tilde{R}_{\text{Padé}}(\omega) - \tilde{R}(\omega)| < 10^{-6}$$

dans le domaine  $\omega$  concerné.

## 2. Résultats de l'analyse de données QENS de la MBP

Pour illustrer la contribution pseudo-élastique à la diffusion élastique, nous considérons un exemple concret lié à notre première étude QENS qui est basée sur la MBP en solution aqueuse [59]. Comme indiqué dans le chapitre précédent, la constante de diffusion  $D$ , de la MBP a été mesurée séparément par diffusion dynamique de la lumière (DLS) et a ensuite été injectée dans l'ajustement, et les spectres QENS ont été ajustés directement avec le modèle (IV.56). L'hypothèse est ici que les mouvements globaux et internes ne sont pas corrélés. Comme mentionné précédemment et dans la référence [59], les ajustements résultants pour  $EISF$ ,  $\alpha$  et  $\tau$  ne varient que très peu si la constante de diffusion est simplement négligée. Ceci est illustré dans la figure VI.2 qui montre un tracé log-log du spectre de Fourier de la fonction de relaxation du modèle ajusté pour le MBP dans  $D_2O$  pour  $T = 283$  et  $q = 1.2/\text{Å}$  (ligne

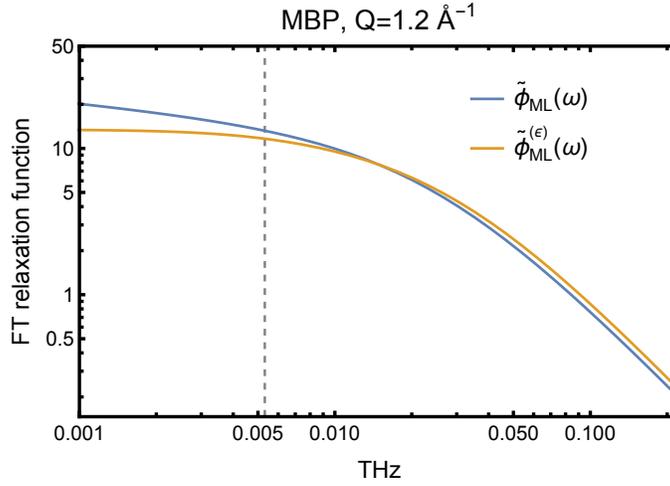


FIGURE VI.2. Impact de la diffusion globale sur le spectre de Fourier de la fonction de relaxation modèle pour la MBP dans le  $D_2O$  à  $T = 283$  K et  $q = 1.2/\text{Å}$  [54].

bleue), ainsi que la contrepartie correspondante, élargie par diffusion, résultant du facteur d'amortissement  $\exp(-Dq^2|t|)$  de la fonction intermédiaire de diffusion (ligne jaune). Nous avons pris ici  $D = 3.3 \text{ Å}^2/\text{ns}$  obtenu par DLS. La ligne verticale indique la résolution instrumentale (FWHM).

L'impact de la diffusion pseudoélastique sur les intensités élastiques observées est illustré dans la figure VI.3, qui montre à nouveau le spectre de Fourier de la fonction de relaxation du modèle ajusté  $\tilde{\phi}_{ML}(\omega)$  pour les mêmes paramètres que dans la figure VI.2 (ligne bleue), ainsi que la fonction de résolution du modèle (ligne jaune), où la résolution instrumentale (HWHM) correspond à  $\sigma = 0.0027$  THz. La ligne rouge est le spectre élargi par la résolution  $\tilde{\phi}_{ML}^{(R)}(\omega)$ , et la zone en rouge clair est la contribution pseudoélastique correspondante, qui est pour cet exemple  $\xi_{ML} \approx 0.47$  avec  $\epsilon = 3\sigma$ . La différence entre le spectre du modèle et sa version élargie à la résolution doit également être remarquée.

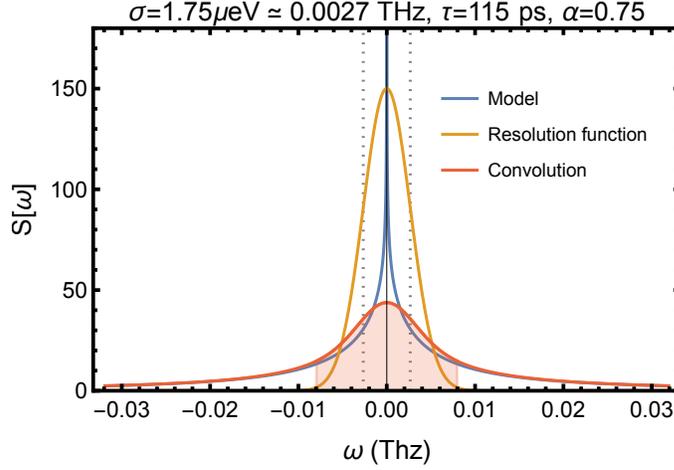


FIGURE VI.3. Fonction de relaxation du modèle  $\tilde{\phi}_{\text{ML}}(\omega)$  (ligne bleue), la fonction de résolution du modèle (ligne jaune) et la convolution  $\tilde{\phi}_{\text{ML}}^{(R)}(\omega)$  (ligne rouge) pour la résolution du spectromètre IN16B à l'ILL [54].

Un résultat important de notre première étude [59] est que l'EISF ajusté est nul. Cela peut s'expliquer par le fait que la MBP en solution est une molécule très flexible, de sorte que  $\langle \exp(i\mathbf{q} \cdot \hat{\mathbf{R}}_j) \rangle \approx 0$ . Dans l'approximation gaussienne [74, 75] ( $q \equiv |\mathbf{q}|$ ) de l'amplitude élastique, on peut écrire

$$EISF(\mathbf{q}) \stackrel{|\mathbf{q}| \rightarrow 0}{\sim} \frac{1}{N} \sum_{j \in \mathbb{H}} e^{-\frac{1}{3}|\mathbf{q}|^2 \langle \hat{\mathbf{u}}_j^2 \rangle} \approx 0, \quad (\text{VI.142})$$

où  $\langle \hat{\mathbf{u}}_j^2 \rangle \equiv \langle (\hat{\mathbf{R}}_j - \langle \hat{\mathbf{R}}_j \rangle)^2 \rangle$  est la fluctuation de position quadratique moyenne de l'atome (d'hydrogène)  $j$ . Pour des valeurs de  $q$  plus petites, la décroissance de l'EISF correspond aux grandes amplitudes de mouvement des atomes. En gardant à l'esprit que l'EISF est une "quantité théorique",  $EISF(\mathbf{q}) = F_{\text{inc}}(\mathbf{q}, \infty)$ , qui ne peut être déterminée qu'en supposant un modèle, nous pouvons maintenant vérifier les ajustements du modèle à trois paramètres (IV.73) avec un ajustement à deux paramètres, où l'EISF est éliminé conformément à l'équation. (VI.114).

Les résultats de la figure VI.4 montrent que  $\tau$  et  $\alpha$  ne changent que légèrement en comparant les ajustements à deux et trois paramètres et, en particulier, que la disparition de l'EISF à partir des ajustements à trois paramètres est confirmée. Cela signifie que l'intensité élastique mesurée (points verts) est entièrement due à la "contribution pseudoélastique" définie par l'équation. (VI.134). Pour son calcul, nous avons utilisé les approximations de Padé décrit dans la section 1.4. Nous notons ici que l'estimation standard des barres d'erreurs de paramètres pour les ajustements à trois paramètres (points jaunes) conduit à des barres d'erreur qui sont parfois à peine visibles sur les graphes et ne peuvent pas être réalisées pour les ajustements à deux paramètres. La raison en est purement technique, à savoir que dans ce cas, le "paramètre" EISF est une fonction compilée qui est transmise comme argument à l'ajustement du logiciel Mathematica [76].

### 3. Résultats de l'analyse de données QENS de la myoglobine

Pour montrer que le modèle (IV.73) ne conduit pas systématiquement à des EISFs proche de zéro pour les protéines en solution, nous avons analysé les données QENS de l'apo-myoglobine à  $T = 284$  K dans le solvant  $D_2O$  (pD = 6), qui ont été collectées sur le spectromètre IN5 à l'Institut Laue-Langevin à Grenoble, avec une résolution (FWHM) de  $11.6 \mu eV$  [43]. Contrairement à la MBP, la myoglobine est une protéine globulaire pliée de manière compacte, d'un poids à peu près identique, mais avec une structure tridimensionnelle bien définie contenant huit hélices  $\alpha$  en tant qu'éléments de structure secondaire.

La figure VI.5, la partie en bas, montre que l'EISF est ici clairement non nul et que les ajustements à trois et deux paramètres donnent à nouveau des

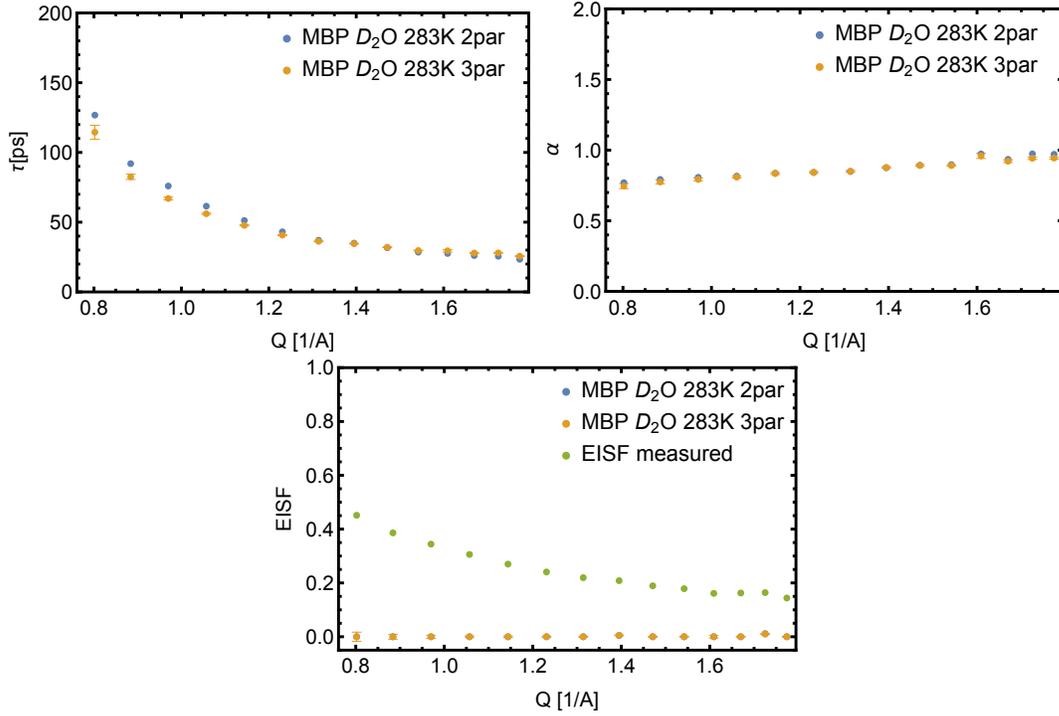


FIGURE VI.4. Les paramètres d'ajustement  $\tau$ ,  $\alpha$  et  $EISF$ , pour la MBP en solution. Les points bleus correspondent à un ajustement à deux paramètres, où  $EISF \equiv EISF(\tau, \alpha, \sigma)$  selon l'équation (VI.120) et les points jaunes à un ajustement à trois paramètres. En ce qui concerne l'EISF, les points verts correspondent à l'intensité élastique mesurée  $F_m(\infty) \equiv F_m(\mathbf{q}, \infty)$  définie dans l'équation (VI.112) [54].

résultats similaires. Le fait que les mouvements atomiques dans une protéine globulaire, repliée d'une manière compacte, soient plus gênés que dans une protéine intrinsèquement désordonnée de type polymère comme la MBP se reflète donc clairement dans les EISF correspondantes, et il convient de noter que cet effet est beaucoup plus prononcé pour les intensités élastiques corrigées en fonction de la résolution que pour les intensités élastiques mesurées. Il convient également de noter que l'échelle de temps caractéristique  $\tau$ , et le paramètre de forme  $\alpha$ , sont systématiquement plus petits pour la myoglobine,

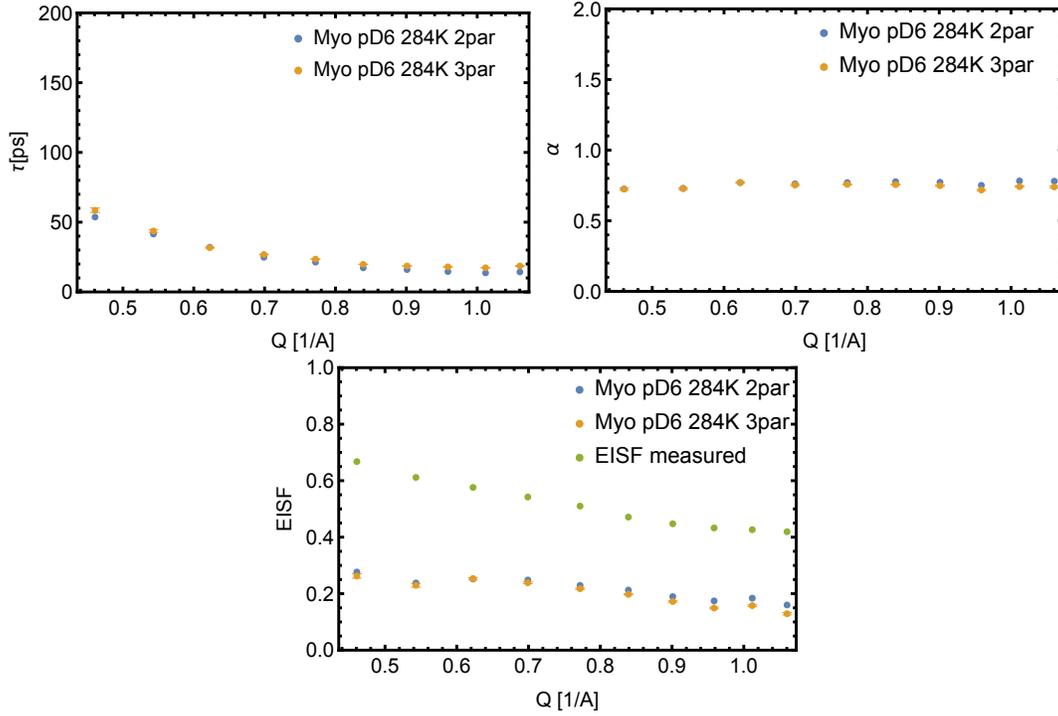


FIGURE VI.5. Les paramètres d’ajustement  $\tau$ ,  $\alpha$  et  $EISF$ , pour la myoglobine en solution [54].

ce qui reflète une dynamique plus gênée des atomes dans cette protéine pliée de manière plus compacte.

La figure VI.6 montre qu’ici, comme pour le MBP, la diffusion globale peut être négligée pour l’analyse des spectres QENS. Nous insérons ici un coefficient de diffusion global de  $D = 10.1 \text{ \AA}^2/\text{ns}$ , qui est la moyenne des valeurs rapportées dans les références [77] et [78] respectivement. Nous notons dans ce contexte qu’une valeur un peu plus faible de  $D \approx 8 \text{ \AA}^2/\text{ns}$  est obtenue par une estimation très approximative à partir de la relation de Stokes-Einstein,  $D = \frac{k_B T}{6\pi\eta R_h}$ , en servant ici un rayon hydrodynamique de  $R_h = 2 \text{ nm}$  qui

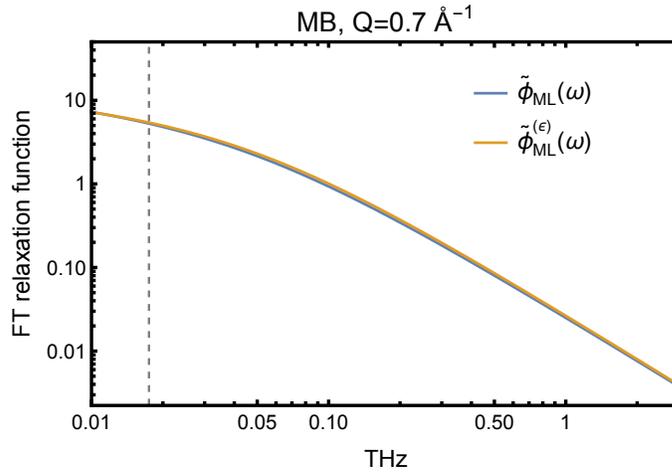


FIGURE VI.6. Impact de la diffusion globale sur le facteur de structure dynamique du modèle à  $q = 0.7/\text{Å}$  pour la myoglobine en solution [54]. Les légendes sont les mêmes que dans la figure. VI.2.

est calculé à partir de la structure PDB "1BVC" pour une molécule de l'apomyoglobine et pour  $\eta$  la viscosité de l'eau à 284 K ( $\eta = 1.306 \text{ mPa s}$ ) [79]. Dans ce contexte, l'impact de la diffusion globale peut être négligé.

**3.1. Correction de la résolution de  $\langle \hat{u}^2 \rangle$ .** L'approximation gaussienne (VI.142) de l'EISF est également connue sous le nom de facteur de Debye-Waller et a été utilisée pendant des décennies pour analyser les "balayages élastiques" de la diffusion incohérente de neutrons à partir des protéines en poudres  $\text{D}_2\text{O}$ -hydratées dans la faible gamme de  $q$ , afin de déduire de ces données les fluctuations moyennes de la position quadratique des atomes d'hydrogène dans la protéine. Nous citons quelques références sur ce sujet [52,80], qui pourraient être considérées comme "historiques" dans ce contexte. Il découle des équations (VI.117) et Eq. (VI.142) que la "vraie" variation de position quadratique moyenne corrigée en fonction de la résolution, moyennée sur tous les

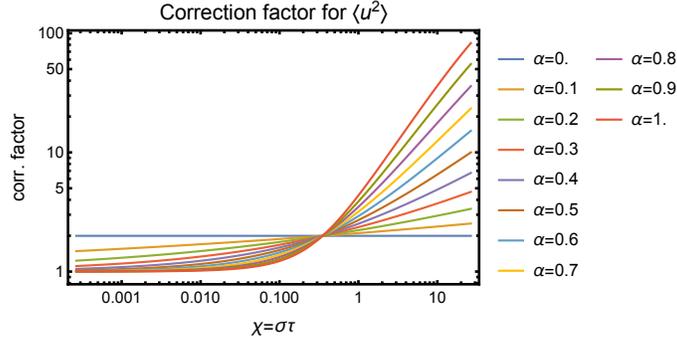


FIGURE VI.7. Facteur de correction pour convertir les fluctuations de position quadratique moyenne mesurées en fluctuations corrigées de la résolution [54].

atomes, est donnée par la formule suivante

$$\overline{\langle \hat{u}^2 \rangle} \stackrel{q \rightarrow 0}{\approx} -\frac{3}{q^2} \log \left( \frac{EISF_m(q) - \xi(q)}{1 - \xi(q)} \right). \quad (\text{VI.143})$$

Nous supposons ici un échantillon isotrope et indiquons explicitement la dépendance de toutes les quantités impliquées à  $q \equiv |\mathbf{q}|$ . En écrivant  $EISF_m(q) = \exp(-q^2 \overline{\langle \hat{u}^2 \rangle}_m / 3)$  et en développant l'expression (VI.143) en une série de puissances en  $\xi(q)$ , nous obtenons

$$\overline{\langle \hat{u}^2 \rangle} \approx \overline{\langle \hat{u}^2 \rangle}_m + \sum_{n=1}^{\infty} \frac{3}{nq^2} \left( e^{\frac{n}{3} q^2 \overline{\langle \hat{u}^2 \rangle}_m} - 1 \right) \xi(q)^n \stackrel{q \rightarrow 0}{\approx} \overline{\langle \hat{u}^2 \rangle}_m \sum_{n=0}^{\infty} \xi(q)^n, \quad (\text{VI.144})$$

où la série géométrique peut être additionnée pour donner

$$\overline{\langle \hat{u}^2 \rangle} \stackrel{q \rightarrow 0}{\approx} \frac{\overline{\langle \hat{u}^2 \rangle}_m}{1 - \xi(q)}. \quad (\text{VI.145})$$

La figure VI.7 montre le facteur de correction  $1/(1 - \xi)$  qui doit être appliqué pour obtenir les "vraies" fluctuations de position quadratique moyenne corrigées de la résolution à partir des fluctuations mesurées. A première vue, les facteurs de correction sont importants pour les "mauvaises résolutions" et la relaxation proche de l'exponentielle peuvent surprendre, mais ils peuvent être

compris en notant que dans ce cas, un balayage élastique contient pratiquement tout l'intégral sur la ligne quasi-élastique.

Dans cette étude, nous avons présenté une méthode quasi-analytique pour calculer la contribution pseudo-élastique de la diffusion quasi-élastique à l'amplitude de diffusion élastique des neutrons, ce qui permet d'estimer la véritable amplitude de diffusion élastique pour le modèle de relaxation de Mittag-Leffler. Grâce à l'efficacité de calcul de cette méthode qui est basée sur les approximations de Padé, la véritable amplitude de diffusion élastique peut être éliminée dans le modèle IV.73 en utilisant l'intensité élastique mesurée et les deux paramètres  $\alpha$  et  $\tau$  de la fonction de relaxation ML ainsi que la résolution de l'instrument  $\sigma$ . Nous avons appliqué la méthode pour confirmer les résultats de notre première étude [59] qui a révélé en particulier un EISF nul. Pour montrer que notre modèle ne conduit pas systématiquement à des EISFs proches de zéro pour les protéines en solution, nous avons réalisé une étude comparative pour la myoglobine en solution qui a à peu près le même poids que la MBP, mais contrairement à cette dernière, il s'agit d'une protéine globulaire avec une structure bien définie. Comme on pouvait s'y attendre, nous trouvons ici un EISF clairement non nul, ce qui montre que les mouvements atomiques sont plus gênés dans une protéine repliée de manière compacte. Un résultat important dans ce contexte est que l'EISF corrigé par la résolution montre ce résultat beaucoup plus clairement que celui mesuré.



## Dynamique d'une enzyme en solution

La compréhension de la dynamique fonctionnelle des enzymes est une question fondamentale en biophysique moléculaire, en biologie et en biochimie. La phosphoglycérate kinase (PGK) est un exemple pour lequel la relation dynamique-fonction a été étudiée de manière intensive à l'aide de diverses méthodes, notamment la RMN structurale, la cristallographie aux rayons X, la diffusion quasi-élastique des neutrons (QENS), la spectroscopie d'écho de spin des neutrons (NSE) et la simulation de dynamique moléculaire (MD) [81–90]. L'idée est de présenter une analyse des spectres QENS à haute résolution de la phosphoglycérate kinase qui permette d'élucider l'activité enzymatique sur la dynamique de la protéine. Plusieurs études ont été réalisées dans le but de mieux comprendre le rôle des mouvements inter-domaines dans la fonction de l'enzyme [84, 86, 87, 89]. Une méthode puissante résolue dans l'espace et dans le temps a été utilisée qui est la spectroscopie d'écho de spin de neutrons (NSE) [87] en combinaison avec la diffusion de neutrons aux petits angles (SANS) et l'analyse en mode normaux, et dans la [89] en combinaison avec la simulation MD. La spectroscopie d'écho de spin (NSE) sonde les mouvements lents et la diffusion globale des protéines sur une échelle de temps de 0.1-100 ns et sur une échelle de longueur de l'ordre du nm. Les résultats des études NSE suggèrent que les mouvements de la charnière des deux domaines de la PGK permettent son activité et que la présence des substrats rigidifie la molécule et accélère sa dynamique interne. Et la présente étude vise à étendre

et à consolider les travaux susmentionnés par une analyse de données QENS du spectromètre à haute résolution IN16B de l'Institut Laue-Langevin de Grenoble [40]. L'instrument sonde l'échelle de temps ns s'il est utilisé en mode BATS et comble l'écart entre les expériences QENS avec des spectromètres de temps de vol et la spectroscopie NSE.

### 1. Modèle utilisé pour l'analyse de données QENS de la PGK

L'analyse de données QENS a été effectuée dans le domaine du temps, en utilisant un modèle stochastique pour la fonction intermédiaire de diffusion classique et en supposant que l'approximation semi-classique de Schofield [59] est valable. Pour le modèle utilisé, nous avons supposé qu'il existe un atome d'hydrogène représentatif "a" dont la dynamique tient compte à la fois de la dynamique de relaxation des atomes d'hydrogène individuels et de leur hétérogénéité de mouvement. Dans l'approximation semi-classique de Schofield, nous avons alors

$$F^{(+)}(\mathbf{q}, t) \approx \left\langle e^{-i\mathbf{q} \cdot \hat{\mathbf{x}}_a(0)} e^{i\mathbf{q} \cdot \hat{\mathbf{x}}_a(t)} \right\rangle_{\text{cl}}, \quad (\text{VII.146})$$

où  $\langle \dots \rangle_{\text{cl}}$  représente une moyenne d'ensemble classique. Nous supposons en outre que les domaines de PGK peuvent être traités comme équivalents et que les mouvements de l'atome diffuseur ne sont pas corrélés avec les mouvements du domaine auquel il est attaché. En écrivant

$$\mathbf{x}_a = \mathbf{R}_a + \mathbf{r}_a, \quad (\text{VII.147})$$

où  $\mathbf{R}_a$  indique le centre du domaine et  $\mathbf{r}_a$  est la position de l'atome diffuseur "représentatif" par rapport à ce point de référence (voir la figure Fig. VII.1), on

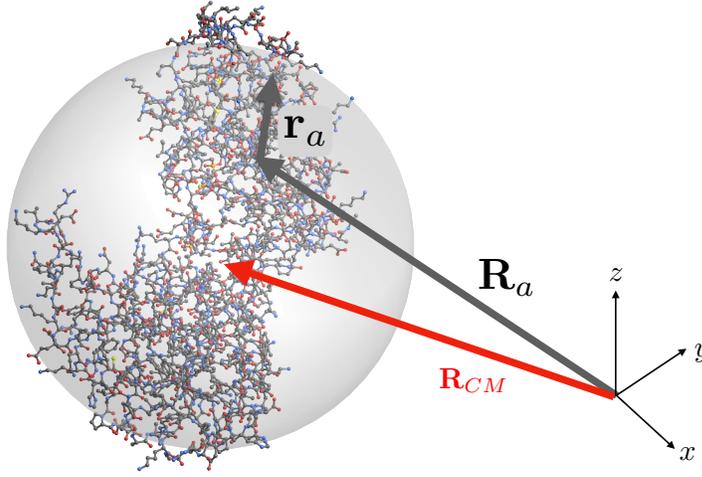


FIGURE VII.1. La molécule PGK (code PDB 3PGK) avec un rayon de sphère  $R_H = 30.5 \text{ \AA}$  qui est utilisé pour la relation de Stokes-Einstein (VII.158) [91]. Ici  $\mathbf{R}_{CM}$  est le centre de masse de la molécule entière et  $\mathbf{R}_a$  et  $\mathbf{r}_a$  sont définis dans le texte.

défini ici et les deux variables dynamiques correspondantes

$$\xi(t) = A_{\mathbf{q}}(t) - \langle A_{\mathbf{q}}(t) \rangle_{cl}, \quad \text{où } A_{\mathbf{q}}(t) = e^{i\mathbf{q} \cdot \mathbf{R}_a(t)}, \quad (\text{VII.148})$$

$$\chi(t) = a_{\mathbf{q}}(t) - \langle a_{\mathbf{q}}(t) \rangle_{cl}, \quad \text{où } a_{\mathbf{q}}(t) = e^{i\mathbf{q} \cdot \mathbf{r}_a(t)}. \quad (\text{VII.149})$$

Si l'on suppose que les mouvements des domaines entiers ne sont pas corrélés avec les mouvements intra-domaines la fonction intermédiaire de diffusion moyennée sur l'orientation de la PGK en solution peut alors être factorisée comme suit

$$\overline{F}^{(+)}(q, t) \approx f(q, t)g(q, t) \quad (q \equiv |\mathbf{q}|), \quad (\text{VII.150})$$

où

$$f(q, t) \equiv \overline{\langle a_{\mathbf{q}}^* a_{\mathbf{q}}(t) \rangle}_{cl} \quad \text{et} \quad g(q, t) \equiv \overline{\langle A_{\mathbf{q}}^* A_{\mathbf{q}}(t) \rangle}_{cl}. \quad (\text{VII.151})$$

Les deux variables  $\xi(t)$  et  $\chi(t)$  doivent être considérés comme des variables stochastiques complexes, dépendant des vecteurs de position respectifs et de

$\mathbf{q}$ . On exige que  $\langle A_{\mathbf{q}} \rangle \neq 0$  bien que cette variable décrive un mouvement global. Concernant la dynamique intra-domaine, on a utilisé le même modèle que pour la dynamique interne utilisé dans notre première étude qui concerne la protéine basique de la myéline [59] où on a supposé que la dynamique de  $\chi$  peut être modélisée comme un processus d'Ornstein-Uhlenbeck fractionnaire (fOU) [71, 72],

$$f(q, t) \equiv EISF(q) + (1 - EISF(q))E_{\alpha}(-(|t|/\tau)^{\alpha}) \quad (\text{VII.152})$$

où  $EISF(q)$ ,  $\tau \equiv \tau(q)$ , et  $\alpha \equiv \alpha(q)$  sont des ajustements qui dépendent de  $q$ .

Passant maintenant aux mouvements inter-domaines dans la PGK où on a utilisé un modèle raffiné comparé à celui que nous avons utilisé dans notre étude de la protéine basique de la myéline [59]. On définit

$$\boldsymbol{\xi} \equiv \begin{pmatrix} \Re\{A_{\mathbf{q}} - \langle A_{\mathbf{q}} \rangle_{\text{cl}}\} \\ \Im\{A_{\mathbf{q}} - \langle A_{\mathbf{q}} \rangle_{\text{cl}}\} \end{pmatrix}, \quad (\text{VII.153})$$

On considère au début que la probabilité conditionnelle,  $P(\boldsymbol{\xi}, t | \boldsymbol{\xi}_0, 0)$  pour une transition  $\boldsymbol{\xi}_0 \rightarrow \boldsymbol{\xi}$  dans le temps  $t$  décrit un processus d'Ornstein-Uhlenbeck (OU) [92], c'est-à-dire la diffusion dans un potentiel harmonique,

$$\partial_t P = \left\{ \eta_{\boldsymbol{\xi}} \frac{\partial}{\partial \boldsymbol{\xi}} \cdot \{ \boldsymbol{\xi} P \} + D_{\boldsymbol{\xi}} \frac{\partial}{\partial \boldsymbol{\xi}} \cdot \frac{\partial P}{\partial \boldsymbol{\xi}} \right\} \quad (\text{VII.154})$$

ici  $D_{\boldsymbol{\xi}}$  est un coefficient de diffusion et  $\eta_{\boldsymbol{\xi}}$  est une constante de relaxation. Le point essentiel est que la fonction d'autocorrélation de  $\boldsymbol{\xi}$  décroît de façon exponentielle,

$$\langle \boldsymbol{\xi}^T(0) \cdot \boldsymbol{\xi}(t) \rangle_{\text{cl}} = \langle |\boldsymbol{\xi}(0)|^2 \rangle_{\text{cl}} e^{-\eta_{\boldsymbol{\xi}} t}, \quad \eta_{\boldsymbol{\xi}} \equiv \eta_{\boldsymbol{\xi}}(\mathbf{q}). \quad (\text{VII.155})$$

puisque  $\langle A_{\mathbf{q}} \rangle \approx 0$  pour une particule qui diffuse librement, et on peut écrire

$$g(q, t) = e^{-D(q)q^2|t|}, \quad (\text{VII.156})$$

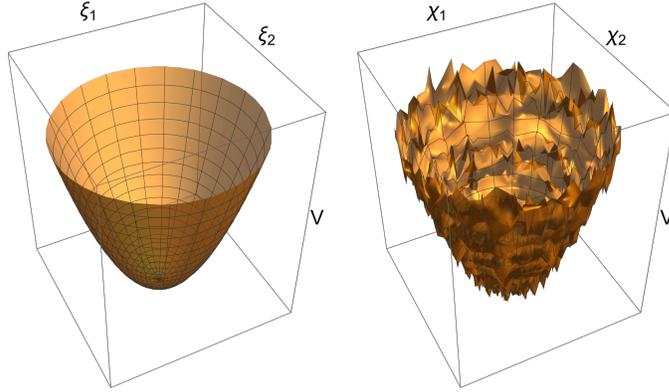


FIGURE VII.2. Potentiel harmonique lisse (partie gauche) et son équivalent rugueux (partie droite) [91].

où le paramètre d'ajustement  $D(q)$  a la dimension d'un coefficient de diffusion ( $\text{m}^2/\text{s}$ ). Pour les petites valeurs de  $q$ , il résulte du développement cumulatif de la fonction intermédiaire de diffusion que

$$g(q, t) \stackrel{q \rightarrow 0}{\equiv} e^{-\frac{q^2}{6} \langle (\mathbf{R}_a(t) - \mathbf{R}_a(0))^2 \rangle_d} \stackrel{t \rightarrow \infty}{\sim} e^{-D_0 q^2 t}. \quad (\text{VII.157})$$

où  $D_0$  est le coefficient de diffusion pour un domaine entier et donc pour la molécule PGK entière. Pour des valeurs de  $q$  finies,  $g(q, t)$  décrit la diffusion de molécules PGK entières qui est modulée par les mouvements inter-domaines.

## 2. Résultats de l'analyse de données QENS de la PGK

Nous commençons la discussion des résultats par la figure VII.3, qui montre un ajustement de la fonction intermédiaire de diffusion déconvoluée par résolution V.101 avec quatre paramètres,  $\tau(q)$ ,  $\alpha(q)$ ,  $EISF(q)$ , et  $D(q)$  pour les valeurs minimales et maximales de  $q$  en absence et en présence de substrats. La figure VII.4 montre les paramètres  $\tau$  et  $\alpha$  en fonction de  $q$  en présence et en l'absence des substrats :  $\tau$  et  $\alpha$  sont systématiquement réduits en

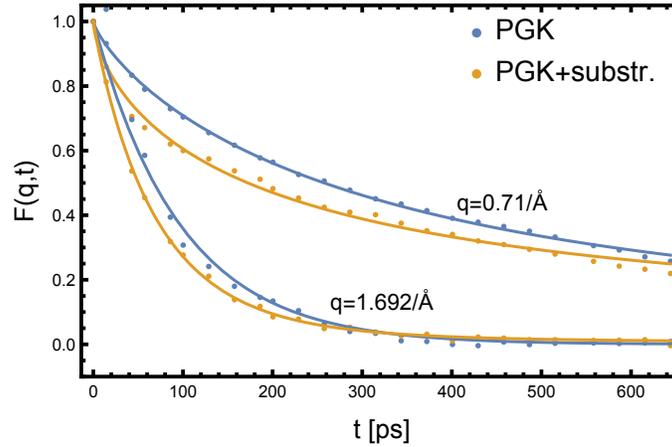


FIGURE VII.3.  $F(q, t)$  déconvoluées de la résolution instrumentale pour la PGK en solution deutérée sans et en présence de substrats (points bleus et jaunes, respectivement) et les ajustements correspondants (ligne bleue et jaune, respectivement) [91].

présence des substrats, ce qui indique que la dynamique interne est accélérée par l'activité enzymatique de la molécule et que la dynamique de relaxation des domaines devient moins exponentielle. Nous remarquons que  $\tau$  et  $\alpha$  en présence et en absence des substrats suivent globalement la même évolution avec  $q$ . Le paramètre d'échelle de temps  $\tau$  devient généralement plus petit avec l'augmentation de  $q$ , ce qui indique simplement que les mouvements localisés sont plus rapides que les mouvements collectifs impliquant un grand nombre d'atomes. Le paramètre de forme  $\alpha$  augmente plutôt avec  $q$  jusqu'à des valeurs proches de 1, ce qui indique une relaxation de plus en plus exponentielle pour des mouvements plus localisés. Cela est peut-être dû à ce comportement qui est au fait que moins de modes de relaxation contribuent aux mouvements localisés qu'aux mouvements de grande amplitude qui sont sondés à de petites valeurs de  $q$ . Les EISFs ajustés ainsi que les contreparties mesurés qui sont obtenus en intégrant l'intensité QENS mesurée sur la largeur de la fonction de résolution sont présentés dans la figure VII.4 (partie en bas). Nous constatons

que l'EISF ajusté est globalement proche de zéro en présence et en absence de ligand, sauf à  $q = 1.3 \text{ \AA}^{-1}$  où l'EISF de PGK en absence de ligand et de substrats est légèrement supérieure à l'EISF en leur présence. La corrélation de cette observation avec la diminution de  $\alpha$  en présence des substrats montre que les domaines sont légèrement rigidifiés, ce qui confirme à nouveau les résultats de la référence [87] qui ont été obtenus par la spectroscopie NSE. La différence entre les intensités élastiques mesurées et ajustées a été observée précédemment et peut être attribuée aux contributions parasites de la diffusion quasi-élastique à l'intensité élastique élargie par l'instrument [59]. À titre de comparaison, nous montrons également les ajustements avec un modèle réduit où  $D(q) \equiv 0$  (carrés bleus et jaunes, respectivement). On peut clairement voir que les résultats sont très similaires, la différence étant les barres d'erreur, qui sont beaucoup plus grandes pour l'ajustement à quatre paramètres et qui ne sont pas montrées ici. Cette observation est en accord avec les résultats de notre première étude pour la protéine basique de la myéline (MBP) intrinsèquement désordonnée [59] et nous présentons les ajustements à trois paramètres pour montrer que les ajustements de  $\tau$ ,  $\alpha$  et  $EISF$  sont stables.

L'impact de la dynamique enzymatique sur le paysage énergétique intradomaine peut être visualisé en comparant les profils des barrières énergétiques  $P_{ML}(\epsilon)$ , décrivant sa "rugosité" [58, 59, 70] qui sont présentés dans la figure VII.5. une différence importante entre les deux profils est observée pour les valeurs de  $q$  correspondant aux amplitudes d'ouverture de la région charnière et indiquent une distribution plus large des barrières énergétiques en présence de substrats. Pour mieux expliquer cette observation, on voit que le paramètre  $\alpha$  pour la PGK avec et sans ligand s'approche de 1, mais celui de la PGK avec le ligand reste encore un peu plus petit, et cette petite différence est

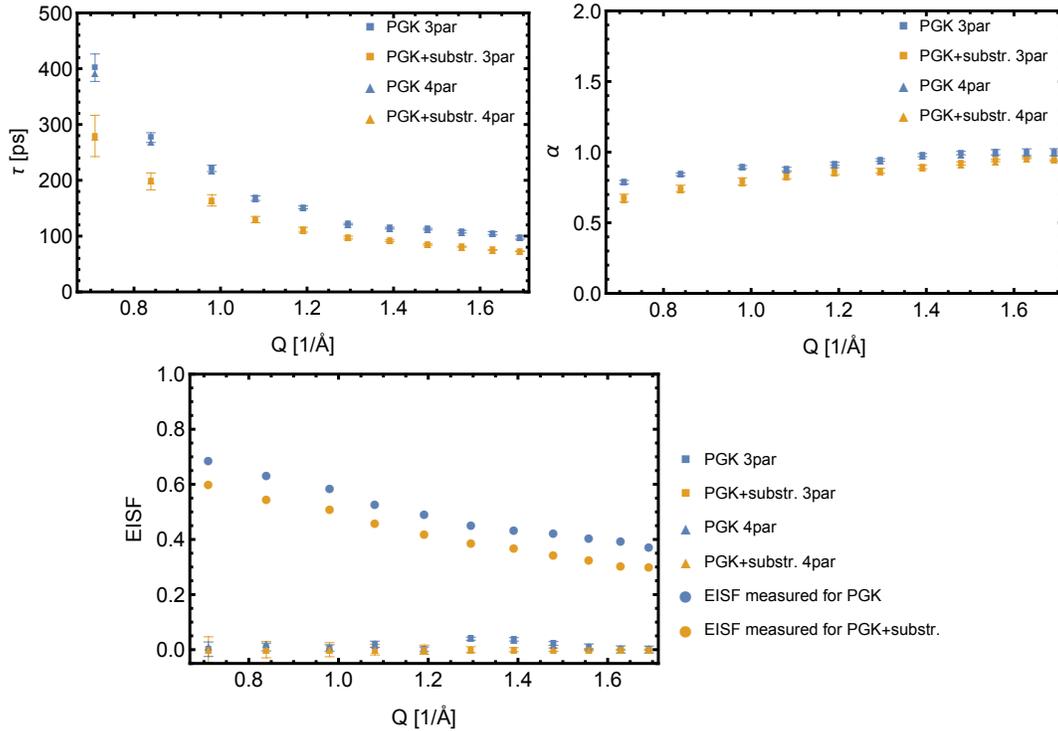


FIGURE VII.4. Les paramètres du modèle  $\tau$ ,  $\alpha$  et  $EISF$  pour la PGK. Les triangles indiquent des ajustements à quatre paramètres (sans barres d'erreur) et les carrés des ajustements à trois paramètres (avec barres d'erreur) [91].

suffisante pour rendre la distribution plus large. Sachant que la largeur de la distribution  $P(\epsilon)$  dépend très fortement d' $\alpha$  et augmente très rapidement lorsque  $\alpha$  diminue. Il y a une grande différence entre  $\alpha=1$  et  $\alpha=0.95$  (voir la figure IV.3).

La figure VII.6 montre le coefficient de diffusion de la PGK avec et sans ligand. On observe que  $D(q)$  présente une modulation prononcée par rapport à ses valeurs pour les petites et les grandes valeurs de  $q$ . Ces valeurs sont similaires et proches de l'estimation du coefficient de diffusion d'une molécule

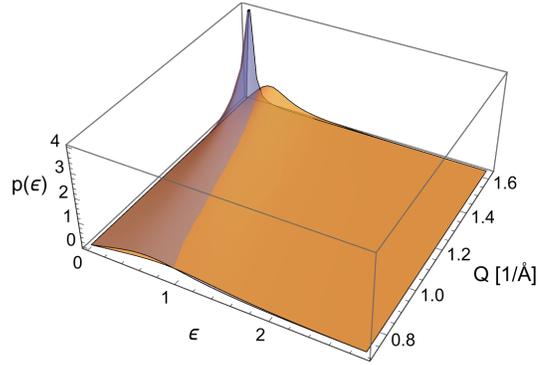


FIGURE VII.5. Spectre de la barrière d'énergie en absence (bleu) et en présence (jaune) de substrats [91].

PGK entière obtenue à partir de la loi de Stokes-Einstein

$$D_0 = \frac{k_B T}{6\pi\eta R_H} \approx 5.1 \times 10^{-3} \text{Å}^2/\text{ps}. \quad (\text{VII.158})$$

À partir de la structure PDB **3PGK**, nous avons utilisé un rayon hydrodynamique effectif de  $R_H = 30.5 \text{Å}$  pour estimer le coefficient de diffusion. Le maximum de  $D(q)$  à environ  $q_{\max} = 1.2/\text{Å}$  correspond à  $2\pi/q_{\max} = 5.2 \text{Å}$  dans l'espace réel, ce qui peut être associé aux mouvements de la région charnière de la PGK causés par son activité enzymatique et qui ont également été observés en combinant la spectroscopie NSE et l'analyse des modes normaux [87].

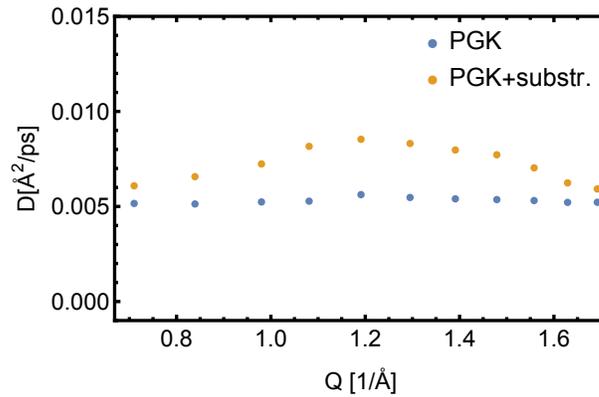


FIGURE VII.6. Le coefficient de diffusion ajusté  $D(q)$  [91].

Dans cette étude, nous avons présenté une analyse des spectres de diffusion quasi-élastique de neutrons à haute résolution de la PGK avec et sans ligand pour une température de  $T = 283 \text{ K}$ . Cette étude donne une image cohérente de la dynamique fonctionnelle de la PGK qui confirme et complète une étude antérieure par spectroscopie NSE [87]. Elle démontre en particulier que la dynamique de relaxation fortement non exponentielle des protéines doit être prise en compte dans les modèles utilisés pour l'analyse de données expérimentales afin d'exploiter pleinement leur riche contenu en informations.

## Chapitre 8

### **Conclusion**

Le travail de thèse concerne essentiellement l'analyse des spectres QENS provenant des protéines en solution qui ont été précédemment mesurées sur deux spectromètres différents IN16B et IN5 à l'Institut Laue-Langevin, à Grenoble. L'objectif du travail présent était d'étudier la dynamique interne et globale des différentes protéines en solution, commençant par une protéine intrinsèquement désordonnée (la protéine basique de la myéline) ensuite une protéine qui a environ la même taille que la MBP, mais qui est une protéine globulaire ayant une structure bien définie (Myoglobine), et enfin une enzyme qui contient deux domaines globulaires (Phosphoglycérate kinase). Le but de ce présent travail était de tester et développer des modèles "minimalistes" multi-échelles pour comprendre la dynamique des protéines en solution.

Le manuscrit de la thèse est globalement structuré autour de sept chapitres principaux. Commenant par la partie théorique de la diffusion de neutrons, qui comporte tous les aspects théoriques de la diffusion quasiélastique de neutrons utilisés dans cette étude. Ensuite dans le deuxième chapitre, les systèmes étudiés ont été présentés, ainsi que les spectromètres utilisés. Le troisième chapitre était consacré à l'analyse de données qui a été effectué dans le domaine du temps. Les études effectuées dans cette thèse et le modèle utilisé ont été détaillé dans le troisième, quatrième et cinquième chapitre, respectivement.

La première étude était basée sur la dynamique de relaxation multi-échelles d'une protéine intrinsèquement désordonnée dont la structure n'est pas ou que partiellement connue. Les IDPs présentent de grandes similitudes avec les polymères classiques. Parmi ces protéines intrinsèquement désordonnées, notre choix était basé sur la protéine basique de la myéline (MBP) qui est un composant essentiel de la gaine de myéline qui se forme autour des axones dans le cerveau humain. Alors que la MBP est bien repliée lorsqu'elle est en contact avec les membranes de myéline chargées négativement, elle est en grande partie désordonnée en solution. Ce travail a illustré le potentiel des modèles "minimalistes" pour dévoiler les changements dynamiques d'une protéine dus à une modification des conditions externes. Ces changements se produisent ici comme une conséquence de l'ajout de 30% trifluoroéthanol (TFE) à la solution de D<sub>2</sub>O. Le TFE est connu pour sa propriété d'induire la formation de contenu hélicoïdal  $\alpha$  dans la région dépliée d'une protéine, et la question était ici de savoir si la MBP se transformerait en une protéine pliée globulaire en présence du TFE. Dans ce contexte, des mesures QENS ont été effectuées sur le spectromètre IN16B de l'Institut Laue-Langevin à Grenoble en deux solvants, D<sub>2</sub>O et en D<sub>2</sub>O avec 30% de Trifluoroéthanol (TFE) deutéré, et à trois températures, 283 K, 303 K et 323 K et ont été analysées avec le modèle générique à trois paramètres et des expériences complémentaires de diffusion dynamique de la lumière et les expériences SRCD ont été réalisées pour estimer respectivement le coefficient de diffusion global de la MBP et le contenu en structure secondaire de la molécule pour les deux solvants en fonction de la température. Nous avons clairement pu montrer que l'ajout de 30% de TFE rigidifie légèrement les molécules de la MBP, mais n'induit pas de changement profond dans la structure des molécules vers celle d'une protéine globulaire, et que les modifications du comportement dynamique dues au TFE disparaissent

rapidement avec l'augmentation de la température. Notre analyse a notamment montré que  $EISF(q) \approx 0$ , ce qui est compatible avec la grande flexibilité interne de cette protéine semblable à un polymère, mais n'est compatible avec aucun type de modèle élastique pour la MBP en solution [93], et n'est non plus compatible avec les spectres QENS des échantillons de poudre, qui devraient gêner non seulement les mouvements internes, mais aussi les mouvements globaux de la protéine entière. Nous notons dans ce contexte que les valeurs de EISF ne sont pas systématiquement petites pour ce modèle et nous nous référons à des travaux récents sur une analyse de données QENS à partir de poudres hydratées d'acétylcholinestérase humaine [58] qui est une enzyme avec une structure bien définie. Ici, l'EISF s'est avéré clairement non nul, comme on peut s'y attendre pour une telle protéine qui est plus compacte et un tel échantillon, où les mouvements globaux de la protéine sont empêchés. Pour confirmer ce résultat, une nouvelle étude a été réalisée [54] afin de mettre au point une procédure d'ajustement dans laquelle cette quantité est éliminée en tant que paramètre d'ajustement. L'idée était de remplacer l'intensité élastique dans la procédure d'ajustement par une estimation qui utilise l'intensité élastique mesurée comme entrée, ainsi que la contribution quasiélastique indésirable due à la résolution finie du spectromètre considéré qui est décrite par les paramètres  $\alpha$  et  $\tau$ . Nous avons en particulier comparé le modèle à deux paramètres au modèle utilisé dans notre première étude (trois paramètres) [59], et les résultats confirment la disparition de l'EISF pour cette protéine (MBP). Pour montrer que ce modèle ne conduit pas systématiquement à la disparition de l'EISF pour les protéines en solution, une analyse des données QENS de la myoglobine a été réalisée. Les données QENS pour cette étude provenaient du spectromètre IN5 de l'Institut Laue-Langevin à Grenoble. Contrairement à la MBP, la myoglobine est une protéine globulaire pliée de façon compacte

mais de poids approximativement identique. Nous avons constaté que l'EISF est clairement non nul pour la myoglobine, ce qui peut se comprendre puisque dans une protéine globulaire compacte, les mouvements atomiques sont plus gênés que dans une protéine intrinsèquement désordonnée comme la MBP.

La troisième étude était basée sur une analyse des spectres de diffusion quasi-élastique de neutrons à haute résolution de la Phosphoglycérate kinase PGK qui élucide l'influence de l'activité enzymatique sur la dynamique de la protéine. Comme notre première étude sur la protéine basique de la myéline [59], les spectres QENS de cette enzyme ont été collectés sur le spectromètre à haute résolution IN16B de l'Institut Laue-Langevin à Grenoble pour une température  $T = 283\text{ K}$  en présence et en absence des substrats. La PGK est une enzyme clé de la glycolyse et catalyse la conversion réversible du 1,3-bisphosphoglycérate (1,3-BPG) en 3-phosphoglycérate (3PG) lors de la synthèse de l'ATP à partir de l'ADP. L'enzyme est composée de deux domaines structurellement homologues qui sont reliés par une région charnière bien conservée. Les sites de liaison du substrat de la PGK se trouvent à l'interface entre les domaines, et des mouvements inter-domaines de grande amplitude sont nécessaires pour l'activité catalytique. L'idée était de voir l'influence de la liaison de l'ATP sur la dynamique interne de la PGK vue par diffusion quasi-élastique de neutrons. Nous constatons que la présence de substrats accélère la dynamique de relaxation de la phosphoglycérate kinase et renforce ses caractéristiques non exponentielles. En utilisant une image de paysage énergétique de la dynamique des protéines, cette observation peut être traduite par une distribution plus large des barrières énergétiques et une rigidification globale de la molécule. Cette étude montre en particulier que le modèle "minimaliste"

utilisé dans ce travail suffit à extraire l'information assez riche contenue des données QENS et d'obtenir une interprétation physique des résultats.



## Chapitre 9

### **Publications**

Dans ce chapitre les trois publications sont citées.

#### **1. Multiscale relaxation dynamics and diffusion of myelin basic protein in solution studied by quasielastic neutron scattering**

Cet article présente une analyse des spectres de diffusion quasi-élastique de neutrons à haute résolution de la protéine basique de la myéline (MBP) en solution, en comparant les spectres à trois températures différentes (283, 303 et 323 K) pour le solvant D<sub>2</sub>O pur et un mélange de D<sub>2</sub>O avec 30% de trifluoroéthanol (TFE) deutéré. Il a été publié dans le "Journal of Chemical Physics".

# Multiscale relaxation dynamics and diffusion of myelin basic protein in solution studied by quasielastic neutron scattering

Cite as: J. Chem. Phys. **156**, 025102 (2022); <https://doi.org/10.1063/5.0077100>

Submitted: 30 October 2021 • Accepted: 19 December 2021 • Accepted Manuscript Online: 22 December 2021 • Published Online: 11 January 2022

 Abir N. Hassani,  Luman Haris,  Markus Appel, et al.



View Online



Export Citation



CrossMark

## ARTICLES YOU MAY BE INTERESTED IN

[Accurate parameterization of the kinetic energy functional for calculations using exact-exchange](#)

The Journal of Chemical Physics **156**, 024107 (2022); <https://doi.org/10.1063/5.0065217>

[Construction of meta-GGA functionals through restoration of exact constraint adherence to regularized SCAN functionals](#)

The Journal of Chemical Physics (2021); <https://doi.org/10.1063/5.0073623>

The Journal  
of Chemical Physics

**SPECIAL TOPIC:** Low-Dimensional  
Materials for Quantum Information Science

Submit Today!

AIP  
Publishing

# Multiscale relaxation dynamics and diffusion of myelin basic protein in solution studied by quasielastic neutron scattering

Cite as: J. Chem. Phys. 156, 025102 (2022); doi: 10.1063/5.0077100

Submitted: 30 October 2021 • Accepted: 19 December 2021 •

Published Online: 11 January 2022



View Online



Export Citation



CrossMark

Abir N. Hassani,<sup>1,2,3</sup>  Luman Haris,<sup>3,4</sup>  Markus Appel,<sup>5</sup>  Tilo Seydel,<sup>5</sup>  Andreas M. Stadler,<sup>3,4,a)</sup>   
and Gerald R. Kneller<sup>1,2,b)</sup> 

## AFFILIATIONS

<sup>1</sup>Centre de Biophysique Moléculaire, CNRS and Université d'Orléans, Rue Charles Sadron, 45071 Orléans, France

<sup>2</sup>Synchrotron Soleil, L'Orme de Merisiers, 91192 Gif-sur-Yvette, France

<sup>3</sup>Jülich Centre for Neutron Science (JCNS-1) and Institute of Biological Information Processing (IBI-8), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

<sup>4</sup>Institute of Physical Chemistry, RWTH Aachen University, Landoltweg 2, 52056 Aachen, Germany

<sup>5</sup>Institut Laue Langevin, 71 Avenue des Martyrs, 38042 Grenoble Cedex 9, France

<sup>a)</sup>E-mail: [a.stadler@fz-juelich.de](mailto:a.stadler@fz-juelich.de)

<sup>b)</sup>Author to whom correspondence should be addressed: [gerald.kneller@cnrs.fr](mailto:gerald.kneller@cnrs.fr)

## ABSTRACT

We report an analysis of high-resolution quasielastic neutron scattering spectra from Myelin Basic Protein (MBP) in solution, comparing the spectra at three different temperatures (283, 303, and 323 K) for a pure D<sub>2</sub>O buffer and a mixture of D<sub>2</sub>O buffer with 30% of deuterated trifluoroethanol (TFE). Accompanying experiments with dynamic light scattering and Circular Dichroism (CD) spectroscopy have been performed to obtain, respectively, the global diffusion constant and the secondary structure content of the molecule for both buffers as a function of temperature. Modeling the decay of the neutron intermediate scattering function by the Mittag-Leffler relaxation function,  $\phi(t) = E_{\alpha}(-(t/\tau)^{\alpha})$  ( $0 < \alpha < 1$ ), we find that trifluoroethanol slows down the relaxation dynamics of the protein at 283 K and leads to a broader relaxation rate spectrum. This effect vanishes with increasing temperature, and at 323 K, its relaxation dynamics is identical in both solvents. These results are coherent with the data from dynamic light scattering, which show that the hydrodynamic radius of MBP in TFE-enriched solutions does not depend on temperature and is only slightly smaller compared to the pure D<sub>2</sub>O buffer, except for 283 K, where it is much reduced. In accordance with these observations, the CD spectra reveal that TFE induces essentially a partial transition from  $\beta$ -strands to  $\alpha$ -helices, but only a weak increase in the total secondary structure content, leaving about 50% of the protein unfolded. The results show that MBP is for all temperatures and in both buffers an intrinsically disordered protein and that TFE essentially induces a reduction in its hydrodynamic radius and its relaxation dynamics at low temperatures.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0077100>

## I. INTRODUCTION

Since the 1990s, the literature has seen a considerable increase in studies targeting the so-called intrinsically disordered proteins (IDPs), which have a well-defined function but not a well-defined structure.<sup>1–4</sup> The discovery of IDPs marked a change of the long-standing paradigm in structural biology that a protein structure is a prerequisite for the protein function. A protein is usually considered intrinsically disordered if it contains at least one region of 40

consecutive amino acids that do not correspond to any particular secondary structure ( $\alpha$ -helix or  $\beta$ -strand). One of the prominent members of the family of IDPs is Myelin Basic Protein (MBP), which is the basic constituent of the myelin sheath of nerves and the second most abundant protein in the central nervous system. The protein plays an important role in understanding diseases of the human nervous system, such as multiple sclerosis. In its predominant form, MBP contains 170 residues and has a molecular weight of 18.5 kDa. In aqueous solution, the protein is intrinsically disordered, and the

present study focusses on the effect of trifluoroethanol (TFE)–water mixtures on its global and internal dynamics. TFE is used since many years as an agent, which is known to induce the formation of  $\alpha$ -helices in IDPs, but its action is still not completely understood. The formation of secondary structure elements does, for instance, not necessarily lead to a well-defined global native structure as for globular proteins.<sup>5</sup> The aim of the present quasielastic neutron scattering (QENS) study is to shed some light on the action of TFE on the global folding and the dynamics of MBP. Thermal neutron scattering is a unique technique to study the structural dynamics of condensed matter at the atomic scale, including soft matter and biological molecules. The energy of thermal neutrons matches by definition the thermal energies of atoms in condensed matter systems, and its wavelength is compatible with typical interatomic distances. High-resolution quasielastic neutron scattering (QENS), with typical resolutions in the  $\mu\text{eV}$  range, corresponding to a few nanoseconds, is, in particular, suitable to study the slow internal relaxation dynamics of complex macromolecular systems.<sup>6–8</sup> We mention here that QENS is complementary to neutron spin-echo spectroscopy, which probes macromolecular conformational dynamics on much longer time scales of up to several hundred nanoseconds and which has also been applied previously to study the dynamics of MBP.<sup>9,10</sup>

The internal dynamics of proteins and complex molecular systems, in general, is characterized by a strongly non-exponential relaxation of the relevant time correlation functions, which exhibit an asymptotic power law decay. The power law decay reflects self-similarity, i.e., forms invariance of the correlation functions under a change of the time scale, which is also typical for glasses.<sup>11–19</sup> The typical properties of internal protein dynamics described above must be taken into account in the model for QENS from MBP, and the challenge is that the dynamics of very slow relaxation modes overlaps with elastic scattering<sup>20</sup> and also with the diffusion of whole molecules, which must be taken into account since MBP is studied in solution. For this purpose, we extend a model for internal protein dynamics, which has been recently used to analyze QENS data from protein powder samples of human acetylcholinesterase.<sup>21,22</sup>

This paper is organized as follows: Sec. II describes the experimental details and is followed by a presentation of the model for the intermediate scattering function, a discussion of the results, and a conclusion, which are, respectively, presented in Secs. III–V.

## II. EXPERIMENTS

### A. QENS spectra

#### 1. Sample preparation

Bovine Myelin Basic Protein (MBP) and all used chemicals were bought from Sigma-Aldrich (St. Louis, MO, USA). MBP powder was dissolved in  $\text{D}_2\text{O}$  (99.9 at. % D) and incubated for one day to allow for exchange of labile protons. The MBP solution was then again freeze dried. The QENS experiments have been performed for  $\text{D}_2\text{O}$ -exchanged MBP in solution, using for the latter a  $\text{D}_2\text{O}$  buffer (20 mM  $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ , pH 4.8) and a corresponding  $\text{D}_2\text{O}$  buffer with 30 vol. % per volume of deuterated 2,2,2-trifluoroethanol- $\text{d}_3$  (TFE). The protein concentration was 55 mg/ml. The protein concentration was determined by optical absorption at 280 nm wavelength with an absorbance of  $E_{1\%} = 5.89$ .<sup>23</sup>

### 2. Instrument

All QENS spectra have been collected on the IN16B spectrometer operated by the Institut Laue-Langevin in Grenoble. IN16B is a neutron backscattering spectrometer with a new “BATS mode” (Backscattering And Time of flight Spectrometer) that considerably extends the accessible energy transfer range.<sup>24,25</sup> The data have been collected for three different temperatures, 283, 303, and 323 K, with an energy resolution of  $3.5 \mu\text{eV}$ , an energy transfer range of  $\pm 150 \mu\text{eV}$ , and a momentum transfer range of  $0.8 \text{ \AA}^{-1} < q < 1.8 \text{ \AA}^{-1}$  in units of  $\hbar$ . All samples and buffers were measured in annular aluminum sample holders with a gap thicknesses of 0.3 mm. The QENS data of the samples were normalized by the vanadium reference, and the data reduction was performed with the MANTID software.<sup>26</sup>

### B. Accompanying experiments

Accompanying experiments with dynamic light scattering (DLS) and Synchrotron Radiation Circular Dichroism (SRCD) spectroscopy have been performed to estimate, respectively, the global diffusion coefficient of MBP and its secondary structure content as a function of temperature and buffer. The details are described in the [supplementary material](#), and we mention here only the respective protein concentrations that were used in the experiments since they are important for the discussion of the results.

Concerning the DLS experiments, we used a protein concentration of 4 mg/ml for the pure  $\text{D}_2\text{O}$  buffer and 7 mg/ml for the  $\text{D}_2\text{O}$  buffer with 30% TFE. We found that MBP coagulates  $T = 323 \text{ K}$  in the pure  $\text{D}_2\text{O}$  buffer, which manifests itself in a drastic decrease in the measured diffusion constant. Therefore, we replaced the latter for subsequent data analysis by an estimate through linear extrapolation from smaller temperatures (see the [supplementary material](#)). The SRCD spectra were recorded with 44 mg/ml for the pure  $\text{D}_2\text{O}$  buffer and with 20 mg/ml for the  $\text{D}_2\text{O}$  buffer with 30% TFE.

## III. THEORY AND DATA ANALYSIS

### A. Scattering functions

Thermal neutron scattering probes the differential scattering cross section of the sample under consideration,

$$\frac{d^2\sigma}{d\Omega d\omega} = \frac{k}{k_0} S(\mathbf{q}, \omega), \quad (1)$$

where  $k_0$  and  $k$  are the momenta of the incident and scattered neutrons, respectively, in units of  $\hbar$  and  $S(\mathbf{q}, \omega)$  is the dynamic structure factor. The latter depends on the kinematical variables  $\mathbf{q}$  and  $\omega$ , which are, respectively, the momentum and energy transfer from the neutron to the sample. It can be written as the time Fourier transform of the intermediate scattering function,

$$S(\mathbf{q}, \omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dt e^{-i\omega t} \mathcal{F}(\mathbf{q}, t), \quad (2)$$

which splits into a coherent and an incoherent part,

$$\mathcal{F}(\mathbf{q}, t) = \mathcal{F}_{\text{coh}}(\mathbf{q}, t) + \mathcal{F}_{\text{inc}}(\mathbf{q}, t), \quad (3)$$

containing, respectively, the information about the collective and average single-atom dynamics of the scattering system,

$$\mathcal{F}_{\text{coh}}(\mathbf{q}, t) = \frac{1}{N} \sum_{ij} b_{i,\text{coh}}^* b_{j,\text{coh}} \left\langle e^{-i\mathbf{q} \cdot \hat{\mathbf{R}}_i(0)} e^{i\mathbf{q} \cdot \hat{\mathbf{R}}_j(t)} \right\rangle, \quad (4)$$

$$\mathcal{F}_{\text{inc}}(\mathbf{q}, t) = \frac{1}{N} \sum_i b_{i,\text{inc}}^2 \left\langle e^{-i\mathbf{q} \cdot \hat{\mathbf{R}}_i(0)} e^{i\mathbf{q} \cdot \hat{\mathbf{R}}_i(t)} \right\rangle. \quad (5)$$

Here,  $\hat{\mathbf{R}}_i(t)$  denotes the position operator of atom  $i$ ,  $\langle \dots \rangle$  denotes a quantum ensemble average, and  $b_{i,\text{coh}}$  and  $b_{i,\text{inc}}$  are the coherent and incoherent scattering lengths of atom  $i$ , respectively.<sup>7,27</sup> Being the quantum time correlation function, the intermediate scattering verifies the symmetry relation

$$\mathcal{F}(\mathbf{q}, t) = \mathcal{F}(-\mathbf{q}, -t + i\beta\hbar), \quad (6)$$

which translates into the detailed balance relation

$$\mathcal{S}(\mathbf{q}, \omega) = \mathcal{S}(-\mathbf{q}, -\omega) e^{\beta\hbar\omega/2} \quad (7)$$

for the dynamic structure factor.

Protein solutions are isotropic systems, and the total observed dynamic structure factor may be expressed as an isotropic angular average over all directions of  $\mathbf{q}$ ,

$$\mathcal{S}_{\text{tot}}(q, t) = \overline{\mathcal{S}_{\text{tot}}(\mathbf{q}, t)} = \frac{1}{4\pi} \iint \sin(\theta) d\theta d\phi \mathcal{S}_{\text{tot}}(q\mathbf{n}(\theta, \phi), t). \quad (8)$$

Here,  $\mathbf{n}_q = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)$  is the unit vector pointing into the direction of  $\mathbf{q}$  and  $q \equiv |\mathbf{q}|$ . The QENS spectra of MBP were then extracted according to

$$\mathcal{S}_{\text{protein}}(q, \omega) \approx \mathcal{S}_{\text{tot}}(q, \omega) - (1 - \phi) \mathcal{S}_{\text{buffer}}(q, \omega), \quad (9)$$

using that the partial specific volume of the protein is  $\bar{v} = 0.73$  ml/g. For a concentration of  $c = 55$  mg/ml, this leads to a volume fraction for the protein of  $\phi = c\bar{v} = 0.04015$  (see Ref. 28). Since the incoherent neutron scattering from hydrogen dominates by far the coherent and incoherent scattering from all other atoms, the dynamic structure factor for MBP may be approximated by

$$\mathcal{S}_{\text{protein}}(q, \omega) \approx b_{H,\text{inc}}^2 \mathcal{S}_{H,\text{protein}}(q, \omega), \quad (10)$$

and we abbreviate in the following  $S(q, \omega) \equiv \mathcal{S}_{H,\text{protein}}(q, \omega)$ , i.e.,

$$S(q, \omega) \equiv \frac{1}{2\pi} \int_{-\infty}^{+\infty} dt e^{-i\omega t} F(\mathbf{q}, t), \quad (11)$$

$$F(q, t) = \frac{1}{N} \sum_{i \in \text{proteinH}} \left\langle e^{-i\mathbf{q} \cdot \hat{\mathbf{R}}_i(0)} e^{i\mathbf{q} \cdot \hat{\mathbf{R}}_i(t)} \right\rangle. \quad (12)$$

The above definition shows that

$$\int_{-\infty}^{+\infty} d\omega S(q, \omega) = 1 = F(q, 0), \quad (13)$$

and this normalization will be used in the following.

## B. Symmetrization of the QENS spectra

For the further data analysis, we will be using the symmetrized dynamic structure factor

$$S^{(+)}(q, \omega) = \frac{e^{-\beta\hbar\omega/2} \mathcal{S}(q, \omega)}{\int_{-\infty}^{+\infty} d\omega' e^{-\beta\hbar\omega'/2} \mathcal{S}(q, \omega')}, \quad (14)$$

which is an even function in  $\omega$  and which is normalized such that

$$\int_{-\infty}^{+\infty} d\omega S^{(+)}(q, \omega) = F^{(+)}(q, 0) = 1. \quad (15)$$

The corresponding intermediate scattering function is a symmetric function in time,

$$F^{(+)}(q, t) = \frac{F(q, t + i\beta\hbar/2)}{F(q, i\beta\hbar/2)}, \quad (16)$$

and we note in this context that  $S^{(+)}(q, \omega) \approx S^{(cl)}(q, \omega)$  and  $F^{(+)}(q, t) \approx F^{(cl)}(q, t)$  correspond to Schofield's semiclassical approximation,<sup>29</sup> where the superscript “(cl)” indicates the corresponding classical scattering function. With this reasoning essentially, classical models can be used for the interpretation of QENS spectra.

## C. Resolution deconvolution

The experimental dynamic structure factor, denoted by an index “ $m$ ,” can be written as the convolution of the “ideal” dynamic structure factor and the instrumental resolution plus a noise,

$$S_m^{(+)}(q, \omega) = \int_{-\infty}^{+\infty} d\omega' S^{(+)}(q, \omega - \omega') \tilde{R}(q, \omega') + \tilde{N}(q, \omega), \quad (17)$$

where  $\tilde{N}(q, \omega)$  and  $\tilde{R}(q, \omega)$  represent the noise and the instrumental resolution, respectively, and are supposed to be symmetric in  $\omega$ . In the context of this paper,  $\tilde{R}(q, \omega) = S_{\text{vanadium}}^{(+)}(q, \omega)$  is the symmetrized dynamic structure factor of a vanadium sample with identical dimensions to the sample of interest, and the noise is essentially unknown. Here, one uses the fact that vanadium can be considered as an ideal isotropic and elastic scatterer. In the time domain, Eq. (17) becomes

$$F_m^{(+)}(q, t) = F^{(+)}(q, t) R(q, t) + N(q, t), \quad (18)$$

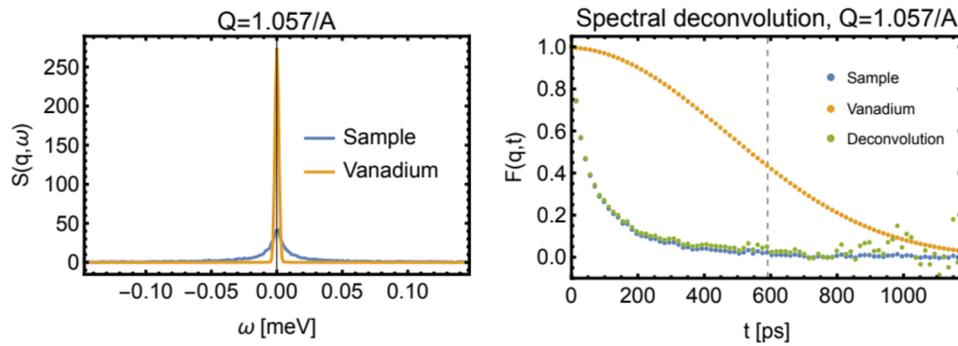
noting that the convolution term in  $\omega$  becomes a normal product. The “ideal” resolution corrected intermediate scattering function is therefore obtained through

$$F^{(+)}(q, t) = \frac{F_m^{(+)}(q, t)}{R(q, t)} - \frac{N(q, t)}{R(q, t)}. \quad (19)$$

The above equation shows that the limit of the instrumental deconvolution by Fourier transform is set by the term  $N(q, t)/R(q, t)$ , which becomes dominant for small values of the resolution window,  $R(q, t)$ . In practice, the threshold for  $R(q, t)$  is attained for

$$t_{\text{max}} \approx \frac{h}{2\Delta\omega}, \quad (20)$$

where  $h$  is Planck's constant and  $\Delta\omega$  is the instrumental energy resolution. For the IN16B spectrometer, on which the QENS data analyzed in this paper have been recorded, we have  $\Delta\omega = 3.5$   $\mu\text{eV}$ , corresponding to  $t_{\text{max}} \approx 600$  ps. Figure 1 illustrates the spectral



**FIG. 1.** Left panel: Symmetrized and normalized dynamic structure factor,  $S_m^{(+)}(q, \omega)$ , of MBP in  $D_2O$  buffer at  $q = 1.06 \text{ \AA}^{-1}$  and  $T = 283 \text{ K}$  (blue line) with the corresponding resolution function  $\tilde{R}(q, \omega)$  (dark yellow line) obtained from a vanadium run. Right panel: Corresponding intermediate scattering function,  $F_m^{(+)}(q, t)$ , in the time domain (dark yellow dots) and resulting deconvoluted intermediate scattering function  $F^{(+)}(q, t)$  (green dots). The vertical dashed line indicates the accessible time domain corresponding to the resolution of the instrument [Eq. (20)].

deconvolution for the QENS spectrum of MBP in  $D_2O$  buffer for  $T = 283 \text{ K}$  at  $q = 1.057/\text{\AA}$ . In the left panel, we show the symmetrized and normalized dynamic structure factor  $S_m^{(+)}(q, \omega)$  (blue line) together with the corresponding spectrum of vanadium,  $\tilde{R}(q, \omega)$ , which defines the instrumental resolution (dark yellow line). Here, individual data points cannot be resolved for the full accessible  $\omega$ -range. The right panel displays the corresponding intermediate scattering functions,  $F_m^{(+)}(q, t)$  (blue dots) and  $R(q, t)$  (dark yellow dots), which have been obtained by discrete (Fast) Fourier transform from equidistantly sampled arrays representing  $S_m^{(+)}(q, \omega)$  and  $\tilde{R}(q, \omega)$ . The resolution-deconvoluted intermediate scattering function,  $F^{(+)}(q, t)$ , is given by the green dots, and the dashed vertical line indicates the threshold of  $t_{\max} \approx 600 \text{ ps}$  corresponding to the instrumental resolution. All displayed data points fulfilling  $t \leq t_{\max}$  have been used for the fit of the model to be described in the following.

## D. QENS model

### 1. Generic form of the intermediate scattering function

The data analysis of the QENS spectra from intrinsically disordered proteins presents the difficulty that the center-of-mass motion and the motions relative to the center of mass are entangled. Here, we use the simplest approximation of the intermediate scattering function that consists in neglecting this entanglement in the functional form,

$$F^{(+)}(q, t) \approx e^{-Dq^2|t|} F_{\text{int}}^{(+)}(q, t), \quad (21)$$

and considering  $D$  to be an *apparent* translational diffusion coefficient that also contains contributions from motions relative to the center of mass. The intermediate scattering function for the internal dynamics of MBP has the generic form

$$F_{\text{int}}^{(+)}(q, t) = EISF(q) + (1 - EISF(q))\phi^{(+)}(q, t), \quad (22)$$

where the symbol  $EISF(q)$  stands for the elastic incoherent structure factor,

$$EISF(q) = \frac{1}{N} \sum_{i \in H} \overline{|e^{iq \cdot \tilde{R}_i}|^2}, \quad (23)$$

and  $\phi^{(+)}(q, t)$  describes the relaxation of  $F^{(+)}(\mathbf{q}, t)$  toward this plateau value. Introducing the fluctuation of the Fourier transformed density of atom  $i$ ,

$$\delta\rho_i(\mathbf{q}, t) = e^{iq \cdot \tilde{R}_i(t)} - \langle e^{iq \cdot \tilde{R}_i} \rangle, \quad (24)$$

the relaxation function  $\phi(q, t)$  can be explicitly expressed as a normalized quantum time correlation function,

$$\phi(q, t) = \frac{\overline{\langle \sum_{i \in H} \delta\rho_i^\dagger(\mathbf{q}, 0) \delta\rho_i(\mathbf{q}, t) \rangle}}{\overline{\langle \sum_{i \in H} \delta\rho_i^\dagger(\mathbf{q}, 0) \delta\rho_i(\mathbf{q}, 0) \rangle}}, \quad (25)$$

and its Fourier transform

$$\tilde{\phi}(q, \omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dt e^{-i\omega t} \phi(q, t) \quad (26)$$

leads to

$$S_{\text{int}}(q, \omega) = EISF(q)\delta(\omega) + (1 - EISF(q))\tilde{\phi}(q, \omega), \quad (27)$$

which shows explicitly that the EISF determines the elastic component of the QENS spectrum for the internal protein dynamics.

### 2. Multiscale relaxation

The internal dynamics in MBP is described by a Mittag-Leffler (ML) relaxation function,<sup>18,19</sup>

$$\phi_{\text{ML}}^{(+)}(t) = E_\alpha(-(|t|/\tau_R)^\alpha), \quad 0 < \alpha < 1, \quad (28)$$

where the  $q$ -dependence is omitted for better readability. Here,  $E_\alpha(\cdot)$  is the Mittag-Leffler (ML) function,<sup>30</sup> which is defined by the Taylor series,

$$E_\alpha(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(1 + \alpha n)} \quad (\alpha > 0), \quad (29)$$

for all arguments in the complex plane. For  $0 < \alpha \leq 1$ , the ML relaxation function (28) decays monotonously, with an asymptotically power law form,

$$\phi_{\text{ML}}(t) \stackrel{t \gg \tau_R}{\sim} \frac{(t/\tau_R)^{-\alpha}}{\Gamma(1-\alpha)}, \quad (30)$$

and exponential relaxation is retrieved for  $\alpha = 1$ ,

$$\lim_{\alpha \rightarrow 1} \phi_{\text{ML}}(t) = \exp(-|t|/\tau), \quad (31)$$

where the power-law long time tail vanishes.

Since the Laplace transform of the ML relaxation function has a simple analytical form,<sup>30</sup>

$$\hat{\phi}_{\text{ML}}(s) = \frac{1}{s(1 + (s\tau_R)^{-\alpha})}, \quad (32)$$

its Fourier and relaxation rate spectra can be easily obtained through

$$\phi^{(+)}(\omega) = \lim_{\epsilon \rightarrow 0} \frac{1}{\pi} \Im \left\{ \hat{\phi}^{(+)}(i\omega + \epsilon) \right\}, \quad (33)$$

which leads to a “generalized Lorentzian,”<sup>31</sup>

$$\tilde{\phi}_{\text{ML}}(\omega) = \frac{1}{\pi} \frac{\sin\left(\frac{\pi\alpha}{2}\right)}{|\omega| \left( (|\omega|\tau_R)^{-\alpha} + (|\omega|\tau_R)^\alpha + 2 \cos\left(\frac{\pi\alpha}{2}\right) \right)}. \quad (34)$$

Here,  $\int_{-\infty}^{+\infty} d\omega \tilde{\phi}_{\text{ML}}(\omega) = 1$ , and the standard Lorentzian is retrieved for  $\alpha \rightarrow 1$ . It should be noted that  $\tilde{\phi}_{\text{ML}}(\omega) \sim |\omega|^{\alpha-1}$  for  $\omega \rightarrow 0$  and becomes thus singular for  $0 < \alpha < 1$ .

The Laplace transform (32) of the ML relaxation function can also be used to obtain an analytical form of its relaxation rate spectrum. For an arbitrary monotonously decaying relaxation function,  $\phi(t)$ , the latter is defined through the relation

$$\phi(t) = \int_0^\infty d\lambda p(\lambda) \exp(-\lambda|t|), \quad (35)$$

where  $p(\lambda) \geq 0$  and  $\int_0^\infty d\lambda p(\lambda) = 1$ . Using that

$$\tilde{\phi}_{\text{ML}}^{(\epsilon)}(\omega) = \frac{\epsilon(\omega^2 + \epsilon^2)^{\alpha/2} + \omega \sin(\alpha \arg(\epsilon + i|\omega|)) + \epsilon \cos(\alpha \arg(\epsilon + i|\omega|))}{(\omega^2 + \epsilon^2) \left( ((\omega^2 + \epsilon^2)^\alpha + 1)(\omega^2 + \epsilon^2)^{-\alpha/2} + 2 \cos(\alpha \arg(\epsilon + i|\omega|)) \right)}. \quad (42)$$

Note that the term  $\epsilon = Dq^2$  makes  $\tilde{\phi}_{\text{ML}}^{(\epsilon)}(\omega)$  regular at  $\omega = 0$ . In this context, it should be noted that the Lorentzian function that is weighted by the EISF in Eq. (41) tends to a Dirac distribution in the limit  $\epsilon \rightarrow 0$ .

#### IV. RESULTS AND DISCUSSION

The results shown in the following have been obtained by the fitting model (38) to the resolution-deconvoluted experimental intermediate scattering functions. To make the model parameters and functions explicitly appear, we write

$$F^{(+)}(q, t) = e^{-Dq^2 t} \left( \text{EISF}(q) + (1 - \text{EISF}(q)) E_\alpha \left( -(t/\tau)^\alpha \right) \right), \quad t \geq 0, \quad (43)$$

where

- $D$  is the apparent translational diffusion coefficient,

$$p(\lambda) = \lim_{\epsilon \rightarrow 0} \frac{1}{\pi} \Im \left\{ \hat{\phi}^{(+)}(-\lambda - i\epsilon) \right\}, \quad (36)$$

one finds for the ML relaxation function<sup>31</sup>

$$p_{\text{ML}}(\lambda) = \frac{\sin(\pi\alpha)}{\pi\lambda \left( (\lambda\tau_R)^{-\alpha} + (\lambda\tau_R)^\alpha + 2 \cos(\pi\alpha) \right)}. \quad (37)$$

The slow power law decay (28) of  $\phi_{\text{ML}}(t)$  leads here to a power law form of the relaxation rate spectrum for small arguments,  $p_{\text{ML}}(\lambda) \stackrel{\lambda \rightarrow 0}{\sim} \lambda^{\alpha-1}$ .

#### 3. Total dynamic structure factor

Using the general form (21) of the total intermediate scattering function, we write the symmetrized version in the form

$$F^{(+)}(t) = e^{-\epsilon|t|} \left( \text{EISF} + (1 - \text{EISF}) \phi_{\text{ML}}(|t|) \right), \quad (38)$$

omitting again the  $q$ -dependence in the formulas and defining

$$\epsilon := Dq^2. \quad (39)$$

The dynamic structure factor can then be expressed in terms of the Laplace transformed intermediate scattering function,

$$S^{(+)}(\omega) = \frac{1}{\pi} \Im \left\{ \hat{F}^{(+)}(i\omega + \epsilon) \right\}, \quad (40)$$

which leads to

$$S^{(+)}(\omega) = \text{EISF} \frac{1}{\pi} \frac{\epsilon}{\omega^2 + \epsilon^2} + (1 - \text{EISF}) \tilde{\phi}_{\text{ML}}^{(\epsilon)}(|\omega|), \quad (41)$$

where  $\tilde{\phi}_{\text{ML}}^{(\epsilon)}(\omega)$  is the Fourier transform of the diffusion-damped ML relaxation function,<sup>32</sup>

- $\text{EISF}(q)$  determines the elastic intensity,
- $\alpha \equiv \alpha(q)$  sets the form of the relaxation function,
- $\tau \equiv \tau(q)$  sets the time scale of the relaxation function.

Fitting all four parameters at the same time is possible but leads to fits with large error bars, in particular, for the apparent translational diffusion coefficient,  $D$ . For this reason, we used the true translational diffusion coefficient obtained from dynamic light scattering (DLS) and, for comparison, the  $q$ -averaged diffusion coefficients,  $\overline{D_{\text{QENS}}(q)}$ , from the four-parameter fit of the QENS spectra. Both fits give very similar values and confidence intervals for  $\tau$ ,  $\alpha$ , and  $\text{EISF}$ , and we show here those for those with a fixed value for  $D_{\text{DLS}}$ . Table I shows both diffusion coefficients, where the error bars of  $\overline{D_{\text{QENS}}(q)}$  are given by the standard deviation of  $D_{\text{QENS}}(q)$  with respect to the mean value  $\overline{D_{\text{QENS}}(q)}$ . The very small value of  $D_{\text{DLS}}$  for  $\text{D}_2\text{O}$  buffer at 323 K indicates the coagulation of MBP, and we

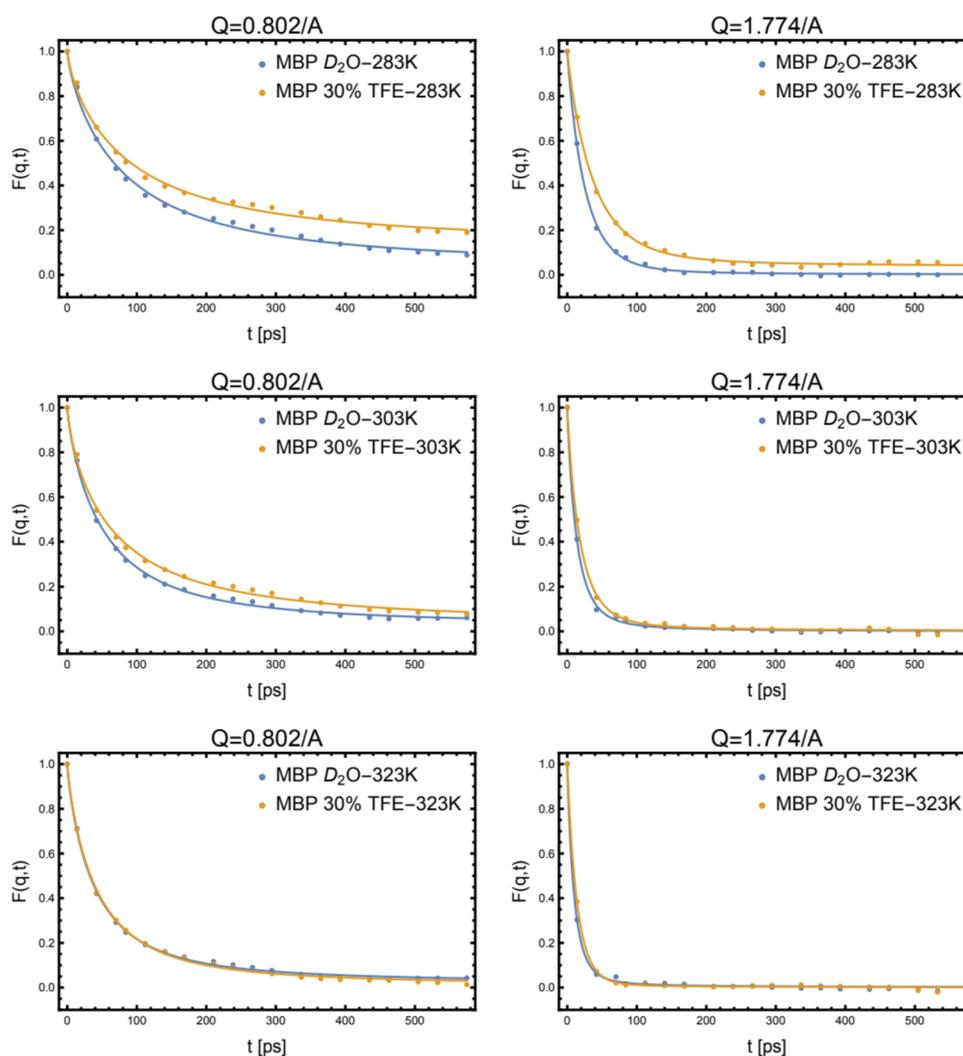
**TABLE I.** Diffusion coefficients from DLS and QENS for MBP in solvents I and II.

Temp (K)	$D_{\text{DLS I}} (\text{\AA}^2/\text{ns})$	$D_{\text{QENS I}} (\text{\AA}^2/\text{ns})$	$D_{\text{DLS II}} (\text{\AA}^2/\text{ns})$	$D_{\text{QENS II}} (\text{\AA}^2/\text{ns})$
283	$3.31 \pm 0.053$	$6.04 \pm 0.46$	$2.05 \pm 0.049$	$5.59 \pm 0.73$
293	$5.72 \pm 0.057$	...	$3.17 \pm 0.145$	...
303	$7.49 \pm 0.090$	$6.15 \pm 0.80$	$4.42 \pm 0.19$	$5.91 \pm 0.70$
313	$9.85 \pm 0.049$	...	$6.32 \pm 0.292$	...
323	$0.31 \pm 0.013$	$6.43 \pm 0.83$	$8.15 \pm 0.19$	$6.29 \pm 0.36$
	11.94 (extrapol.)			

show in the same position also the extrapolated value from lower temperatures for a non-coagulated solution.

The resulting three-parameter fits of the resolution-deconvoluted and symmetrized intermediate scattering function,

$F^{(+)}(q, t)$ , are shown in Fig. 2. All fits have been performed for the two solvents,  $\text{D}_2\text{O}$  buffer (solvent I) and  $\text{D}_2\text{O}$  buffer with 30% deuterated trifluoroethanol (TFE) (solvent II), and the three different temperatures, 283, 303, and 323 K. The blue data points



**FIG. 2.** Plot of the deconvoluted intermediate scattering function  $F^{(+)}(q, t)$  of MBP at different temperatures, comparing solvents I and II (blue and dark yellow points, respectively), and the corresponding fits with model (43) (solid lines) for the minimum and maximum value of  $q$ .

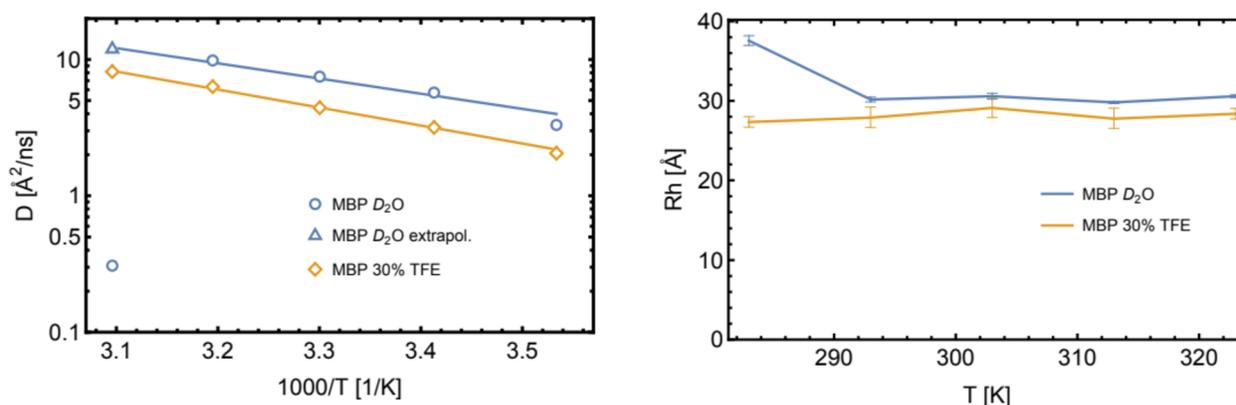


FIG. 3. Left panel: Arrhenius plot of the diffusion coefficient for MBP in solvent I (blue symbols) and solvent II (dark yellow symbols). The solid lines represent the corresponding fitted Arrhenius laws. Right panel: Effective hydrodynamic radii of MBP in the two solvents as a function of temperature.

correspond here to solvent I, the dark yellow points correspond to solvent II, and the corresponding solid lines label the fits. We note that not all data points have been used for the fits but a selection of points, which takes all data for  $F(q, t)$  at short time lags and increasingly less points at large time lags. The reason is that the evaluation of the ML relaxation function is time consuming. We verified that the resulting fit parameters are only weakly influenced by this choice.

We discuss first the results for the diffusion constants in the two buffers at the three temperatures, which have been obtained by DLS and QENS. We recall that MBP in  $D_2O$  buffer coagulates at 323 K, and we note in this context that the DLS experiments have been performed at lower protein concentrations than the QENS experiments (4 mg/ml for solvent I and 7 mg/ml for solvent II), which are thus certainly affected by coagulation. We are though confident that this phenomenon does not affect our results concerning the internal dynamics probed by the IN16B spectrometer since even setting  $D = 0$ , i.e., neglecting completely global diffusion in the fits of  $\tau$ ,  $\alpha$ , and  $EISF$ , does only slightly alter the resulting values for  $\tau$ ,  $\alpha$ , and  $EISF$ . The fits are though sensitive to deviating too much from the DLS and QENS values toward higher values. Replacing  $D_{DLS}$  for solvent I at 323 K by the extrapolated value from lower temperatures is not tolerated since the corresponding width  $Dq^2$  for all  $q$ -values lies clearly outside the resolution of IN16B. The left panel of Fig. 3 shows an Arrhenius plot of the diffusion constants, where for solvent I at 323 K the extrapolated triangular data point has been used instead of the measured one. Here, the blue symbols refer again to solvent I and the dark yellow symbols refer to solvent II. The corresponding solid lines denote the fits of an Arrhenius law,<sup>33,34</sup>

$$D(T) = D_0 e^{-\frac{\Delta G}{k_B T}}, \quad (44)$$

where the activation energy is  $\Delta G = 5.10$  kcal/mol for solvent I and  $\Delta G = 6.05$  kcal/mol for solvent II. We note in this context that the activation energy for solvent I is close to the values that have been found for other IDPs<sup>34</sup> and corresponds to the activation energy for self-diffusion in liquid water.<sup>35,36</sup> The right panel of Fig. 3 shows an estimation of the effective hydrodynamical radius,  $R_h$ , of MBP

as a function of temperature and solvent using the Stokes–Einstein relation

$$D = \frac{k_B T}{6\pi\eta R_h}, \quad (45)$$

with the diffusion coefficients from DLS and experimental values for the dynamic viscosity,  $\eta$ , which are given in the [supplementary material](#). The viscosity was measured with a rolling ball viscometer for both solvents and the relevant temperature range (see the [supplementary material](#)). The results show that  $R_h$  is essentially constant with increasing temperature, and between 283 and 303 K, it

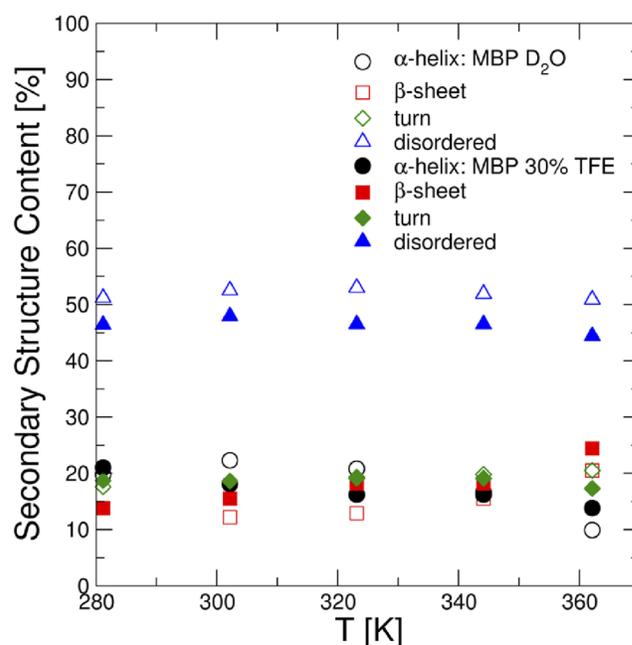


FIG. 4. Secondary structure content of MBP in solvent I (hollow symbols) and solvent II (full symbols).

decreases for MBP in  $D_2O$  buffer. Such a behavior has been, in fact, reported for IDPs,<sup>37,38</sup> but we will here not comment on that. The essential point is that no unfolding transition is observed with increasing temperature, which is characteristic for folded globular proteins. Figure 4 shows that this result is in agreement with the observation from CD spectroscopy that the secondary structure content is essentially constant over the whole temperature range, being slightly larger for the TFE-enriched buffer (solvent II). TFE is known to induce  $\alpha$ -helical motives, and one can observe that at 283 K, but with increasing temperature, this effect disappears continuously, and at high temperatures it is even inverted. The effect of TFE consists thus in moving  $\beta$ -strands into  $\alpha$ -helical motives and vice versa, keeping the total secondary structure content constant.

In Fig. 5, we show the fit parameters concerning the internal dynamics of MBP, comparing again the three different temperatures, 283, 303, and 323 K, and solvents I and II. The time scale parameter  $\tau$  varies considerably with temperature and  $q$ , where the general rule is that it decreases with temperature and with  $q$ . This can be understood since lowering the temperature means slowing

down the dynamics and since the higher values of  $q$  probe increasingly more localized and faster motions. The form parameter,  $\alpha$ , of the relaxation function varies, in contrast, weakly with temperature and increases with  $q$  to values close to 1, indicating a close to exponential relaxation for localized motions. The particularly interesting result is that the relaxation dynamics of MBP in the two solvents becomes similar with increasing the temperature and is practically identical at 323 K. The observation is coherent with the observation that both the hydrodynamic radii and the content and composition of  $\alpha$ -helices and  $\beta$ -strands are comparable. We also see that the coagulation of MBP, which is found by DLS in solvent I at 323 K and certainly present in the QENS samples of higher concentration, has no influence on the internal dynamics of the protein.

Concerning the EISF parameter, we observe that it is generally close to zero for the accessible  $q$ -range, except for solvent II at 283 K and small  $q$ -values, which indicates that at low temperatures solvent II hinders somewhat more large-amplitude motions compared to solvent I. We note in this context that EISFs are not systematically small for our model and refer to recent work on a QENS data

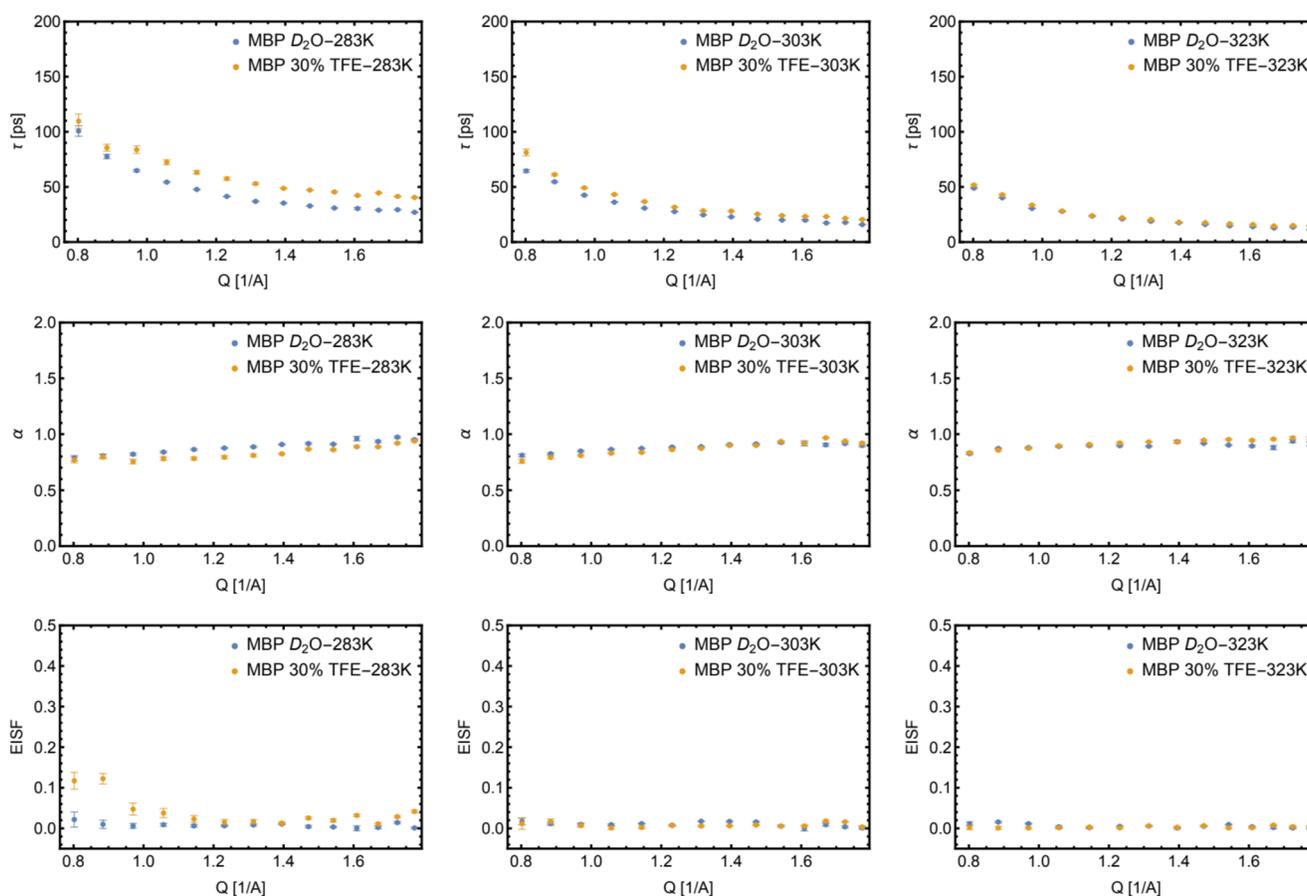


FIG. 5. From top to bottom, the model parameters  $\tau$ ,  $\alpha$ , and  $EISF$  for MBP at different temperatures comparing the two solvents,  $D_2O$  buffer and  $D_2O$  buffer with 30% deuterated TFE.

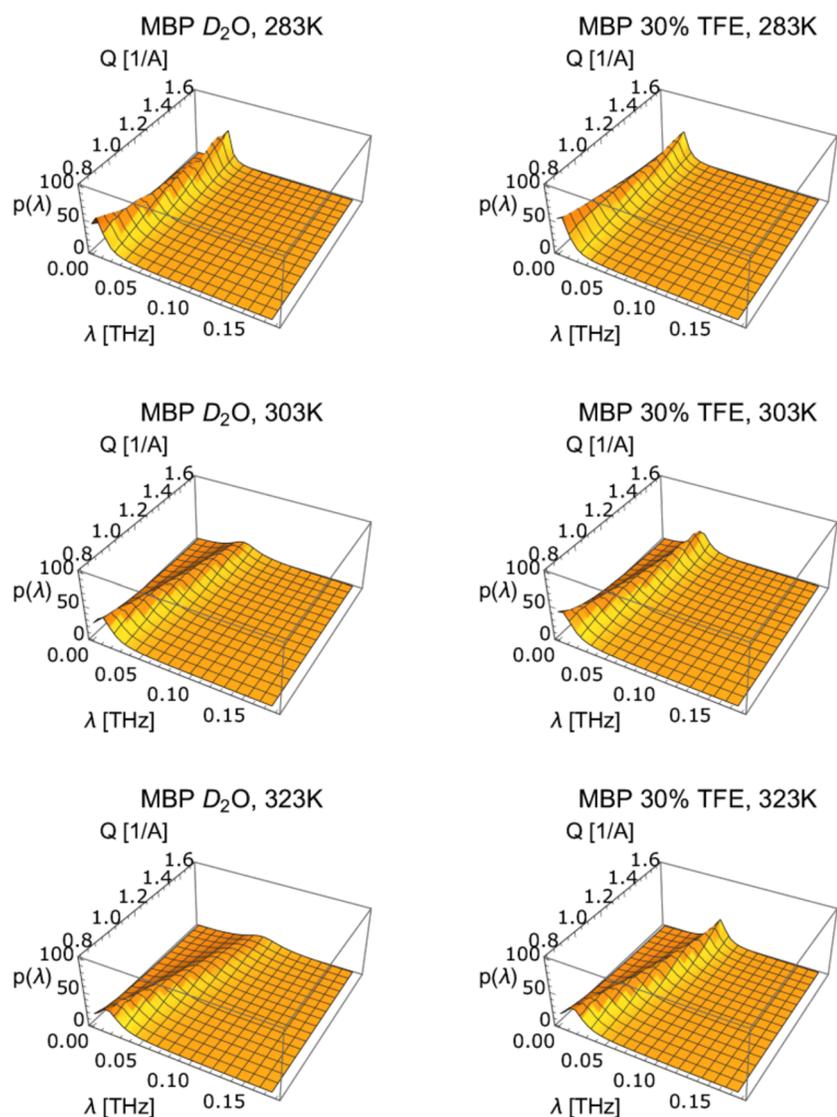
analysis from hydrated powders of human acetylcholinesterase,<sup>19,20</sup> which is an enzyme with a well-defined structure. Here, the EISF was found to be clearly non-zero, as one would expect for such a protein that is more compact and such a sample, where global protein motions are prevented.

To complement the discussion of the fit parameters, we show in Fig. 6 the corresponding relaxation rate spectra,  $p(\lambda)$ , which are defined through Eq. (37). For both solvents and all temperatures, we observe that the peak of the relaxation rate spectrum shifts to higher values with increasing  $q$ , and its width decreases. This reflects the fact that the form parameter  $\alpha$  approaches 1 with increasing  $q$ , noting that  $\alpha = 1$  corresponds to exponential relaxation and a perfectly monodisperse relaxation rate spectrum of the form  $p(\lambda) = \delta(\lambda - 1/\tau)$ ; we are now in the position to give a physical interpretation of the form parameter  $\alpha$ , referring to the concept of protein

energy landscapes proposed by Frauenfelder *et al.*<sup>13</sup> Proceeding as in Ref. 19, we consider the variable

$$\delta\xi(\mathbf{q}, t) = \xi(\mathbf{q}, t) - \langle \xi(\mathbf{q}, t) \rangle, \quad \text{where} \quad \xi(\mathbf{q}, t) = e^{i\mathbf{q}\cdot\mathbf{R}_j(t)}, \quad (46)$$

which is the spatial Fourier transform of the atomic density  $\delta(\mathbf{r} - \mathbf{R}_j(t))$  of atom  $j$ , and assume that the dynamics of  $\delta\xi(\mathbf{q}, t)$  can be modeled as a fractional Ornstein–Uhlenbeck (fOU) process. The latter can be interpreted as the diffusion process in a “rough” parabolic potential, or “energy landscape,” which tends to restore  $\xi(\mathbf{q}, t)$  to its mean value,  $\langle \xi(\mathbf{q}, t) \rangle$ . Since  $\xi(\mathbf{q}, t)$  is a complex function, the real and the imaginary part should here be considered as independent variables, and the left panel of Fig. 7 depicts a sketch of the “rough” parabolic potential as a function of  $\Re\{\delta\xi\}$  and  $\Im\{\delta\xi\}$ . If the variable  $\delta\xi$  was diffusing in a smooth parabolic potential, then



**FIG. 6.** From top to bottom, relaxation rate spectra  $p(\lambda)$  [see Eq. (37)] for  $T = 283$  K,  $T = 303$  K, and  $T = 323$  K, respectively, in  $D_2O$  buffer (left panel) and  $D_2O$  buffer with 30% TFE (right panel).

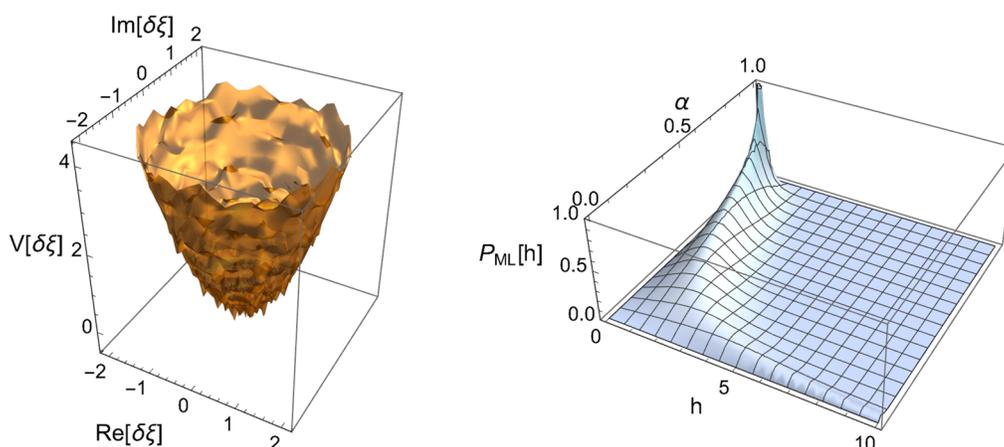


FIG. 7. Left panel: Sketch of a rugged harmonic potential for the real and imaginary part of  $\delta\xi(\mathbf{q}, t)$ . Right panel: Sketch of the energy barrier distribution,  $P(h, \alpha)$ , given in Eq. (48).

its time autocorrelation function would be an exponential function,  $\langle \delta\xi^*(0)\delta\xi(t) \rangle \propto \exp(-t/\tau)$ , and diffusion in the rugged potential leads to longtime memory effects, which are described by a ML relaxation function, i.e.,  $\langle \delta\xi^*(0)\delta\xi(t) \rangle \propto E_\alpha(-t/\tau)^\alpha$ . As discussed in Ref. 19, one can find the distribution of the energy barriers in the “rough” parabolic potential by using the diffusion model proposed by Zwanzig,<sup>39</sup> which describes diffusion of a particle in a rough harmonic potential with a *fixed* energy barrier. Introducing the dimensionless energy barrier height

$$h = \frac{\Delta E}{k_B T}, \quad (47)$$

the distribution for the energy barriers corresponding to the ML relaxation function is found to be<sup>19</sup>

$$P_{\text{ML}}(h) = \frac{2h \sin(\pi\alpha)}{\pi(e^{-\alpha h^2} + e^{\alpha h^2} + 2 \cos(\pi\alpha))}. \quad (48)$$

The right panel of Fig. 7 displays this distribution of energy barriers as a function of  $u$  and  $\alpha$ . It shows that  $\lim_{\alpha \rightarrow 1} P_{\text{ML}}(h) = \delta(h)$ , i.e., becomes more and more the energy barrier distribution for a smooth potential with vanishing barrier height, whereas for  $\alpha \rightarrow 0$ , the distribution  $P_{\text{ML}}(h)$  becomes very broad, including also barriers of almost infinite height. Having this interpretation in mind, the plots of the fit parameters shown in Fig. 5 show that localized motions are characterized by a smoother potential than global ones ( $\alpha$  increases with  $q$ ). This simply reflects that the relaxation rate spectrum of large-amplitude motions in a protein is more disperse than for localized motion since they require the concerted motion of many modes with different relaxation time scales. Concerning the influence of the solvent on the effective local potential of the hydrogen atoms, one sees that solvent II leads to a slightly rougher effective potential for  $\delta\xi$  at  $T = 283$  K compared to solvent I. This difference vanishes with increasing temperature, which makes the differences in the relaxation dynamics of MBP disappear.

## V. CONCLUSIONS

The present analysis of high-resolution quasielastic neutron scattering spectra from myelin basic protein shows that “minimalistic” multiscale models for the relaxation dynamics of proteins combined with results from dynamic light scattering yield meaningful insights into the dynamics of intrinsically disordered proteins and permit, in particular, the detection and interpretation of changes in the relaxation dynamics as a function of solvent, temperature, and spatial resolution. With our choice of the Mittag-Leffler relaxation function as a model for the relaxation of the intermediate scattering function, the relaxation dynamics of MBP can be interpreted in terms of a simple energy landscape model. The QENS analysis shows that a trifluoroethanol-enriched solvent buffer at 283 K makes this energy landscape rougher, which leads to slower relaxation dynamics with a more disperse relaxation rate spectrum, and that this effect disappears continuously with increasing temperature. In accordance with this observation, the complementary experiments with dynamic light scattering and SRCD experiments show that TFE reduces at 283 K the radius of hydration of MBP and increases the  $\alpha$ -helix content at the expense of  $\beta$ -strands. These effects disappear with increasing temperature, where the radii of hydration are comparable and constant, and at higher temperatures even more  $\beta$ -strands appear. Independent of if trifluoroethanol is added to the  $D_2O$  buffer or not, there is no unfolding transition with increasing temperature, which would manifest itself by a sudden increase in the hydrodynamic radius and which is typical for folded proteins. Our CD measurements confirm this observation, showing that the global secondary structure content in the presence of TFE is constant with temperature and only slightly higher compared to a pure  $D_2O$  buffer. To confirm our findings, small-angle x-ray scattering/small-angle neutron scattering (SAXS/SANS) will certainly be useful to obtain information about the radii of gyration and possible aggregation of MBP at exactly the same protein concentrations and solvents, which have been used for the QENS experiments.

From a technical point of view, we think that the high quality of the data and the extended energy transfer range of the IN16B

spectrometer with the new “BATS mode” (Backscattering And Time of flight Spectrometer) were crucial to obtain the results quoted above. We finally remark that all numerical and several analytical calculations have been performed with the Wolfram Mathematica package.<sup>40</sup>

## SUPPLEMENTARY MATERIAL

In the [supplementary material](#), we give information about complementary experiments by dynamic light scattering, viscosimetry, and synchrotron radiation circular dichroism spectroscopy on MBP as a function of solvent and temperature. We show, in particular, the global diffusion constants measured by DLS, the corresponding viscosities that have been used to compute the respective radii of hydration, and the SRCD spectra that have been used to evaluate the secondary structure content of MBP.

## ACKNOWLEDGMENTS

We gratefully acknowledge beam time allocation for the neutron scattering experiments on the IN16B spectrometer at the Institut Laue-Langevin, Grenoble, France (<http://doi.ill.fr/10.5291/ILL-DATA.8-04-874>), and for the synchrotron radiation spectroscopy experiments on the DISCO beam line at the Synchrotron SOLEIL, Gif-sur-Yvette, France. In this context, we thank Frank Wien for his help with recording and analyzing the CD spectra.

A.N.H. acknowledges financial support from the Région Centre and the Jülich Center for Neutron Science.

## AUTHOR DECLARATIONS

### Conflict of Interest

The authors have no conflicts to disclose.

## DATA AVAILABILITY

The neutron scattering data are available at <http://doi.ill.fr/10.5291/ILL-DATA.8-04-874>, and the SRCD data are available from the corresponding author upon reasonable request.

## REFERENCES

- 1 P. Tompa, *Trends Biochem. Sci.* **27**, 527 (2002).
- 2 H. J. Dyson and P. E. Wright, *Nat. Rev. Mol. Cell Biol.* **6**, 197 (2005).
- 3 P. Tompa, *Trends Biochem. Sci.* **37**, 509 (2012).
- 4 P. E. Wright and H. J. Dyson, *Nat. Rev. Mol. Cell Biol.* **16**, 18 (2015).
- 5 M. Vincenzi, F. A. Mercurio, and M. Leone, *Curr. Protein Pept. Sci.* **20**, 425 (2019).
- 6 T. Springer, *Quasielastic Neutron Scattering for the Investigation of Diffusive Motions in Solids and Liquids* (Springer, Berlin, Heidelberg, 1972), Vol. 64, ISBN: 978-3-540-05808-3; 978-3-540-37457-2.
- 7 M. Bée, *Quasielastic Neutron Scattering: Principles and Applications in Solid State Chemistry, Biology and Materials Science* (Adam Hilger, Bristol, 1988).
- 8 J. Colmenero and A. Arbe, *J. Polym. Sci., Part B: Polym. Phys.* **51**, 87 (2013).
- 9 A. M. Stadler, L. Stingaciu, A. Radulescu, O. Holderer, M. Monkenbusch, R. Biehl, and D. Richter, *J. Am. Chem. Soc.* **136**, 6987 (2014).
- 10 L. R. Stingaciu, R. Biehl, D. Changwoo, D. Richter, and A. M. Stadler, *J. Phys. Chem. Lett.* **11**, 292 (2020).
- 11 N. Alberding, R. H. Austin, S. S. Chan, L. Eisenstein, H. Frauenfelder, I. C. Gunsalus, and T. M. Nordlund, *J. Chem. Phys.* **65**, 4701 (1976).
- 12 R. H. Austin, K. W. Beeson, L. Eisenstein, H. Frauenfelder, and I. C. Gunsalus, *Biochemistry* **14**, 5355 (1975).
- 13 W. Doster, S. Cusack, and W. Petry, *Nature* **337**, 754 (1989).
- 14 I. E. T. Iben, D. Braunstein, W. Doster, H. Frauenfelder, M. K. Hong, J. B. Johnson, S. Luck, P. Ormos, A. Schulte, P. J. Steinbach *et al.*, *Phys. Rev. Lett.* **62**, 1916 (1989).
- 15 H. Frauenfelder, S. G. Sligar, and P. G. Wolynes, *Science* **254**, 1598 (1991).
- 16 W. G. Glöckle and T. F. Nonnenmacher, *Biophys. J.* **68**, 46 (1995).
- 17 H. Yang and X. S. Xie, *J. Chem. Phys.* **117**, 10965 (2002).
- 18 H. Yang, G. Luo, P. Karnchanaphanurach, T.-M. Louie, I. Rech, S. Cova, L. Xun, and X. S. Xie, *Science* **302**, 262 (2003).
- 19 S. C. Kou and X. S. Xie, *Phys. Rev. Lett.* **93**, 180603 (2004).
- 20 G. R. Kneller, *Proc. Natl. Acad. Sci. U. S. A.* **115**, 9450 (2018).
- 21 M. Saouessi, J. Peters, and G. R. Kneller, *J. Chem. Phys.* **150**, 161104 (2019).
- 22 M. Saouessi, J. Peters, and G. R. Kneller, *J. Chem. Phys.* **151**, 125103 (2019).
- 23 E. H. Eylar and M. Thompson, *Arch. Biochem. Biophys.* **129**, 468 (1969).
- 24 M. Appel, B. Frick, and A. Magerl, *Sci. Rep.* **8**, 13580 (2018).
- 25 M. Appel, B. Frick, and A. Magerl, *Physica B* **562**, 6 (2019).
- 26 O. Arnold, J. C. Bilheux, J. M. Borreguero, A. Buts, S. I. Campbell, L. Chapon, M. Doucet, N. Draper, R. Ferraz Leal, M. A. Gigg *et al.*, *Nucl. Instrum. Methods Phys. Res., Sect. A* **764**, 156 (2014).
- 27 S. Lovesey, *Theory of Neutron Scattering from Condensed Matter* (Clarendon Press, Oxford, 1984), Vol. I.
- 28 A. M. Stadler, F. Demmel, J. Ollivier, and T. Seydel, *Phys. Chem. Chem. Phys.* **18**, 21527 (2016).
- 29 P. Schofield, *Phys. Rev. Lett.* **4**, 239 (1960).
- 30 *Mittag-Leffler Functions, Related Topics and Applications*, Springer Monographs in Mathematics, edited by R. Gorenflo, A. A. Kilbas, F. Mainardi, and S. V. Rogosin (Springer, Heidelberg, 2014), ISBN: 978-3-662-43929-6; 978-3-662-43930-2.
- 31 G. R. Kneller, *Phys. Chem. Chem. Phys.* **7**, 2641 (2005).
- 32 V. Calandrini, V. Hamon, K. Hinsen, P. Calligari, M.-C. Bellissent-Funel, and G. R. Kneller, *Chem. Phys.* **345**, 289 (2008).
- 33 K. Sozański, A. Wiśniewska, T. Kalwarczyk, and R. Hołyst, *Phys. Rev. Lett.* **111**, 228301 (2013).
- 34 I. V. Nesmelova, D. L. Melnikova, V. Ranjan, and V. D. Skirda, *Progress in Molecular Biology and Translational Science* (Elsevier, 2019), Vol. 166, pp. 85–108, ISBN: 978-0-12-816851-6.
- 35 D. W. McCall, D. C. Douglass, and E. W. Anderson, *J. Chem. Phys.* **31**, 1555 (1959).
- 36 R. Mills, *J. Phys. Chem.* **77**, 685 (1973).
- 37 M. Kjaergaard, A.-B. Norholm, R. Hendus-Altenburger, S. F. Pedersen, F. M. Poulsen, and B. B. Kragelund, *Protein Sci.* **19**, 1555 (2010).
- 38 R. Wuttke, H. Hofmann, D. Nettels, M. B. Borgia, J. Mittal, R. B. Best, and B. Schuler, *Proc. Natl. Acad. Sci. U. S. A.* **111**, 5213 (2014).
- 39 R. Zwanzig, *Proc. Natl. Acad. Sci. U. S. A.* **85**, 2029 (1988).
- 40 Wolfram Research, Inc., Mathematica, Version 12.0, Champaign, IL, 2019.

## **2. Quasi-analytical resolution-correction of elastic neutron scattering from proteins**

Cet article présente une analyse des spectres de diffusion quasi-élastique de neutrons de la protéine basique de la myéline (MBP) et de la Myoglobine (Mb) en solution, cette étude a été réalisée pour confirmer le résultat de la première étude où on a trouvé un EISF proche de zéro, et pour montrer que notre modèle ne conduit pas systématiquement à la disparition de l'EISF pour les protéines en solution, une analyse de données QENS de la myoglobine a été réalisée et on a trouvé des valeurs de l'EISF clairement non nuls. Il a été publié dans le "Journal of Chemical Physics".

# Quasi-analytical resolution-correction of elastic neutron scattering from proteins

Cite as: J. Chem. Phys. **157**, 134103 (2022); <https://doi.org/10.1063/5.0103960>

Submitted: 17 June 2022 • Accepted: 07 September 2022 • Accepted Manuscript Online: 07 September 2022 • Published Online: 05 October 2022

 Abir N. Hassani,  Andreas M. Stadler and  Gerald R. Kneller



View Online



Export Citation



CrossMark

## ARTICLES YOU MAY BE INTERESTED IN

Multiscale relaxation dynamics and diffusion of myelin basic protein in solution studied by quasielastic neutron scattering

The Journal of Chemical Physics **156**, 025102 (2022); <https://doi.org/10.1063/5.0077100>

Relating dynamic free volume to cooperative relaxation in a glass-forming polymer composite

The Journal of Chemical Physics **157**, 131101 (2022); <https://doi.org/10.1063/5.0114902>

Communication: Electronic transition of the I-C<sub>6</sub><sup>+</sup> cation at 417 nm

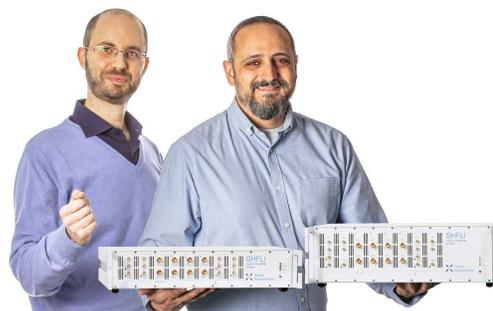
The Journal of Chemical Physics **157**, 121102 (2022); <https://doi.org/10.1063/5.0106183>

Webinar

Meet the Lock-in Amplifiers  
that measure microwaves

Oct. 6th – Register now

 Zurich  
Instruments



# Quasi-analytical resolution-correction of elastic neutron scattering from proteins

Cite as: J. Chem. Phys. 157, 134103 (2022); doi: 10.1063/5.0103960

Submitted: 17 June 2022 • Accepted: 7 September 2022 •

Published Online: 5 October 2022



View Online



Export Citation



CrossMark

Abir N. Hassani,<sup>1,2,3</sup>  Andreas M. Stadler,<sup>3,4,a)</sup>  and Gerald R. Kneller<sup>1,2,b)</sup> 

## AFFILIATIONS

<sup>1</sup>Centre de Biophysique Moléculaire, CNRS and Univ. d'Orléans, Rue Charles Sadron, 45071 Orléans, France

<sup>2</sup>Synchrotron Soleil, L'Orme de Merisiers, 91192 Gif-sur-Yvette, France

<sup>3</sup>Jülich Centre for Neutron Science (JCNS-1) and Institute of Biological Information Processing (IBI-8), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

<sup>4</sup>Institute of Physical Chemistry, RWTH Aachen University, Landoltweg 2, 52056 Aachen, Germany

<sup>a)</sup>E-mail: [a.stadler@fz-juelich.de](mailto:a.stadler@fz-juelich.de)

<sup>b)</sup>Author to whom correspondence should be addressed: [gerald.kneller@cnrs.fr](mailto:gerald.kneller@cnrs.fr)

## ABSTRACT

Elastic neutron scattering from proteins reflects the motional amplitudes resulting from their internal collective and single-atom dynamics and is observable if the global diffusion of whole molecules is either blocked or cannot be resolved by the spectrometer under consideration. Due to finite instrumental resolution, the measured elastic scattering amplitude always contains contaminations from quasielastic neutron scattering and some model must be assumed to extract the resolution-corrected counterpart from corresponding experimental spectra. Here, we derive a quasi-analytical method for that purpose, assuming that the intermediate scattering function relaxes with a “stretched” Mittag-Leffler function,  $E_\alpha(-t/\tau)^\alpha$  ( $0 < \alpha < 1$ ), toward the elastic amplitude and that the instrumental resolution function has Gaussian form. The corresponding function can be integrated into a fitting procedure and allows for eliminating the elastic intensity as a fit parameter. We illustrate the method for the analysis of two proteins in solution, the intrinsically disordered Myelin Basic Protein, confirming recently published results [Hassani *et al.*, J. Chem. Phys. 156, 025102 (2022)], and the well-folded globular protein myoglobin. We also briefly discuss the consequences of our findings for the extraction of mean square position fluctuations from elastic scans.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0103960>

## I. INTRODUCTION

Thermal neutron scattering is a powerful and versatile spectroscopic method to probe the structural dynamics of condensed matter systems.<sup>1</sup> An important application concerns quasielastic neutron scattering (QENS) from proteins, which gives information about the diffusion and the relaxation dynamics of these macromolecules.<sup>2–6</sup> To probe the internal non-exponential multiscale relaxation dynamics, which is crucial for their function and typical for complex systems, in general,<sup>7–10</sup> one can either use hydrated powder samples, where global diffusional motions are simply blocked, or probe a protein solution with a spectrometer that will not resolve these motions. In both cases, information about the motional amplitudes of internal protein dynamics is contained in the elastic amplitude and elastic scans are thus, in principle, sufficient to obtain this information. One must, however, be aware that the extracted motional amplitudes are underestimated due to the unavoidable contamination of

the elastic amplitude by contributions from quasielastic scattering, and this correction can be particularly important for slowly relaxing systems.<sup>11</sup> Noting that the “true” elastic amplitude defines the asymptotic form of the neutron intermediate scattering function at infinite time, it can only be obtained by assuming some model for that function. A corresponding “minimalistic” model has been recently proposed and motivated in Ref. 12 and was then applied in a few subsequent QENS studies of protein dynamics,<sup>13–15</sup> as well as for confined water molecules in clays.<sup>16</sup> In all these studies the elastic amplitude was a fit parameter, which left some ambiguity about the physical significance of the resulting fits, in particular, since fit parameters are quite interdependent. The goal of this paper is to replace the elastic intensity as a fit parameter by an estimation on the basis of its experimentally measured counterpart, the assumed model for the relaxation function, and the resolution of the instrument under consideration. Computational efficiency is here a fundamental aspect since it enables the integration of the corresponding function into

the fitting procedure for the remaining parameters of the relaxation function.

This paper is organized as follows: The core of the paper is contained in Secs. II and III, which describe the theoretical background and the method, followed by Sec. IV showing some applications, and the conclusions in Sec. V.

## II. THEORETICAL BACKGROUND

### A. Scattering functions

In standard neutron scattering experiments one measures the dynamic structure factor,

$$\tilde{F}(\mathbf{q}, \omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dt e^{-i\omega t} F(\mathbf{q}, t), \quad (1)$$

which is the time Fourier transform of the intermediate scattering function containing the information about the structural dynamics of the system under consideration,

$$F(\mathbf{q}, t) = \frac{1}{N} \sum_{j,k} \Gamma_{jk} \left\langle e^{-i\mathbf{q} \cdot \hat{\mathbf{R}}_j(0)} e^{i\mathbf{q} \cdot \hat{\mathbf{R}}_k(t)} \right\rangle. \quad (2)$$

Usually, the dynamic structure factor is denoted by  $S(\mathbf{q}, \omega)$ , but we use the symbol  $\tilde{F}(\mathbf{q}, \omega)$  to label Fourier transforms in a uniform way. The scattering-related quantities are, respectively, the momentum and energy transfer from the neutron to the sample,  $\mathbf{q}$  and  $\omega$ , in units of  $\hbar$ ,  $N$  is the total number of atoms in the scattering system and for each pair  $\{j, k\}$  of them,  $\{\hat{\mathbf{R}}_j(t), \hat{\mathbf{R}}_k(t)\}$  denote the associated time-dependent position operators. The symbol  $\langle \dots \rangle$  stands for a quantum ensemble average and the weighting factors  $\Gamma_{jk}$  have the form

$$\Gamma_{jk} = \overline{b_j^* b_k} + \delta_{jk} \overline{|b_j - b_k|^2}, \quad (3)$$

where  $b_j$  and  $b_k$  are the (complex) scattering lengths<sup>1,17</sup> of the atoms  $j$  and  $k$ , respectively. For a given atom, the average runs over all isotopes and combinations of the nuclear and neutron spins and we note that  $b_{j,\text{coh}} \equiv \overline{b_j}$  and  $b_{j,\text{inc}} \equiv \overline{(|b_j - \overline{b_j}|^2)^{1/2}}$  are, respectively, the coherent and incoherent scattering lengths of atom  $j$ . Coherent and incoherent scattering probe, respectively, the collective and average single atom dynamics of the system under consideration, but since these scattering types are not separable without special spin-polarization experiments,<sup>18–20</sup> we will not explicitly distinguish between them.

The intermediate scattering function fulfills the symmetry relations of a quantum time correlation function,

$$F^*(\mathbf{q}, t) = F(\mathbf{q}, -t), \quad (4)$$

$$F(\mathbf{q}, -t) = F(-\mathbf{q}, t + i\beta\hbar), \quad (5)$$

where  $\beta = 1/k_B T$  is the inverse Boltzmann temperature. For the dynamic structure factor, Eq. (5) translates into

$$\tilde{F}(\mathbf{q}, \omega) = e^{\beta\hbar\omega} \tilde{F}(-\mathbf{q}, -\omega), \quad (6)$$

which is the well-known detailed-balance relation.

### B. Elastic and inelastic scattering

Noting that

$$e^{i\mathbf{q} \cdot \hat{\mathbf{R}}_j(t)} = \int d^3r \exp(-i\mathbf{q} \cdot \mathbf{r}) \delta(\mathbf{r} - \hat{\mathbf{R}}_j(t))$$

is the spatially Fourier-transformed single particle density for atom  $j$ , we introduce the deviation of this quantity with respect to its mean value,

$$\delta\tilde{\alpha}_k(\mathbf{q}, t) = e^{i\mathbf{q} \cdot \hat{\mathbf{R}}_j(t)} - \left\langle e^{i\mathbf{q} \cdot \hat{\mathbf{R}}_j(t)} \right\rangle, \quad (7)$$

to split the intermediate scattering function into a static and a time-dependent component,

$$F(\mathbf{q}, t) = F(\mathbf{q}, \infty) + \delta F(\mathbf{q}, t), \quad (8)$$

which are given by

$$F(\mathbf{q}, \infty) = \frac{1}{N} \sum_{j,k} \Gamma_{jk} \left\langle e^{i\mathbf{q} \cdot \hat{\mathbf{R}}_j} \right\rangle^* \left\langle e^{i\mathbf{q} \cdot \hat{\mathbf{R}}_k} \right\rangle, \quad (9)$$

$$\delta F(\mathbf{q}, t) = \frac{1}{N} \sum_{j,k} \Gamma_{jk} \left\langle \delta\tilde{\alpha}_j^\dagger(\mathbf{q}, 0) \delta\tilde{\alpha}_k(\mathbf{q}, t) \right\rangle. \quad (10)$$

Making the physically reasonable assumption

$$\lim_{t \rightarrow \infty} \delta F(\mathbf{q}, t) = 0 \quad (11)$$

shows that  $F(\mathbf{q}, \infty)$  is the asymptotic form of the intermediate scattering function and it follows by the Fourier transform of Eq. (8) that

$$\tilde{F}(\mathbf{q}, \omega) = F(\mathbf{q}, \infty) \delta(\omega) + \delta\tilde{F}(\mathbf{q}, \omega). \quad (12)$$

Therefore,  $F(\mathbf{q}, \infty)$  represents the elastic amplitude of the Fourier spectrum and  $\delta\tilde{F}(\mathbf{q}, \omega)$  its inelastic component. Here, “inelastic” is to be understood as “non-elastic” and includes also the quasielastic component of the spectrum, which is very close to the elastic line and describes relaxation and diffusion processes.

### C. Generic form of the scattering functions

For modeling purposes, it is convenient to introduce the normalized relaxation function

$$\phi(\mathbf{q}, t) = \delta F(\mathbf{q}, t) / \delta F(\mathbf{q}, 0), \quad (13)$$

noting that this function does not monotonously decay for short times. This leads to the generic form

$$F(\mathbf{q}, t) = F(\mathbf{q}, \infty) + (F(\mathbf{q}, 0) - F(\mathbf{q}, \infty)) \phi(\mathbf{q}, t) \quad (14)$$

of the intermediate scattering function, which translates into the corresponding generic form

$$\tilde{F}(\mathbf{q}, \omega) = F(\mathbf{q}, \infty) \delta(\omega) + (F(\mathbf{q}, 0) - F(\mathbf{q}, \infty)) \tilde{\phi}(\mathbf{q}, \omega), \quad (15)$$

of the dynamic structure factor. We note that

$$F(\mathbf{q}, 0) = \frac{1}{N} \sum_{j,k} \Gamma_{jk} \left\langle e^{i\mathbf{q} \cdot (\hat{\mathbf{R}}_k - \hat{\mathbf{R}}_j)} \right\rangle \quad (16)$$

is the total static structure factor, which tends for  $q \equiv |\mathbf{q}| \rightarrow \infty$  to a constant value,

$$\lim_{q \rightarrow \infty} F(\mathbf{q}, 0) = \frac{1}{N} \sum_k \Gamma_{kk}, \quad (17)$$

and oscillates for smaller  $q$ -values around that constant. For modeling purposes, it is convenient to normalize the intermediate scattering function such that

$$\frac{1}{N} \sum_k \Gamma_{kk} = 1. \quad (18)$$

#### D. Hydrogen-rich systems

We finally consider the frequent case of neutron scattering from hydrogen-rich systems, such as proteins and polymers. Because of the exceptionally large cross section for incoherent scattering from hydrogen, it is sufficient to consider only these atoms, setting  $\Gamma_{jk} = \delta_{jk}$ . With the normalization (18) it then follows that

$$F(\mathbf{q}, 0) = 1, \quad (19)$$

and the generic form of the intermediate scattering function simplifies to

$$F(\mathbf{q}, t) = EISF(\mathbf{q}) + (1 - EISF(\mathbf{q}))\phi(\mathbf{q}, t), \quad (20)$$

where

$$EISF(\mathbf{q}) = \frac{1}{N} \sum_{j \in \mathbb{H}} \left| \left\langle e^{i\mathbf{q} \cdot \mathbf{R}_j} \right\rangle \right|^2 \quad (21)$$

is referred to as elastic incoherent structure factor and the relaxation function has the form

$$\phi(\mathbf{q}, t) = \frac{\sum_{j \in \mathbb{H}} \left\langle \delta \tilde{\alpha}_j^\dagger(\mathbf{q}, 0) \delta \tilde{\alpha}_j(\mathbf{q}, t) \right\rangle}{\sum_{j \in \mathbb{H}} \left\langle \delta \tilde{\alpha}_j^\dagger(\mathbf{q}, 0) \delta \tilde{\alpha}_j(\mathbf{q}, 0) \right\rangle}. \quad (22)$$

### III. PSEUDOELASTIC SCATTERING

#### A. Measured and true elastic intensity

We will now consider a measured dynamic structure factor, which is always broadened due to finite instrumental resolution. Defining  $\tilde{R}(\omega)$  to be the instrumental resolution function and omitting for simplicity the  $q$ -dependence of the relevant quantities, the measured dynamic structure factor is given by the frequency convolution of the true dynamic structure factor and the resolution function,

$$\tilde{F}_m(\omega) = (\tilde{R} * \tilde{F})(\omega) \equiv \int_{-\infty}^{+\infty} d\omega' \tilde{R}(\omega - \omega') \tilde{F}(\omega'). \quad (23)$$

With (15), we obtain then in a first step

$$\tilde{F}_m(\omega) = F(\infty) \tilde{R}(\omega) + (F(0) - F(\infty)) (\tilde{R} * \tilde{\phi})(\omega). \quad (24)$$

We define now the measured elastic intensity through the integral

$$F_m(\infty) \equiv \int_{-\epsilon}^{+\epsilon} d\omega \tilde{F}_m(\omega), \quad (25)$$

where  $\epsilon > 0$  is defined such that

$$\int_{-\epsilon}^{+\epsilon} d\omega \tilde{R}(\omega) \lesssim 1, \quad (26)$$

and the measured total static structure factor through

$$F_m(0) \equiv \int_{\omega_{\min}}^{\omega_{\max}} d\omega \tilde{F}_m(\omega), \quad (27)$$

where  $[\omega_{\min}, \omega_{\max}]$  is the dynamical range of the instrument. It follows then from the generic form (24) of the measured, resolution-broadened dynamic structure factor that

$$F_m(\infty) \approx F(\infty) + (F_m(0) - F(\infty))\xi, \quad (28)$$

where  $\xi$  is the pseudoelastic contribution due to finite instrumental resolution,

$$\xi = \int_{-\epsilon}^{+\epsilon} d\omega (\tilde{R} * \tilde{\phi})(\omega). \quad (29)$$

Supposing that this contribution can be reliably computed on the basis of appropriate models for the relaxation function and the instrumental resolution, the “true” elastic intensity may be estimated through

$$F(\infty) \approx \frac{F_m(\infty) - \xi F_m(0)}{1 - \xi}. \quad (30)$$

For essentially incoherent scattering, the measured total structure factor is not needed, since one knows that the incoherent static structure factor is simply a constant. Assuming the normalization (18), one can replace  $F_m(0) \rightarrow 1$  in this case. It is also worthwhile noting that the standard definition of the elastic amplitude<sup>21,22</sup> corresponds in our notation to the measured one.

#### B. Model

##### 1. Symmetrized correlation function

The symmetry relations (4) and (5) show that the intermediate scattering function becomes a real symmetrical function in time if one considers the classical limit  $\hbar \rightarrow 0$  and if one can assume that the scattering functions are invariant with respect to the parity operation  $\mathbf{q} \rightarrow -\mathbf{q}$ . Based on this observation, Schofield proposed to use the time-symmetrized real function  $F^{(+)}(\mathbf{q}, t) \equiv F(\mathbf{q}, t + i\beta\hbar/2)$  to define the semiclassical approximation  $F^{(+)}(\mathbf{q}, t) \approx F^{(cl)}(\mathbf{q}, t)$ .<sup>23</sup> To be able to work with classical relaxation models, we consider now the time-symmetrized generic form

$$F^{(+)}(t) = F(\infty) + (F(0) - F(\infty))\phi^{(+)}(t) \quad (31)$$

of the intermediate scattering function, where the relaxation function is defined as

$$\phi^{(+)}(t) \equiv \frac{\phi(t + i\beta\hbar/2)}{\phi(i\beta\hbar/2)}, \quad (32)$$

in order to ensure its correct normalization.

## 2. “Minimalistic” model

We assume now that the symmetrized relaxation function is well represented by the model

$$\phi^{(+)}(t) = \phi_{\text{ML}}(t), \quad (33)$$

where  $\phi_{\text{ML}}(\cdot)$  is the “stretched” Mittag-Leffler (ML) function,

$$\phi_{\text{ML}}(t) \equiv E_{\alpha}(-(|t|/\tau)^{\alpha}) \quad (0 < \alpha \leq 1). \quad (34)$$

The Mittag-Leffler function,  $E_{\alpha}(z)$ , is an entire function in the complex plane,<sup>24</sup> and the Taylor series  $E_{\alpha}(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(1+n\alpha)}$  shows that  $E_1(z) = \exp(z)$ . The most important property of the model relaxation function (34) is that it decays asymptotically with a power law,

$$\phi_{\text{ML}}(t) \stackrel{t \rightarrow \infty}{\sim} \frac{(t/\tau)^{-\alpha}}{\Gamma(1-\alpha)}. \quad (35)$$

Inserting (34) into the generic form (31) of the symmetrized intermediate scattering leads to the model

$$F^{(+)}(t) = F(\infty) + (F(0) - F(\infty))E_{\alpha}(-(|t|/\tau)^{\alpha}), \quad (36)$$

which has *a priori* three parameters:

1. the time scale parameter,  $\tau$ ,
2. the form parameter,  $\alpha$ , and
3. the plateau value  $F(\infty) \equiv \lim_{t \rightarrow \infty} F^{(+)}(t)$ .

In this form the model has been used in recent publications,<sup>13–15</sup> noting that only incoherent scattering has been considered, where  $F(0) = 1$ . In the following this restriction will not be made, assuming that an estimation for  $F(0) \equiv F(\mathbf{q}, 0)$  can be provided according to Eq. (27).

The dynamic structure factor corresponding to the model (36) has then the form

$$S^{(+)}(\omega) = F(\infty)\delta(\omega) + (F(0) - F(\infty))\tilde{\phi}_{\text{ML}}(\omega), \quad (37)$$

where the Fourier transformed relaxation function is a “generalized Lorentzian,”<sup>25</sup>

$$\tilde{\phi}_{\text{ML}}(\omega) = \frac{\sin\left(\frac{\pi\alpha}{2}\right)}{\pi\omega\left((\tau\omega)^{-\alpha} + (\tau\omega)^{\alpha} + 2\cos\left(\frac{\pi\alpha}{2}\right)\right)}, \quad (38)$$

which follows from the even simpler analytical form of its Laplace transform,<sup>24</sup>

$$\hat{\phi}_{\text{ML}}(s) = \frac{1}{s(1 + (s\tau)^{-\alpha})}, \quad (39)$$

by using the identity  $\tilde{\phi}_{\text{ML}}(\omega) = \Re\{\hat{\phi}_{\text{ML}}(i\omega)\}$ . The Fourier spectrum (38) becomes a “normal” Lorentzian function for  $\alpha \rightarrow 1$ .

## 3. Choice of the model

The ML relaxation function has the remarkable property of being close to self-similar (“weakly self-similar”)<sup>26</sup> for every  $t > 0$  and the physical reason for using it as model relaxation function are developed in Ref. 27. We resume here the essential points. From a

mathematical point of view the ML relaxation verifies a fractional differential equation of the form<sup>24,28</sup>

$$\partial_t \phi_{\text{ML}}(t) + \tau^{-\alpha} \underbrace{\frac{d}{dt} \int_0^t dt' \frac{(t-t')^{\alpha-1}}{\Gamma(\alpha)} \phi_{\text{ML}}(t')}_{\partial_t^{1-\alpha} \phi_{\text{ML}}(t)} = 0, \quad (40)$$

where  $\partial_t^{1-\alpha}$  denotes a fractional derivative<sup>29</sup> of order  $1 - \alpha$ . From a physical point of view Eq. (40) can be considered as a special form of the general equation of motion,

$$\partial_t \phi^{(+)}(t) + \int_0^t dt' \kappa^{(+)}(t-t') \phi^{(+)}(t') = 0, \quad (41)$$

that any symmetric time correlation function fulfills according to the Mori–Zwanzig theory of the Generalized Langevin equation.<sup>30–32</sup> Here,  $\kappa^{(+)}(t)$  is the associated memory kernel, which is itself a time autocorrelation function that can be formally derived from the Hamiltonian of the dynamical system under consideration. The only point that matters here is to consider that the memory kernel has essentially two characteristic time scales,  $\tau$  and  $\tau^*$ , where  $\tau$  characterizes the asymptotic regime of the correlation function  $\phi^{(+)}(t)$  and  $\tau^*$  the transition to that regime. In Ref. 27, it is then shown that the ML relaxation function emerges whenever  $\tau^* \ll \tau$ . Writing  $\kappa^{(+)}(t) \equiv \kappa^{(+)}(t; \tau, \tau^*)$ , we have then

$$\kappa^{(+)}(t; \tau, \tau^*) \stackrel{\tau^* \rightarrow 0}{\sim} \phi_{\text{ML}}^{(+)}(t), \quad (42)$$

and the fractional derivative in Eq. (40) thus represents the asymptotic form of a memory kernel.

## 4. Pseudoelastic model contribution

The estimation of the plateau value of a function from experimental data with an instrument-limited finite time range is clearly a delicate task and it is desirable to be to have some consistency check in which experimental data are used. This can be achieved if the pseudoelastic contribution,  $\xi$ , can be efficiently corrected for the given model relaxation function, such that the estimation (30) can be integrated into the fitting procedure. For this purpose, we will assume that the resolution function is well represented by a Gaussian function,

$$\tilde{R}(\omega) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{\omega^2}{2\sigma^2}} \longleftrightarrow R(t) = e^{-\frac{\sigma^2 t^2}{2}}, \quad (43)$$

where  $\sigma$  is approximately the half width at half maximum (HWHM) of the instrument under consideration. Working with symmetrized neutron scattering spectra and the model relaxation function  $\phi_{\text{ML}}(t)$ , the resulting pseudoelastic contribution becomes then a function of the parameters  $\alpha$ ,  $\tau$ , and  $\sigma$ ,

$$\xi_{\text{ML}}(\tau, \alpha, \sigma) = \int_{-\epsilon}^{+\epsilon} d\omega (\tilde{R} * \tilde{\phi}_{\text{ML}})(\omega), \quad (44)$$

such that

$$F(\infty; \tau, \alpha, \sigma) \approx \frac{F_m(\infty) - \xi_{\text{ML}}(\tau, \alpha, \sigma)F_m(0)}{1 - \xi_{\text{ML}}(\tau, \alpha, \sigma)} \quad (45)$$

replaces the fit parameter  $F(\infty)$ . For a Gaussian function, we have  $\int_{-3\sigma}^{+3\sigma} d\omega \tilde{R}(\omega) \approx 0.9978$  such that  $\epsilon = 3\sigma$  is a good choice.

### C. Computing $\xi_{ML}$

In order to obtain a quasi-analytical formula for the pseudoelastic contribution  $\xi_{ML}$  defined in Eq. (44), we introduce the boxcar function

$$\tilde{W}(\omega) = \Theta\left(1 - \frac{|\omega|}{\epsilon}\right) \longleftrightarrow W(t) = \frac{2 \sin(\epsilon t)}{t}, \quad (46)$$

where  $\epsilon > 0$  is the cutoff frequency. With Parseval's theorem and the convolution theorem of the Fourier transform,

$$(\tilde{R} * \tilde{\phi}_{ML})(\omega) \longleftrightarrow R(t)\phi_{ML}(t), \quad (47)$$

we write in a first step

$$\begin{aligned} \xi_{ML} &= \int_{-\infty}^{+\infty} d\omega \tilde{W}(\omega)(\tilde{R} * \tilde{\phi}_{ML})(\omega) \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} dt W(t)R(t)\phi_{ML}(t). \end{aligned}$$

Defining the auxiliary function

$$H(t) \equiv W(t)R(t)\phi_{ML}(t), \quad (48)$$

and noting that all functions on the r.h.s. are even in time, it follows that

$$\xi_{ML} = \frac{1}{\pi} \int_0^{\infty} dt e^{-st} H(t) \Big|_{s=0}. \quad (49)$$

The pseudoelastic contribution  $\xi_{ML}$  may, thus, be written as Laplace transform of a product of three functions, evaluated at  $s = 0$ . We use now that for a pair of functions,  $f(t)$  et  $g(t)$ ,

$$\int_0^{\infty} dt e^{-st} f(t)g(t) = \frac{1}{2\pi i} \oint_C ds' \hat{f}(s-s')\hat{g}(s'), \quad (50)$$

where  $C$  encircles all singularities of the integrand, and use this formula in two steps:

1. Compute

$$\hat{\phi}_{ML}^{(R)}(s) = \frac{1}{2\pi i} \oint_C ds' \hat{\phi}_{ML}(s-s')\hat{R}(s'). \quad (51)$$

2. Compute

$$\xi_{ML} = \frac{1}{\pi} \left\{ \frac{1}{2\pi i} \oint_C ds' \hat{\phi}_{ML}^{(R)}(-s')\hat{W}(s') \right\}. \quad (52)$$

Here, an analytical form is known for the Laplace transformed model relaxation function,  $\hat{\phi}_{ML}(s)$  [see Eq. (39)], for the Laplace transform of the model resolution time window,

$$\hat{R}(s) = \frac{1}{\sigma} \sqrt{\frac{\pi}{2}} e^{\frac{s^2}{2\sigma^2}} \operatorname{erfc}\left(\frac{s}{\sqrt{2}\sigma}\right), \quad (53)$$

and for the Laplace transform of  $W(t)$ ,

$$\hat{W}(s) = 2 \operatorname{arccot}\left(\frac{s}{\epsilon}\right). \quad (54)$$

An analytical form of the contour integrals (51) and (52) cannot be found, but a good quasi-analytical approximation can be obtained by replacing  $\hat{R}(s)$  and  $\hat{W}(s)$  by Padé approximants,<sup>33</sup>

$$\hat{R}(s) \approx \frac{1}{\sigma} \frac{P(s/\sigma)}{Q(s/\sigma)}, \quad (55)$$

$$\hat{W}(s) \approx \frac{P'(s/\epsilon)}{Q'(s/\epsilon)}, \quad (56)$$

where  $P(\cdot)$ ,  $Q(\cdot)$ ,  $P'(\cdot)$ ,  $Q'(\cdot)$  are polynomials. We note here that the method has been recently used to compute a good approximation for the resolution-broadened Fourier transform  $\hat{\phi}_{ML}^{(R)}(\omega)$ .<sup>14</sup> Introducing appropriately scaled integration variables, the contour integrals (51) and (52) may then be evaluated by the residue theorem of complex analysis. The details are described in the Appendix, and the result is

$$\xi_{ML} \approx \frac{1}{\pi} \sum_{j=1}^m \sum_{k=1}^{m'} \chi \lambda c_j d_k \hat{\Phi}_{ML}(-\chi(u_j + \lambda v_k)), \quad (57)$$

where  $\hat{\Phi}_{ML}(\cdot)$  is the scale-free version of the Laplace-transformed relaxation function  $\hat{\phi}_{ML}(\cdot)$ ,

$$\hat{\Phi}_{ML}(u) = \frac{1}{u(1+u^{-\alpha})}, \quad (58)$$

$\chi$  and  $\lambda$  are the dimensionless scaling parameters,

$$\chi = \sigma\tau \quad \text{and} \quad \lambda = \frac{\epsilon}{\sigma}, \quad (59)$$

and  $\{u_j\}$  and  $\{v_k\}$  are the roots of the polynomials  $Q(u)$  and  $Q'(v)$ , respectively. The coefficients

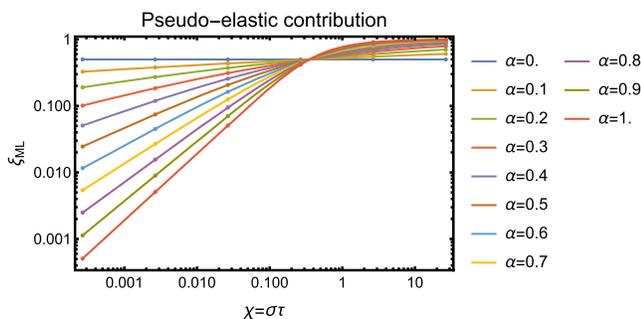
$$c_j = \frac{P(u_j)}{\prod_{k=1, k \neq j}^m (u_j - u_k)}, \quad (60)$$

$$d_k = \frac{P'(v_k)}{\prod_{l=1, l \neq k}^{m'} (v_k - v_l)} \quad (61)$$

are the residues of the dimensionless expressions  $P(u)/Q(u)$  and  $P'(v)/Q'(v)$ , evaluated at the respective roots of the denominator polynomials. We have, thus,  $Q(u_j) = 0$  and  $Q'(v_k) = 0$ . The final expression (57) for  $\xi_{ML}$  is, thus, the linear superposition of simple terms of the form (58). Coding  $\xi_{ML}$  as a compiled function leads to sufficiently short execution times, which allow for integrating this function into a fitting procedure.

### D. Numerical test

To obtain a systematic picture of the pseudoelastic contribution as a function of  $\tau$  and  $\alpha$ , we compute it according to Eq. (57) for  $\alpha = k \times 0.1$ ,  $k = 0, \dots, 10$ . In view of later applications, we define  $\sigma$  to be the resolution (HWHM) of the IN16B spectrometer,  $\sigma = 1.75 \mu\text{eV}$ , and we set  $\epsilon = 3\sigma$ . We vary then  $\chi \in [\sigma\tau_{\min}, \sigma\tau_{\max}]$ , where  $\tau_{\min} = 0.1 \text{ ps}$  and  $\tau_{\max} = 10^4 \text{ ps}$  (solid lines). For comparison, we compute  $\xi_{ML}$  by numerical integration of  $\hat{\phi}_{ML}^{(R)}(\omega)$ , choosing the same values for  $\alpha$ ,  $\sigma$ , and  $\epsilon$  and fixing  $\tau$  to the



**FIG. 1.** Pseudoelastic contribution  $\xi_{ML}$  as a function of  $\chi \equiv \sigma\tau$  and  $\alpha$ . The solid lines correspond to the calculation according to Eq. (57) and the points to control calculations by numerical integration. More explanations are given in the text.

discrete values  $\tau = 10^j$  ps,  $j = -1, \dots, 4$  (points). Defining the matrices  $\xi_{ML} \equiv (\xi_{ML}[j, k])$  and  $\xi_{ML}^{(n.i.)} \equiv (\xi_{ML}^{(n.i.)}[j, k])$ , where “n.i.” stands for “numerical integration,” we find that

$$\|\xi_{ML}^{(n.i.)} - \xi_{ML}\| = 3.45 \times 10^{-6},$$

where  $\|\dots\|$  is defined as the maximum singular value of the difference matrix. The results of the calculations are shown in Fig. 1. One observes that for “good resolutions,” where  $\sigma\tau < 1$ , the pseudoelastic contributions increase with decreasing  $\alpha$ , and the opposite is true for “bad resolutions,” where  $\sigma\tau > 1$ . We note that  $\lim_{\alpha \rightarrow 0} \phi_{ML}(t) = 1/2$  for any  $t > 0$ , which explains the results for  $\alpha = 0$ . All calculations have been performed with Padé-approximations of order  $m = 8$  for the denominator polynomials  $Q(u)$  and  $Q'(u)$  and order  $n = 8$  for the corresponding numerator polynomials,  $P(u)$  and  $P'(u)$ , choosing  $s = 1$  as the reference point. Constructing the Padé-approximant for the resolution function through

$$\tilde{R}_{\text{Padé}}(\omega) \equiv \frac{1}{\pi} \Re \left\{ \frac{P(i\omega)}{Q(i\omega)} \right\},$$

we find that

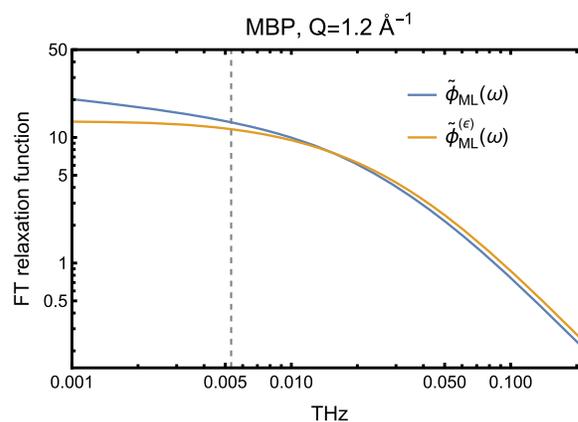
$$|\tilde{R}_{\text{Padé}}(\omega) - \tilde{R}(\omega)| < 10^{-6}$$

in the relevant  $\omega$ -domain. All computations have been performed with the Wolfram Mathematica software.<sup>34</sup>

## IV. APPLICATIONS

### A. QENS analysis of Myelin Basic Protein

To illustrate the pseudoelastic contribution to elastic scattering we consider now a concrete example related to a recently published QENS study of Myelin Basic Protein (MBP) in an aqueous solution.<sup>15</sup> Myelin Basic Protein is an elementary constituent of the myelin sheath of nerves and in aqueous solution, it is an intrinsically disordered protein (IDP). The incoherent QENS spectra for the study cited above have been recorded on the new IN16B spectrometer of the Institut Laue-Langevin, using the BATS option (Backscattering And Time-of-flight Spectrometer) with an instrumental resolution (FWHM) of  $3.5 \mu\text{eV}$ . The translational diffusion constant,  $D$ , of MBP was measured separately by dynamic light

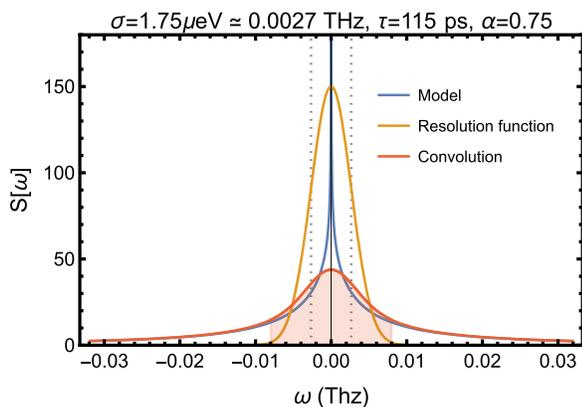


**FIG. 2.** Impact of global diffusion on the Fourier spectrum of the model relaxation function for MBP in  $D_2O$  buffer at  $T = 283$  K and  $q = 1.2/\text{\AA}$ . The solid blue line labels the relaxation function and the yellow line the corresponding function with the diffusion damping factor, where  $D = 3.3 \text{\AA}^2/\text{ns}$  from DLS. The vertical dashed indicates the instrumental resolution (FWHM).

scattering (DLS) and was then injected into the fit, writing  $F^{(+)}(q, t) = \exp(-Dq^2|t|)F_{\text{int}}^{(+)}(q, t)$ , where  $F_{\text{int}}^{(+)}(q, t)$  is the symmetrized intermediate scattering function for internal motions, the generic form of which is given by Eq. (36). The implicit assumption is here that global and internal motions are not correlated. As mentioned in Ref. 15, the resulting fits for EISF,  $\alpha$ , and  $\tau$  vary only a little if the diffusion constant is simply neglected. This is illustrated in Fig. 2, which shows a log-log plot of the Fourier spectrum of the fitted model relaxation function for MBP in  $D_2O$  buffer ( $T = 283$  K,  $q = 1.2/\text{\AA}$ ) for the dynamical range of the instrument, together with the corresponding diffusion-broadened counterpart resulting from the damping factor  $\exp(-Dq^2|t|)$  of the intermediate scattering function. We take here  $D = 3.3 \text{\AA}^2/\text{ns}$  from DLS. Having this figure in mind, the resolution-deconvolved intermediate scattering function can, therefore, be fitted directly with the model (36). It is worthwhile mentioning that more sophisticated treatments of global protein motions have been developed,<sup>35–38</sup> but we consider that our estimation, which is exact for small  $q$ -values, is sufficient to estimate their impact.

The impact of pseudoelastic scattering on the observed elastic intensities is illustrated in Fig. 3, which shows again the Fourier spectrum of the fitted model relaxation function  $\tilde{\phi}_{ML}(\omega)$  for the same parameters as in Fig. 2 (blue line), together with the model resolution function (yellow line), where the instrumental resolution (HWHM) corresponds to  $\sigma = 0.0027$  THz, and the resulting resolution-broadened spectrum,  $\tilde{\phi}_{ML}^{(R)}(\omega)$  (red line). The (dimensionless) area in light red is the corresponding pseudoelastic contribution, which is for this example  $\xi_{ML} \approx 0.47$  with  $\epsilon = 3\sigma$ . The difference between the model spectrum and its resolution-broadened version should also be noticed.

An important result of the study in Ref. 15 was that the fitted EISF vanishes. This can be explained by the fact that MBP in solution is a very flexible molecule, such that  $\langle \exp(i\mathbf{q} \cdot \mathbf{R}_j) \rangle \approx 0$ . In the Gaussian approximation<sup>39,40</sup> (in  $q \equiv |\mathbf{q}|$ ) of the elastic amplitude one would write



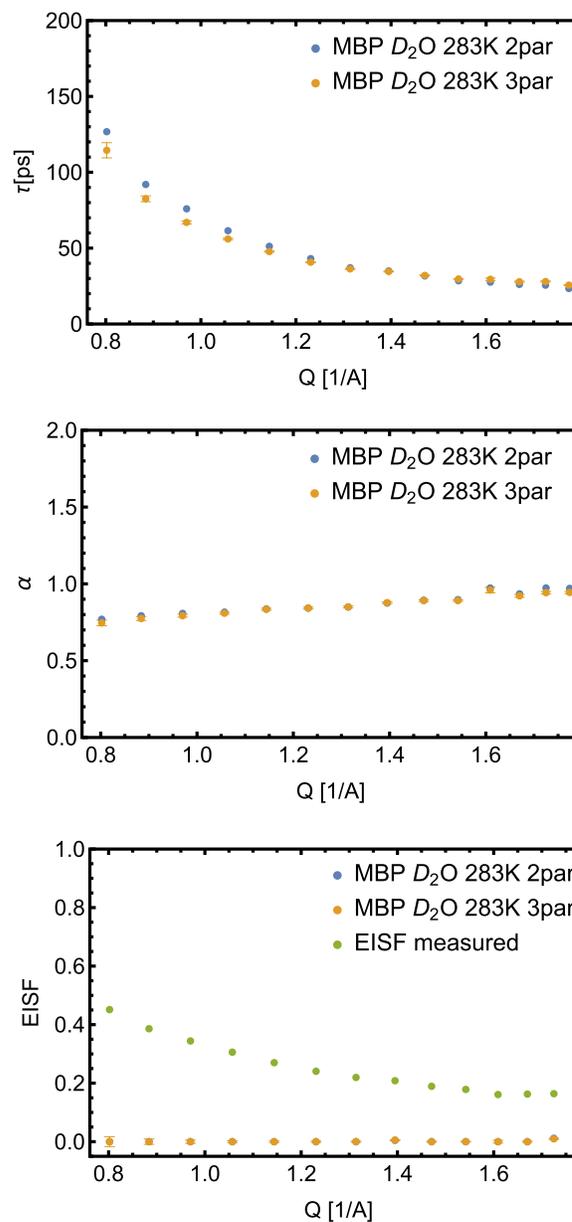
**FIG. 3.** Model relaxation function  $\tilde{\phi}_{\text{ML}}(\omega)$  (blue line), model resolution function (yellow line) and the resulting convolution  $\tilde{\phi}_{\text{ML}}^{(R)}(\omega)$  (red line) for the resolution of the IN16B spectrometer at ILL. The values for  $\tau$  and  $\alpha$  are taken from the fit parameters shown in Fig. 4 (three-parameter fit for  $|\mathbf{q}| = 0.8/\text{\AA}^{-1}$ ). The vertical dotted lines indicate the instrumental resolution and the light red area corresponds to the pseudoelastic contribution of  $\xi_{\text{ML}} \approx 0.47$  for this example.

$$\text{EISF}(\mathbf{q}) \stackrel{|\mathbf{q}| \rightarrow 0}{\sim} \frac{1}{N} \sum_{j \in \mathbb{H}} e^{-\frac{1}{3}|\mathbf{q}|^2 \langle \hat{u}_j^2 \rangle} \approx 0, \quad (62)$$

where  $\langle \hat{u}_j^2 \rangle \equiv \langle (\hat{\mathbf{R}}_j - \langle \hat{\mathbf{R}}_j \rangle)^2 \rangle$  is the mean square position fluctuation of (hydrogen) atom  $j$ . For smaller  $q$ -values a vanishing EISF stands thus for large motional amplitudes of the atoms. Keeping in mind that the EISF is a “theoretical quantity,”  $\text{EISF}(\mathbf{q}) = F_{\text{inc}}(\mathbf{q}, \infty)$ , which can only be determined by assuming a model, we can now check the fits of the three-parameter model (36) with a two-parameter fit, where the EISF is eliminated according to Eq. (27). The results in Fig. 4 show that  $\alpha$  and  $\tau$  change only slightly comparing the two- and three-parameter fits, where the decrease of  $\alpha$  with  $q$  indicates that increasingly slower relaxation modes are mixed in with decreasing spatial resolution, making the relaxation function less exponential. The vanishing EISF from the three-parameter fits is, in particular, confirmed if one considers the relevant scale for this quantity. The measured elastic intensity (bottom panel, green points) is this entirely determined by the “pseudoelastic contribution” defined in Eq. (57). For its calculation, we used the same Padé approximations as for the numerical test described in Sec. III D. We note here that the standard estimation of the parameter errors for the three-parameter fits (yellow dots) leads to error bars, which are sometimes hardly visible in the plots and cannot be performed for the two-parameter fits. The reason is purely technical—namely that in the latter case the EISF “parameter” is a compiled function, which is passed as an argument to the fit routine of the Mathematica software.<sup>34</sup> Concerning this point, we think that the error of the fit parameters is anyway better estimated by comparing the results of the three- and two-parameter fits.

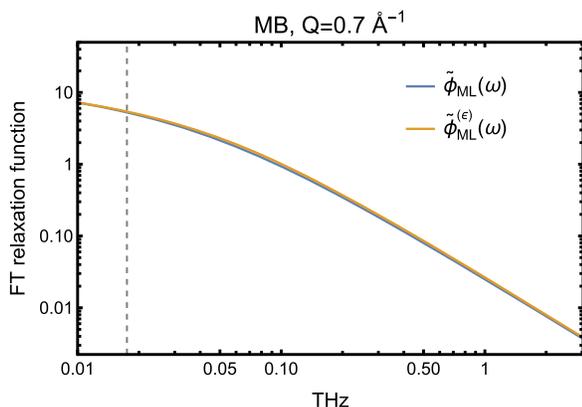
## B. QENS analysis of myoglobin

To show that the model (36) does not systematically lead to vanishing EISFs for proteins in solution, we have analyzed QENS data from apo-myoglobin at  $T = 284$  K in  $\text{D}_2\text{O}$  buffer ( $\text{pD} = 6$ ), which have been recorded on the IN5 spectrometer at the Institut



**FIG. 4.** From top to bottom: Characteristic time scale,  $\tau$ , form parameter,  $\alpha$ , and elastic scattering amplitude,  $\text{EISF}$ , for incoherent QENS from MBP in solution. Blue dots correspond to a two-parameter fit, where  $\text{EISF} \equiv \text{EISF}(\tau, \alpha, \sigma)$  according to Eq. (45) and yellow points to a three-parameter fit. Concerning the EISF, the green points correspond to the measured elastic intensity  $F_m(\infty) \equiv F_m(\mathbf{q}, \infty)$  defined in Eq. (25), and on the scale of the plot the results in two- and three-parameter plots are indistinguishable.

Laue-Langevin in Grenoble, at a resolution (FWHM) of  $11.6 \mu\text{eV}$ .<sup>41</sup> In contrast to MBP, myoglobin is a compactly folded globular protein of about the same weight, but with a well-defined three-dimensional structure containing eight  $\alpha$ -helices as secondary structure elements. Figure 5 shows that here, as for MBP, global diffusion can be neglected for the analysis of the QENS spectra. We insert here



**FIG. 5.** Impact of global diffusion on the model dynamic structure factor at  $q = 0.7/\text{Å}$  for myoglobin in solution. The legends are the same as in Fig. 2.

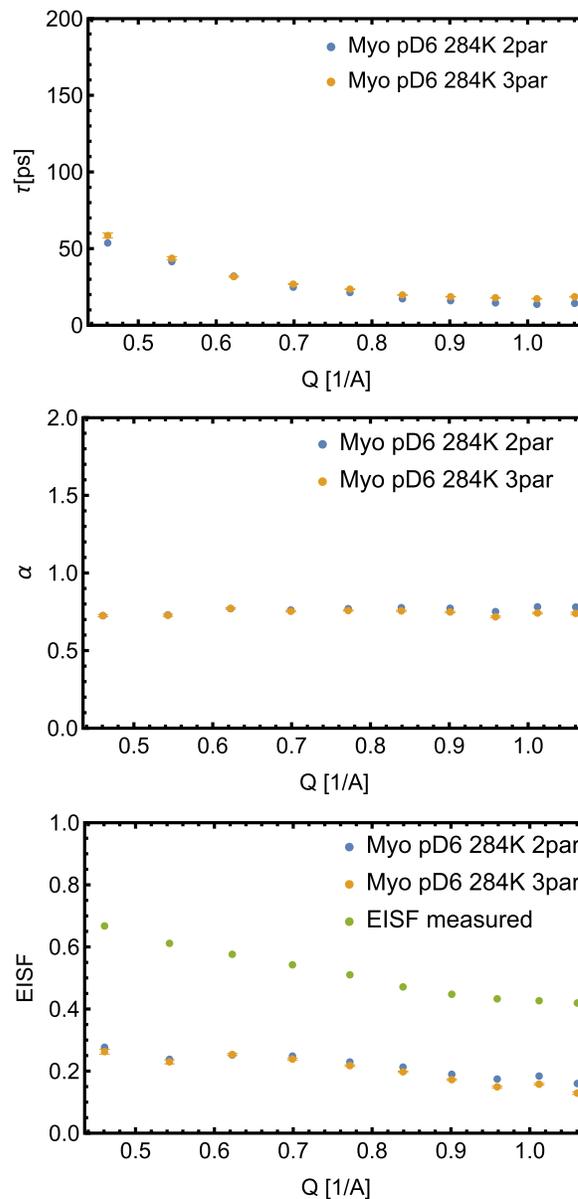
a global diffusion coefficient of  $D = 10.1 \text{ Å}^2/\text{ns}$ , which is the mean of the values reported in Refs. 42 and 43, respectively, but the exact value is here not important. We note in this context that a somewhat smaller value of  $D \approx 8 \text{ Å}^2/\text{ns}$  is obtained by a very rough estimation from the Stokes-Einstein relation,  $D = k_B T / 6\pi\eta R_h$ , inserting here a hydrodynamic radius of  $R_h = 2 \text{ nm}$  for a myoglobin molecule and for  $\eta$  the kinematic viscosity of water at 284 K ( $\eta = 1.306 \text{ mPa s}$ ).<sup>44</sup> The hydrodynamic radius is here calculated from PDB structure 1BVC for apo-myoglobin, assuming the latter to be a sphere.

Since impact of global diffusion can be neglected, model (36) has again been fitted directly to the resolution-deconvolved intermediate scattering function. Figure 6, bottom panel, shows that the EISF for myoglobin is clearly non-vanishing and that both the three- and the two-parameter fits give again similar results. The fact that the atomic motions in a globular, compactly folded protein are more hindered than in a polymer-like intrinsically disordered protein like MBP is thus clearly reflected in the corresponding EISFs, and it should be noted that this effect is much more pronounced for the resolution-corrected elastic intensities than for the measured ones. It should also be noted that the characteristic time scale,  $\tau$ , is systematically smaller compared with MBP, reflecting faster localized motions of the atoms in the more compactly folded myoglobin molecule, and that the  $\alpha$ -values are very similar compared with MBP in the common  $q$ -interval ( $[0.8/\text{Å} < q < 1.06/\text{Å}]$ ).

### C. Resolution correction for $\langle \dot{\mathbf{u}}^2 \rangle$

The Gaussian approximation (62) of the EISF is also known as Debye-Waller factor and has been used for decades to analyze “elastic scans” of incoherent neutron scattering from  $\text{D}_2\text{O}$ -hydrated protein powders in the low  $q$ -region, in order to infer the average mean square position fluctuations of the (hydrogen) atoms in the protein from these data. There is a large bulk of literature on that subject and we cite here only Refs. 7, 45, and 46. It follows then from Eqs. (30) and (62) that the “true,” resolution-corrected mean square position fluctuation, averaged over all atoms, is given by

$$\overline{\langle \dot{\mathbf{u}}^2 \rangle} \stackrel{q \rightarrow 0}{\approx} -\frac{3}{q^2} \log \left( \frac{\text{EISF}_m(q) - \xi(q)}{1 - \xi(q)} \right). \quad (63)$$



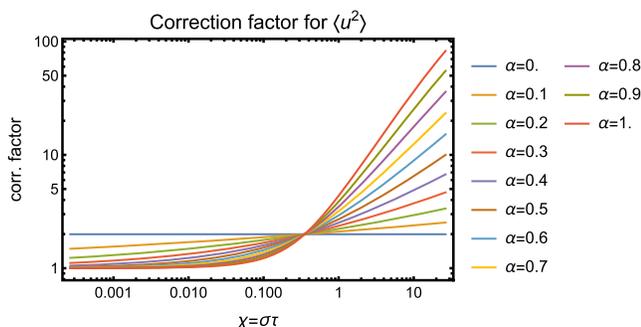
**FIG. 6.** The same quantities as in Fig. 6 for myoglobin in solution. More details are given in the text.

We assume here an isotropic sample and indicate explicitly the dependency of all involved quantities on  $q \equiv |\mathbf{q}|$ . Writing  $\text{EISF}_m(q) = \exp(-q^2 \overline{\langle \dot{\mathbf{u}}^2 \rangle}_m / 3)$  and developing Expression (63) into a power series in  $\xi(q)$ , we obtain

$$\overline{\langle \dot{\mathbf{u}}^2 \rangle} \approx \overline{\langle \dot{\mathbf{u}}^2 \rangle}_m + \sum_{n=1}^{\infty} \frac{3}{nq^2} \left( e^{\frac{n}{3} q^2 \overline{\langle \dot{\mathbf{u}}^2 \rangle}_m} - 1 \right) \xi(q)^n$$

$$\stackrel{q \rightarrow 0}{\approx} \overline{\langle \dot{\mathbf{u}}^2 \rangle}_m \sum_{n=0}^{\infty} \xi(q)^n,$$

where the geometrical series can be summed up to give



**FIG. 7.** Correction factor to convert the measured mean square position fluctuations into the resolution-corrected ones. More explanations are given in the text.

$$\overline{\langle \dot{\mathbf{u}}^2 \rangle} \stackrel{q \rightarrow 0}{\approx} \frac{\overline{\langle \dot{\mathbf{u}}^2 \rangle}_m}{1 - \xi(q)}. \quad (64)$$

Figure 7 shows the correction factor  $1/(1 - \xi)$ , which needs to be applied to obtain the “true,” resolution-corrected mean square position fluctuations from the measured ones. At a first glance, the large correction factors for “bad resolutions” and close to exponential relaxation may surprise, but they can be understood noting that in this case, an elastic scan contains practically the whole integral over the quasielastic line such that  $\xi \lesssim 1$ .

## V. CONCLUSIONS

In this paper, we have presented a quasi-analytical method for computing the pseudoelastic contribution of quasielastic scattering to the elastic neutron scattering amplitude, which enables the estimation of the true elastic scattering amplitude for the Mittag-Leffler relaxation model. Due to the computational efficiency of the method, which is based on Padé approximants for the Laplace transformed relaxation and resolution functions, the true elastic scattering amplitude can be eliminated in the model (36) by using the measured elastic intensity and the two parameters  $\alpha$  and  $\tau$  of the ML relaxation function together with the instrument resolution,  $\sigma$ , as input. We have applied the method to confirm the results of a recent analysis of incoherent QENS spectra from the internal dynamics of Myelin Basic Protein in solution,<sup>15</sup> which revealed, in particular, a vanishing EISF. MBP is an intrinsically disordered protein and to demonstrate that the vanishing EISF is the result of its “floppiness” and the corresponding large motional amplitudes of the atoms, we performed a comparative study for myoglobin in solution. Myoglobin has about the same weight as MBP, but in contrast to the latter, it is a globular protein with a well defined structure. As one would expect, we find here a clearly non-vanishing EISF as a result of the more hindered atomic motions in a compactly folded protein. An important result in this context is that the resolution-corrected EISF shows this result much more clearly than the measured one. We have also shown that the resolution corrections of the elastic intensity may also strongly impact atomic mean square position fluctuations, which are routinely extracted from so-called elastic scans.

It should be kept in mind that the elastic intensity is a “theoretical quantity,” which can only be extracted from experimental QENS spectra by assuming a model for the relaxation dynamics of the protein under consideration. In that light, our analysis is a consistency check for the ML relaxation model, which does, in principle, not exclude other models for the relaxation function.

We finally emphasize that our method is prepared to deal with a coherent scattering, in general. This aspect is, for instance, important for the increasing number of QENS studies of proteins in deuterated aqueous solutions ( $D_2O$ -buffer), where coherent scattering from the solvent becomes visible at higher momentum transfers  $\gtrsim 1.5 \text{ \AA}^{-1}$ , if the solvent contribution is not subtracted.

## ACKNOWLEDGMENTS

A.N.H. acknowledges financial support from the Région Centre-Val de Loire, France, and the Jülich Center for Neutron Science, Forschungszentrum Jülich, Germany.

## AUTHOR DECLARATIONS

### Conflict of Interest

The authors have no conflicts to disclose.

### Author Contributions

**Abir N. Hassani:** Investigation (equal); Validation (lead); Visualization (lead); Writing – original draft (supporting). **Andreas M. Stadler:** Resources (lead); Writing – original draft (supporting). **Gerald R. Kneller:** Conceptualization (lead); Methodology (lead); Software (equal); Supervision (lead); Writing – original draft (equal).

## DATA AVAILABILITY

The analysis code is available at <https://doi.org/10.5281/zenodo.7058345> and the corresponding prepared input data for myoglobin at <https://doi.org/10.5281/zenodo.7108355>. The original neutron scattering data for Myelin Basic Protein are available at <https://doi.ill.fr/10.5291/ILL-DATA.8-04-874>.

## APPENDIX: DERIVATION OF EXPRESSION (57)

We start from Eq. (51) and insert Eq. (55) for the Padé approximation of the Laplace transformed instrumental time window. In addition, we use that

$$\hat{\phi}_{ML}(s) = \tau \hat{\Phi}_{ML}(s\tau)$$

to obtain in a first step

$$\begin{aligned} \hat{\phi}_{ML}^{(R)}(s) &= \frac{1}{2\pi i} \oint_C \frac{ds'}{\sigma} \hat{\phi}_{ML}(s-s') \frac{P(s'/\sigma)}{Q(s'/\sigma)} \\ &\stackrel{s'/\sigma \rightarrow u}{=} \frac{1}{2\pi i} \oint_C du \tau \hat{\Phi}_{ML}(\tau(s-\sigma u)) \frac{P(u)}{Q(u)} \\ &= \frac{1}{\sigma} \left\{ \frac{1}{2\pi i} \oint_C du \chi \hat{\Phi}_{ML} \left( \chi \left( \frac{s}{\sigma} - u \right) \right) \frac{P(u)}{Q(u)} \right\}. \end{aligned}$$

Writing  $Q(u) = \prod_{k=1}^m (u - u_k)$  and defining the coefficients

$$c_j = \frac{P(u_j)}{\prod_{k=1, k \neq j}^m (u_j - u_k)},$$

it follows then from the residue theorem of complex analysis that

$$\hat{\phi}_{ML}^{(R)}(s) = \frac{1}{\sigma} \sum_{j=1}^m \chi c_j \hat{\Phi} \left( \chi \left( \frac{s}{\sigma} - u_j \right) \right). \quad (\text{A1})$$

This expression is now inserted into the contour integral (52), and appropriate changes of the integration variables lead to

$$\begin{aligned} \xi_{ML} &= \frac{1}{\pi} \left\{ \frac{1}{2\pi i} \oint_C ds' \left\{ \frac{1}{\sigma} \sum_{j=1}^m \chi c_j \hat{\Phi} \left( \chi \left( -\frac{s'}{\sigma} - u_j \right) \right) \right\} \frac{P'(s'/\epsilon)}{Q'(s'/\epsilon)} \right\}, \\ &\stackrel{s'/\sigma \rightarrow u}{=} \frac{1}{\pi} \left\{ \frac{1}{2\pi i} \oint_C du \left\{ \sum_{j=1}^m \chi c_j \hat{\Phi}(\chi(-u - u_j)) \right\} \frac{P'(u\sigma/\epsilon)}{Q'(u\sigma/\epsilon)} \right\}, \\ &\stackrel{\epsilon/\sigma \rightarrow \lambda, u/\lambda \rightarrow v}{=} \frac{1}{\pi} \left\{ \frac{1}{2\pi i} \oint_C dv \left\{ \sum_{j=1}^m \lambda \chi c_j \hat{\Phi}(\chi(-\lambda v - u_j)) \right\} \frac{P'(v)}{Q'(v)} \right\}. \end{aligned}$$

Writing  $Q'(v) = \prod_{l=1}^{m'} (v - v_l)$  and defining the coefficients

$$d_k = \frac{P(u_k)}{\prod_{l=1, l \neq k}^{m'} (v_k - v_l)},$$

we get again from the residue theorem the final result

$$\xi_{ML} = \frac{1}{\pi} \sum_{j=1}^m \sum_{k=1}^{m'} \lambda \chi c_j d_k \hat{\Phi}(-\chi(u_j + \lambda v_k)), \quad (\text{A2})$$

where  $\chi \equiv \sigma\tau$  and  $\lambda \equiv \epsilon/\sigma$ .

## REFERENCES

- 1 S. Lovesey, *Theory of Neutron Scattering from Condensed Matter* (Clarendon Press, Oxford, 1984), Vol. I.
- 2 *Neutron Scattering in Biology: Techniques and Applications*, Biological and Medical Physics, Biomedical Engineering, edited by J. Fitter, T. Gutberlet, and J. Katsaras (Springer, Berlin; New York, 2006), Chaps. 15 and 16, ISBN: 978-3-540-29108-4.
- 3 V. G. Sakai and A. Arbe, *Curr. Opin. Colloid Interface Sci.* **14**, 381 (2009).
- 4 J. Colmenero and A. Arbe, *J. Polym. Sci., Part B: Polym. Phys.* **51**, 87 (2013).
- 5 R. Ashkar, H. Z. Bilheux, H. Bordallo, R. Briber, D. J. E. Callaway, X. Cheng, X.-Q. Chu, J. E. Curtis, M. Dadmun, P. Fenimore *et al.*, *Acta Crystallogr., Sect. D: Struct. Biol.* **74**, 1129 (2018).
- 6 J. C. Smith, P. Tan, L. Petridis, and L. Hong, *Annu. Rev. Biophys.* **47**, 335 (2018).
- 7 W. Doster, S. Cusack, and W. Petry, *Nature* **337**, 754 (1989).
- 8 I. E. T. Iben, D. Braunstein, W. Doster, H. Frauenfelder, M. K. Hong, J. B. Johnson, S. Luck, P. Ormos, A. Schulte, P. J. Steinbach *et al.*, *Phys. Rev. Lett.* **62**, 1916 (1989).
- 9 H. Frauenfelder, S. G. Sligar, and P. G. Wolynes, *Science* **254**, 1598 (1991).
- 10 P. W. Fenimore, H. Frauenfelder, B. H. McMahon, and F. G. Parak, *Proc. Natl. Acad. Sci. U. S. A.* **99**, 16047 (2002).
- 11 G. R. Kneller and V. Calandrini, *J. Chem. Phys.* **126**, 125107 (2007).
- 12 G. R. Kneller, *Proc. Natl. Acad. Sci. U. S. A.* **115**, 9450 (2018).
- 13 M. Saouessi, J. Peters, and G. R. Kneller, *J. Chem. Phys.* **150**, 161104 (2019).
- 14 M. Saouessi, J. Peters, and G. R. Kneller, *J. Chem. Phys.* **151**, 125103 (2019).
- 15 A. N. Hassani, L. Haris, M. Appel, T. Seydel, A. M. Stadler, and G. R. Kneller, *J. Chem. Phys.* **156**, 025102 (2022).
- 16 M. H. Petersen, N. Vernet, W. P. Gates, F. J. Villacorta, S. Ohira-Kawamura, Y. Kawakita, M. Arai, G. Kneller, and H. N. Bordallo, *J. Phys. Chem. C* **125**, 15085 (2021).
- 17 M. Bée, *Quasielastic Neutron Scattering: Principles and Applications in Solid State Chemistry, Biology and Materials Science* (Adam Hilger, Bristol, 1988).
- 18 P. Gerlach, O. Schärpf, W. Prandl, and B. Dorner, *J. Phys. Colloq.* **43**, C7 (1982).
- 19 A. M. Gaspar, S. Busch, M.-S. Appavou, W. Haeussler, R. Georgii, Y. Su, and W. Doster, *Biochim. Biophys. Acta, Proteins Proteomics* **1804**, 76 (2010).
- 20 T. Burankova, R. Hempelmann, A. Wildes, and J. P. Embs, *J. Phys. Chem. B* **118**, 14452 (2014).
- 21 *Neutron Scattering in Biology: Techniques and Applications*, Biological and Medical Physics, Biomedical Engineering, edited by J. Fitter, T. Gutberlet, and J. Katsaras (Springer, Berlin; New York, 2006), Chap. 15, ISBN: 978-3-540-29108-4.
- 22 Q. Berrod, K. Lagrené, J. Ollivier, and J.-M. Zanotti, *EPJ Web Conf.* **188**, 05001 (2018).
- 23 P. Schofield, *Phys. Rev. Lett.* **4**, 239 (1960).
- 24 *Mittag-Leffler Functions, Related Topics and Applications*, Springer Monographs in Mathematics, edited by R. Gorenflo, A. A. Kilbas, F. Mainardi, and S. V. Rogosin (Springer, Heidelberg, 2014), ISBN: 978-3-662-43929-6; 978-3-662-43930-2.
- 25 G. R. Kneller, *Phys. Chem. Chem. Phys.* **7**, 2641 (2005).
- 26 G. R. Kneller and M. Saouessi, *J. Phys. A: Math. Theor.* **53**, 20LT01 (2020).
- 27 G. R. Kneller and M. Saouessi, *Acta Phys. Pol., B* **53**(2), 2 (2022).
- 28 W. G. Glöckle and T. F. Nonnenmacher, *Biophys. J.* **68**, 46 (1995).
- 29 K. Oldham and J. Spanier, *The Fractional Calculus* (Academic Press, New York, London, 1974).
- 30 R. Zwanzig, *Statistical Mechanics of Irreversibility*, Lectures in Theoretical Physics (Wiley-Interscience, New York, 1961), pp. 106–141.
- 31 H. Mori, *Prog. Theor. Phys.* **33**, 423 (1965).
- 32 R. Zwanzig, *Nonequilibrium Statistical Mechanics* (Oxford University Press, 2001).
- 33 F. W. J. Olver, D. W. Lozier, and R. F. Boisvert, *NIST Handbook of Mathematical Functions* (Cambridge University Press, 2010), ISBN: 978-0-521-14063-8.
- 34 Wolfram Research, Inc., *Mathematica*, Version 13.0, Champaign, IL, 2019, code available at <https://doi.org/10.5281/zenodo.7058345> and corresponding data under <https://doi.org/10.5281/zenodo.7108355>.
- 35 J. Pérez, J.-M. Zanotti, and D. Durand, *Biophys. J.* **77**, 454 (1999).
- 36 A. M. Stadler, I. Digel, G. M. Artmann, J. P. Embs, G. Zaccai, and G. Büldt, *Biophys. J.* **95**, 5449 (2008).
- 37 F. Roosen-Runge, M. Hennig, F. Zhang, R. M. J. Jacobs, M. Sztucki, H. Schober, T. Seydel, and F. Schreiber, *Proc. Natl. Acad. Sci. U. S. A.* **108**, 11815 (2011).
- 38 M. Grimaldo, F. Roosen-Runge, F. Zhang, F. Schreiber, and T. Seydel, *Q. Rev. Biophys.* **52**, e7 (2019).
- 39 A. Rahman, K. S. Singwi, and A. Sjölander, *Phys. Rev.* **126**, 986 (1962).
- 40 G. R. Kneller, *J. Chem. Phys.* **145**, 044103 (2016).
- 41 A. M. Stadler, F. Demmel, J. Ollivier, and T. Seydel, *Phys. Chem. Chem. Phys.* **18**, 21527 (2016).
- 42 V. Riveros-Moreno and J. B. Wittenberg, *J. Biol. Chem.* **247**, 895 (1972).
- 43 M. Rocco and O. Byron, *Eur. Biophys. J.* **44**, 417 (2015).
- 44 M. L. Huber, R. A. Perkins, A. Laescke, D. G. Friend, J. V. Sengers, M. J. Assael, I. N. Metaxa, E. Vogel, R. Mareš, and K. Miyagawa, *J. Phys. Chem. Ref. Data* **38**, 101 (2009).
- 45 G. Zaccai, *Science* **288**, 1604 (2000).
- 46 *Neutron Scattering in Biology: Techniques and Applications*, Biological and Medical Physics, Biomedical Engineering, edited by J. Fitter, T. Gutberlet, and J. Katsaras (Springer, Berlin; New York, 2006), Chap. 18, ISBN: 978-3-540-29108-4.

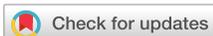
### **3. Signature of functional enzyme dynamics in quasielastic neutron scattering spectra : The case of phosphoglycerate kinase**

Cet article concerne une analyse des spectres de diffusion quasi-élastique de neutrons à haute résolution de la Phosphoglycérate kinase PGK qui élucide l'influence de l'activité enzymatique sur la dynamique de la protéine. Il a été publié en tant que communication dans le "Journal of Chemical Physics".

RAPID COMMUNICATION | OCTOBER 11 2023

## Signature of functional enzyme dynamics in quasielastic neutron scattering spectra: The case of phosphoglycerate kinase

Abir N. Hassani ; Luman Haris ; Markus Appel ; Tilo Seydel ; Andreas M. Stadler ; Gerald R. Kneller  



*J. Chem. Phys.* 159, 141102 (2023)

<https://doi.org/10.1063/5.0166124>



CrossMark

### Articles You May Be Interested In

A functional integral formalism for quantum spin systems

*J. Math. Phys.* (July 2008)

Modes selection in polymer mixtures undergoing phase separation by photochemical reactions

*Chaos* (June 1999)

Spreading of a surfactant monolayer on a thin liquid film: Onset and evolution of digitated structures

*Chaos* (March 1999)

11 October 2023 13:24:09

500 kHz or 8.5 GHz?  
And all the ranges in between.

Lock-in Amplifiers for your periodic signal measurements



Find out more



# Signature of functional enzyme dynamics in quasielastic neutron scattering spectra: The case of phosphoglycerate kinase

Cite as: J. Chem. Phys. 159, 141102 (2023); doi: 10.1063/5.0166124

Submitted: 3 July 2023 • Accepted: 25 September 2023 •

Published Online: 11 October 2023



View Online



Export Citation



CrossMark

Abir N. Hassani,<sup>1,2</sup>  Luman Haris,<sup>2,3</sup>  Markus Appel,<sup>4</sup>  Tilo Seydel,<sup>4</sup>  Andreas M. Stadler,<sup>2,3,a)</sup>   
and Gerald R. Kneller<sup>1,5,b)</sup> 

## AFFILIATIONS

<sup>1</sup>Centre de Biophysique Moléculaire, CNRS and Université d'Orléans, Rue Charles Sadron, 45071 Orléans, France

<sup>2</sup>Jülich Centre for Neutron Science (JCNS-1) and Institute of Biological Information Processing (IBI-8), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

<sup>3</sup>Institute of Physical Chemistry, RWTH Aachen University, Landoltweg 2, 52056 Aachen, Germany

<sup>4</sup>Institut Laue Langevin, 71 Avenue des Martyrs, 38042 Grenoble Cedex 9, France

<sup>5</sup>Laboratoire des biomolécules, Département de chimie, Ecole Normale Supérieure, 75005 Paris, France

<sup>a)</sup>E-mail: [a.stadler@fz-juelich.de](mailto:a.stadler@fz-juelich.de)

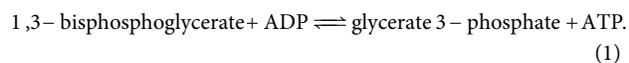
<sup>b)</sup>Author to whom correspondence should be addressed: [gerald.kneller@cnrs.fr](mailto:gerald.kneller@cnrs.fr)

## ABSTRACT

We present an analysis of high-resolution quasi-elastic neutron scattering spectra of phosphoglycerate kinase which elucidates the influence of the enzymatic activity on the dynamics of the protein. We show that in the active state the inter-domain motions are amplified and the intra-domain asymptotic power-law relaxation  $\propto t^{-\alpha}$  is accelerated, with a reduced coefficient  $\alpha$ . Employing an energy landscape picture of protein dynamics, this observation can be translated into a widening of the distribution of energy barriers separating conformational substates of the protein.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0166124>

Understanding the functional dynamics of enzymes is a fundamental issue in molecular biophysics, biology, and biochemistry. Phosphoglycerate kinase (PGK) is one example for which the dynamics-function relationship has been intensively studied with various methods, including structural nuclear magnetic resonance (NMR), x-ray crystallography, quasielastic neutron scattering (QENS), neutron spin echo (NSE) spectroscopy, and molecular dynamics (MD) simulation.<sup>1–10</sup> PGK is a monomeric enzyme which is fundamental for the metabolism of all living organisms. By converting 1,3-bisphosphoglycerate to 3-phosphoglycerate it catalyzes one of the two ATP-producing reactions of the glycolytic pathway and it participates also in gluconeogenesis by catalyzing the opposite reaction to produce 1,3-bisphosphoglycerate and adenosine diphosphate phosphate (ADP),<sup>11</sup>



Yeast PGK has a weight of about 45 kDa and is composed of two domains which are connected by a well conserved hinge region where the catalytic reactions take place. Several of the studies cited above have been performed with the particular goal to better understand the role of the inter-domain motions for the function of the enzyme.<sup>4,6,7,9</sup> A powerful space and time-resolved method for this purpose is neutron spin echo (NSE) spectroscopy, which has been used in Ref. 7 in combination with small-angle neutron scattering (SANS) and normal mode analysis and in Ref. 9 in combination with MD simulation. Standard NSE probes the slow motions and global diffusion of proteins on a 0.1–100 ns time scale and on a nm length scale. The results of the NSE studies suggest that the hinge-bending motions of the two domains in PGK enable its enzymatic activity and that the presence of the substrates rigidifies the molecular and accelerates its internal dynamics.

The present article aims at extending and consolidating the abovementioned work with an analysis of QENS data from the high-resolution spectrometer IN16B at the Institut Laue-Langevin

in Grenoble. The instrument probes the ns time scale if operated in BATS mode (backscattering and time of flight spectroscopy)<sup>12</sup> and closes the gap between QENS experiments with standard time-of-flight spectrometers and NSE spectroscopy.

The QENS experiments on PGK were performed at 283 K in presence and absence of the substrates (13 mM MgATP, 41 mM 3PG, 20 mM MOPS, 50 mM NaCl, 2 mM EDTA, pD 7.4, 99.9% atom D deuteriumoxide) using a PGK concentration of 50 mg/ml. PGK from yeast and all chemicals were obtained commercially from Sigma-Aldrich. These conditions are the same as in Ref. 7 and warrant that PGK in presence of substrates is more than 90% in the ligand-bound state. Prior to data analysis, the solvent-contributions were subtracted. Since about 50% of the atoms in a protein are hydrogen atoms, which have a strongly dominant cross section for incoherent neutron scattering, the dynamic structure factor for QENS from PGK can be written in the form

$$S(\mathbf{q}, \omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dt e^{-i\omega t} F(\mathbf{q}, t), \quad (2)$$

$$F(\mathbf{q}, t) \approx \frac{1}{N} \sum_{j \in \mathbb{H}} \left\langle e^{-i\mathbf{q} \cdot \hat{\mathbf{x}}_j(0)} e^{i\mathbf{q} \cdot \hat{\mathbf{x}}_j(t)} \right\rangle. \quad (3)$$

Here  $N$  is the number of hydrogen atoms,  $\hat{\mathbf{x}}_j(t)$  is the time-dependent position operator of hydrogen atom  $j$ ,  $\mathbf{q}$  the scattering vector, and the symbol  $\langle \dots \rangle$  denotes a quantum ensemble average. The latter leads to the symmetry relations  $F^*(\mathbf{q}, t) = F(\mathbf{q}, -t)$  and  $F(\mathbf{q}, t) = F(-\mathbf{q}, -t + i\beta\hbar)$ .

As in several previous studies,<sup>13–16</sup> the analysis of the QENS data has been performed in the time domain, employing a stochastic model for the classical intermediate scattering function and assuming that Schofield's semiclassical approximation<sup>17,18</sup> is valid.

For model building purposes, we assume that there is a representative hydrogen atom “ $a$ ” whose dynamics accounts for both the relaxation dynamics of the individual hydrogen atoms and their motional heterogeneity. Within Schofield's semiclassical approximation we have then

$$F^{(+)}(\mathbf{q}, t) \approx \left\langle e^{-i\mathbf{q} \cdot \hat{\mathbf{x}}_a(0)} e^{i\mathbf{q} \cdot \hat{\mathbf{x}}_a(t)} \right\rangle_{\text{cl}}, \quad (4)$$

where  $\langle \dots \rangle_{\text{cl}}$  stands for a classical ensemble average. We assume furthermore that the domains in PGK can be treated as equivalent and that the motions of the scattering atom are uncorrelated with the motions of the domain to which it is attached. Writing  $\mathbf{x}_a = \mathbf{R}_a + \mathbf{r}_a$ , where  $\mathbf{R}_a$  points to the center of the domain and  $\mathbf{r}_a$  to the position of the scattering atom with respect to that reference point (see Fig. 1), the orientation-averaged intermediate scattering function of PGK in solution can then be factorized as

$$\bar{F}^{(+)}(q, t) \approx f(q, t)g(q, t) \quad (q \equiv |\mathbf{q}|), \quad (5)$$

where

$$f(q, t) \equiv \overline{\langle a_a^*(t) a_a(t) \rangle}_{\text{cl}} \quad \text{and} \quad g(q, t) \equiv \overline{\langle A_a^*(t) A_a(t) \rangle}_{\text{cl}} \quad (6)$$

are the orientation-averaged autocorrelation functions related to the respective dynamical variables

$$A_{\mathbf{q}} \equiv e^{i\mathbf{q} \cdot \mathbf{R}_a} \quad \text{and} \quad a_{\mathbf{q}} \equiv e^{i\mathbf{q} \cdot \mathbf{r}_a}. \quad (7)$$

Both are to be understood as complex stochastic variables, depending parametrically on the respective position vectors and on  $\mathbf{q}$ , whose time evolution is described by appropriate diffusion models to be described in the following. Conceptually our approach can be compared to modeling reaction coordinates describing protein folding,<sup>19</sup> and it has been used so far to model QENS from the internal dynamics of proteins and from water (see Refs. 13–16 cited above). We apply it here in addition to account for inter-domain motions in PGK by a refined model compared to the one we used in our recent study of Myelin Basic Protein (MBP).<sup>14</sup> Defining

$$\xi \equiv \begin{pmatrix} \Re\{A_{\mathbf{q}} - \langle A_{\mathbf{q}} \rangle_{\text{cl}}\} \\ \Im\{A_{\mathbf{q}} - \langle A_{\mathbf{q}} \rangle_{\text{cl}}\} \end{pmatrix}, \quad (8)$$

we consider in particular that the conditional probability,  $P(\xi, t | \xi_0, 0)$  for a transition  $\xi_0 \rightarrow \xi$  within time  $t$  describes an Ornstein–Uhlenbeck (OU) process,<sup>20</sup> i.e., diffusion in a harmonic potential (left part of Fig. 2),

$$\partial_t P = \left\{ \eta_{\xi} \frac{\partial}{\partial \xi} \cdot \{ \xi P \} + D_{\xi} \frac{\partial}{\partial \xi} \cdot \frac{\partial P}{\partial \xi} \right\}. \quad (9)$$

Here  $\eta_{\xi}$  is a relaxation constant and  $D_{\xi}$  is a diffusion coefficient which can be related to  $\eta_{\xi}$  via  $D_{\xi} = \eta_{\xi} k_B T / K_{\xi} = \eta_{\xi} \langle |\xi|^2 \rangle_{\text{cl}}$ , with  $K_{\xi}$  being the force constant of the harmonic potential. We assume in particular that  $\langle A_{\mathbf{q}} \rangle \rightarrow 0+$ , without vanishing identically, such that  $\xi_1$  and  $\xi_2$  can be formally considered as independent dynamical variables. The essential point is that the autocorrelation function of  $\xi$  decays exponentially,

$$\langle \xi^T(0) \cdot \xi(t) \rangle_{\text{cl}} = \langle |\xi(0)|^2 \rangle_{\text{cl}} e^{-\eta_{\xi} t}, \quad \eta_{\xi} \equiv \eta_{\xi}(\mathbf{q}). \quad (10)$$

Since  $\langle A_{\mathbf{q}} \rangle \approx 0$  for a freely diffusing particle it follows that  $\langle |\xi(0)|^2 \rangle_{\text{cl}} \approx 1$ , and we write therefore

$$g(q, t) = e^{-D(q)q^2|t|}, \quad (11)$$

assuming that the relaxation coefficient  $\eta_{\xi}(\mathbf{q})$  does not depend on the direction of  $\mathbf{q}$ . The parameter  $D(q)$  is here a  $q$ -dependent diffusion coefficient. For small  $q$ -values it follows from the cumulant expansion of the intermediate scattering function that

$$g(q, t) \stackrel{q \rightarrow 0}{=} e^{-\frac{q^2}{6} \langle (\mathbf{R}_a(t) - \mathbf{R}_a(0))^2 \rangle_{\text{cl}}} \stackrel{t \rightarrow \infty}{\sim} e^{-D_0 q^2 t}. \quad (12)$$

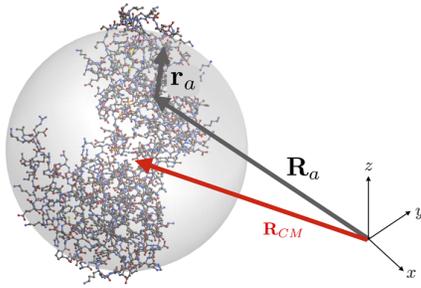
where  $D_0$  is the diffusion coefficient for a whole domain and thus for the whole PGK molecule. For finite  $q$ -values  $g(q, t)$  describes the diffusion of whole PGK molecules which is modulated by inter-domain motions.

For the intra-domain dynamics we use the same model as for the internal protein dynamics in Refs. 13 and 14 and in the light of the preceding discussions the dynamics of the vector variable

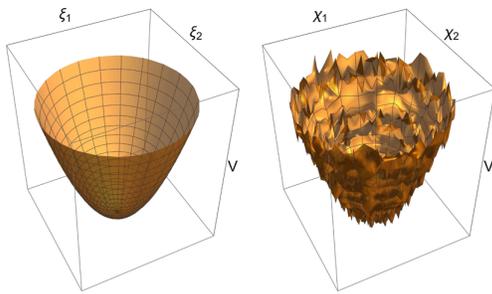
$$\chi \equiv \begin{pmatrix} \Re\{a_{\mathbf{q}} - \langle a_{\mathbf{q}} \rangle_{\text{cl}}\} \\ \Im\{a_{\mathbf{q}} - \langle a_{\mathbf{q}} \rangle_{\text{cl}}\} \end{pmatrix}, \quad (13)$$

is described by a fractional Ornstein Uhlenbeck process,<sup>21,22</sup>

$$\partial_t p = {}_0\partial_t^{1-\alpha} \left\{ \eta_{\chi}^{(\alpha)} \frac{\partial}{\partial \chi} \cdot \{ \chi p \} + D_{\chi}^{(\alpha)} \frac{\partial}{\partial \chi} \cdot \frac{\partial p}{\partial \chi} \right\}. \quad (14)$$



**FIG. 1.** The PGK molecule (PDB code 3PGK) together with a sphere of radius  $R_H = 30.5 \text{ \AA}$  which is used for the Stokes–Einstein relation (18) and the definitions of  $R_g$  and  $r_g$ . The red arrow points to the center-of-mass.



**FIG. 2.** Smooth harmonic potential (left) and its rough counterpart (right) steering the diffusion of  $\xi$  and  $\chi$ , respectively.

Here  $0 < \alpha < 1$  and  ${}_0\partial_t^{1-\alpha}$  denotes a fractional derivative of order  $1 - \alpha$  with respect to  $t$ ,<sup>23</sup>  $\eta_\chi^{(\alpha)}$  is a fractional relaxation constant, and  $D_\chi^{(\alpha)}$  the corresponding fractional diffusion coefficient. The autocorrelation function of  $\chi$  evolves according to

$$\langle \chi^T(0) \cdot \chi(t) \rangle_{cl} = \langle |\chi|^2 \rangle_{cl} E_\alpha(-\eta_\alpha t^\alpha), \quad \eta_\alpha \equiv \eta_\alpha(\mathbf{q}), \quad (15)$$

where  $E_\alpha(-\eta_\alpha t^\alpha)$  is a stretched Mittag-Leffler (ML) function,<sup>23,24</sup> which interpolates between a stretched exponential at short times and an inverse power law at long times and converges to a simple exponential (Debye relaxation) for  $\alpha \rightarrow 1$ . Noting that  $\langle |\chi|^2 \rangle_{cl} = 1 - EISF(q)$ , where  $EISF(q) \equiv |a_q|_{cl}^2$  is the elastic intensity of an isotropic system, we arrive at

$$f(q, t) \equiv EISF(q) + (1 - EISF(q))E_\alpha(-(|t|/\tau)^\alpha). \quad (16)$$

Here  $EISF(q)$ ,  $\tau \equiv \tau(q)$ , and  $\alpha \equiv \alpha(q)$  are considered as  $q$ -dependent fit parameters. The fractional Ornstein–Uhlenbeck process can be visualized as diffusion process in a “rough” two-dimensional harmonic potential (right part Fig. 2), which is characterized by a wide distribution of energy barriers separating various minima or “conformational substates.”<sup>25</sup> In the framework of the Generalized Langevin Equation<sup>26</sup> the correlation function (15) is characterized by a memory function with an algebraic long time tail  $\propto t^{\alpha-2}/\Gamma(\alpha - 1)$  whose amplitude vanishes for  $\alpha \rightarrow 1$ . We note here

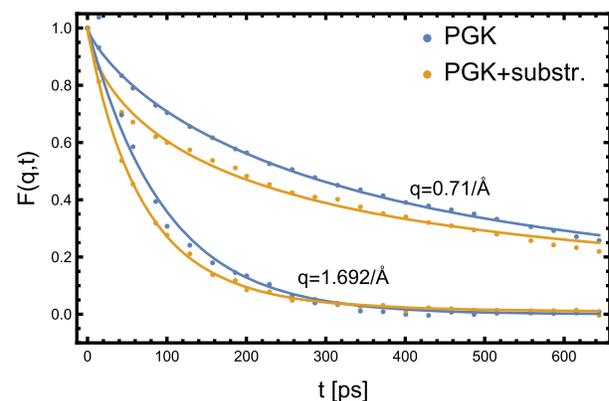
that the parameter  $\alpha$  entirely determines the distribution function for the dimensionless barrier,  $\epsilon = \Delta E/k_B T$ ,

$$P_{ML}(\epsilon) = \frac{2e \sin(\pi\alpha)}{\pi(e^{-\alpha\epsilon^2} + e^{\alpha\epsilon^2} + 2 \cos(\pi\alpha))}. \quad (17)$$

For  $\alpha \rightarrow 1$  the barrier distribution,  $P_{ML}(\epsilon)$ , is entirely concentrated on  $\epsilon = 0$ .

We start the discussion of the results with the right panel of Fig. 3, which shows a fit of the resolution-deconvolved intermediate scattering function,  $\bar{F}^{(+)}(q, t) \approx f(q, t)g(q, t)$ , with four parameters,  $\tau(q)$ ,  $\alpha(q)$ ,  $EISF(q)$ , and  $D(q)$  for the minimum and maximum  $q$ -values in absence and presence of substrates. The data reduction has been performed with the MANTID package<sup>27</sup> and the code for the data analysis has been developed with Wolfram Mathematica.<sup>28</sup> Figure 4 displays the parameters  $\tau$  (left panel) and  $\alpha$  (right panel) as a function of  $q$  in presence and absence of the substrates (blue and yellow triangles, respectively). A clear impact of the presence of substrates on the intra-domain dynamics can be seen: Both  $\tau$  and  $\alpha$  are systematically reduced in presence of the substrates, which indicates that the internal molecular dynamics is accelerated by the enzymatic activity of the molecule and that the relaxation dynamics of the domains becomes less exponential. We note that  $\tau$  and  $\alpha$  in presence and absence of the substrates follow globally the same evolution with  $q$ . The time scale parameter  $\tau$  becomes generally smaller with increasing  $q$ , which simply indicates that localized motions are faster than collective motions implying a large number of atoms. The form parameter,  $\alpha$ , increases instead with  $q$  to values close to 1, indicating increasingly exponential relaxation for more localized motions. We attribute this behavior to the fact that less relaxation modes contribute to localized motions than to large amplitude motions which are probed at small values of  $q$ .

Figure 5 presents the fitted EISFs together with the integrated measured QENS intensity over the width of the resolution function, which can be considered as measured counterpart. The difference



**FIG. 3.** Resolution-deconvolved  $F(q, t)$  of PGK in deuterated solution without and in presence of substrates (blue and yellow dots, respectively) and the corresponding fits (blue and yellow line, respectively).

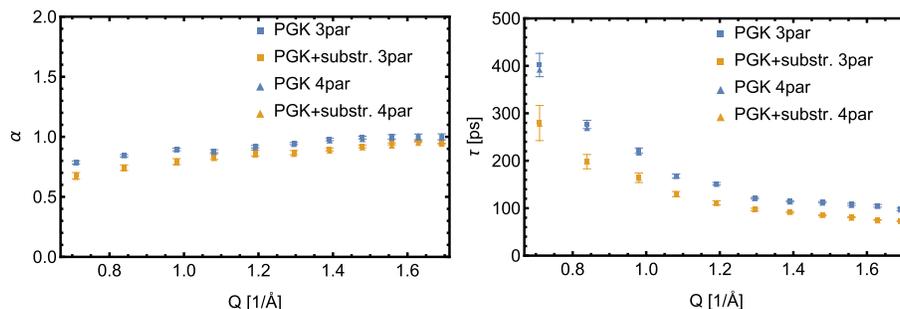


FIG. 4. The model parameters  $\tau$  and  $\alpha$  for PGK. Triangles indicate four-parameter fits (no error bars given) and squares three-parameter fits (with error bars). More explanations are given in the text.

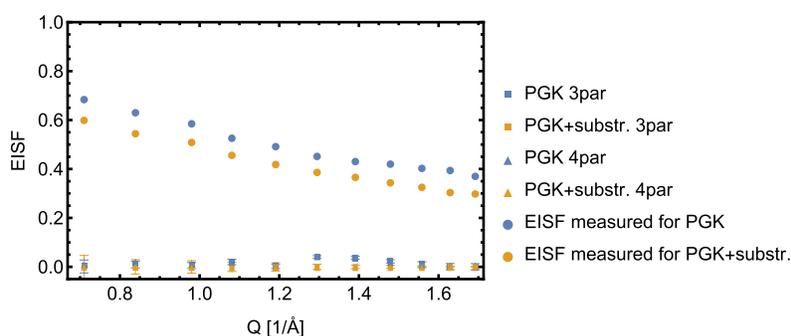


FIG. 5. The EISF parameter for PGK in absence and presence of substrates.

between the two can be attributed to unavoidable spurious contributions from quasielastic scattering, which are due to finite instrumental resolution. They are here estimated on the basis of the underlying model for the relaxation function,  $\phi(q, t) \equiv E_{\alpha}(-|t|/\tau)^{\alpha}$ , and the measured elastic intensity. The details can be found in Ref. 29. We find that the fitted EISF is globally close to zero in the presence and absence of the ligand, except at  $q = 1.3 \text{ \AA}^{-1}$  where the EISF of PGK in absence of substrates ligand is slightly larger than the EISF in their presence. Correlating this observation with the decrease of  $\alpha$  in presence of the substrates shows that the domains are slightly stiffened, which confirms again the findings in Ref. 7 which were obtained by NSE spectroscopy.

For comparison we show also fits with a reduced model where  $D(q) \equiv 0$  (blue and yellow squares, respectively). It can be clearly seen that the results are very similar, the difference being the error bars, which are much larger for the fit of all four parameters and which are not shown here. This observation is in line with the findings in Ref. 14 for the intrinsically disordered Myelin Basic Protein (MBP) and we present the three-parameter fits to show that the fits of  $\tau$ ,  $\alpha$  and  $EISF$  are stable.

The impact of the enzymatic dynamics on the intra-domain energy landscape can be visualized by comparing the energy barrier profiles,  $P_{ML}(\epsilon)$ , describing its “roughness,”<sup>13,14</sup> which are displayed in the left panel of Fig. 6. Important differences between the two profiles are again observed for  $q$ -values corresponding to opening

amplitudes of the hinge region and indicate a wider distribution of energy barriers in presence of the substrates. This corresponds to the decrease of the alpha parameter described above, indicating a stiffening of PGK in its active mode.

The  $q$ -dependent diffusion coefficient is displayed in the right panel of Fig. 6. One observes that  $D(q)$  displays a pronounced modulation with respect to its values at small and large  $q$ -values. The latter are close to the estimation for the diffusion coefficient of a whole PGK molecule obtained from the Stokes–Einstein law,

$$D_0 = \frac{k_B T}{6\pi\eta R_H} \approx 5.1 \times 10^{-3} \text{ \AA}^2/\text{ps}. \quad (18)$$

For this estimation we used an effective hydrodynamic radius of  $R_H = 30.5 \text{ \AA}$  calculated from the PDB structure 3PGK, including the diameter of a water molecule (see Fig. 1). The maximum of  $D(q)$  at about  $q_{\max} \approx 1.2 \text{ \AA}^{-1}$  corresponds to  $2\pi/q_{\max} \approx 5 \text{ \AA}$  in real space, which can be associated with breathing motions of the hinge-region in PGK caused by its enzymatic activity and which have also been observed by combining NSE spectroscopy, normal mode analysis, and hydrodynamic molecular modelling.<sup>7</sup> The fact that  $D(q) \approx D_0$  for smaller  $q$ -values is in line with the requirement that  $g(q, t)$  must here describe diffusion of whole PGK molecules and  $D(q) \approx D_0$  at

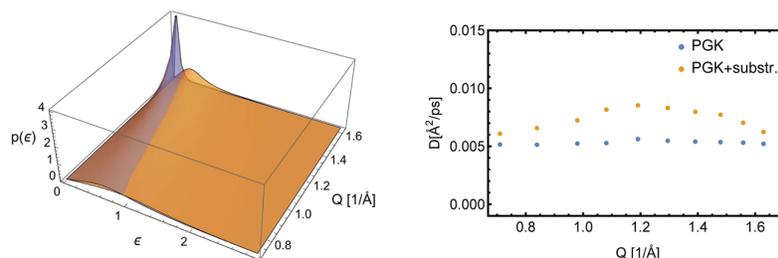


FIG. 6. Left panel: Energy barrier spectrum in absence (blue) and presence (yellow) of substrates. Right panel: The fitted diffusion coefficient  $D(q)$ .

higher  $q$ -values reflects that more localized motions do not affect the global diffusive dynamics of PGK.

To resume, we can say that our QENS study gives a consistent picture of the functional dynamics of PGK which confirms and completes an earlier study by NSE spectroscopy. It shows in particular that the “minimalistic” model used in this work suffices to extract the essential information in the QENS data through a physical interpretation of  $q$ -dependent model parameters. In this context the strongly non-exponential relaxation dynamics in proteins is a key element which must be accounted for to fully exploit the rich information content in the QENS data. We finally remark that the concept can be easily extended to describe also coherent scattering. In this case one simply considers collective dynamical variables of the form  $\alpha_q \equiv \sum_j \exp(iq \cdot x_j)/\sqrt{N}$ .

We gratefully acknowledge beam time allocation on the IN16B neutron spectrometer at the Institut Laue-Langevin, Grenoble, France. A.N.H. acknowledges financial support from the Région Centre - Val de Loire and the Jülich Centre for Neutron Science.

#### AUTHOR DECLARATIONS

##### Conflict of Interest

The authors have no conflicts to disclose.

##### Author Contributions

**Abir N. Hassani:** Formal analysis (lead); Investigation (equal); Methodology (supporting); Software (equal); Visualization (lead); Writing – original draft (supporting). **Luman Haris:** Investigation (supporting). **Markus Appel:** Investigation (equal); Resources (equal). **Tilo Seydel:** Investigation (equal); Resources (equal). **Andreas M. Stadler:** Conceptualization (equal); Investigation (lead); Resources (equal); Writing – original draft (supporting). **Gerald R. Kneller:** Conceptualization (lead); Formal analysis (supporting); Investigation (equal); Methodology (lead); Software (equal); Supervision (lead); Writing – original draft (lead).

##### DATA AVAILABILITY

The data that support the findings of this study are openly available at <http://doi.ill.fr/10.5291/ILL-DATA.8-04-916> and the analysis software under <https://doi.org/10.5281/zenodo.8396244>.

#### REFERENCES

- P. Tanswell, E. W. Westhead, and R. J. P. Williams, *Eur. J. Biochem.* **63**, 249 (1976).
- R. D. Banks *et al.*, *Nature* **279**, 773 (1979).
- M. A. Sinev, O. I. Razgulyaev, M. Vas, A. A. Timchenko, and O. B. Ptitsyn, *Eur. J. Biochem.* **180**, 61 (1989).
- B. E. Bernstein, P. A. M. Michels, and W. G. J. Hol, *Nature* **385**, 275 (1997).
- E. Balog *et al.*, *Phys. Rev. Lett.* **93**, 028103 (2004).
- Z. Palmi *et al.*, *Proteins: Struct., Funct., Bioinf.* **77**, 319 (2009).
- R. Inoue *et al.*, *Biophys. J.* **99**, 2309 (2010).
- E. Balog, D. Perahia, J. C. Smith, and F. Merzel, *J. Phys. Chem. B* **115**, 6811 (2011).
- N. Smolin, R. Biehl, G. R. Kneller, D. Richter, and J. C. Smith, *Biophys. J.* **102**, 1108 (2012).
- X. Hu *et al.*, *Nat. Phys.* **12**, 171 (2016).
- P. D. Boyer, *The Enzymes. 8: Group Transfer. Part A: Nucleotidyl Transfer, Nucleosidyl Transfer, Acyl Transfer, Phosphoryl Transfer*, 3rd ed. (Academic Press, New York, 1973).
- M. Appel, B. Frick, and A. Magerl, *Physica B* **562**, 6 (2019).
- M. Saouessi, J. Peters, and G. R. Kneller, *J. Chem. Phys.* **150**, 161104 (2019).
- A. N. Hassani *et al.*, *J. Chem. Phys.* **156**, 025102 (2022).
- M. H. Petersen *et al.*, *J. Phys. Chem. C* **125**, 15085 (2021).
- M. H. Petersen, M. T. F. Telling, G. Kneller, and H. N. Bordallo, *EPJ Web Conf.* **272**, 01012 (2022).
- G. R. Kneller, *Proc. Natl. Acad. Sci. U. S. A.* **115**, 9450 (2018).
- P. Schofield, *Phys. Rev. Lett.* **4**, 239 (1960).
- J. N. Onuchic, Z. Luthey-Schulten, and P. G. Wolynes, *Annu. Rev. Phys. Chem.* **48**, 545 (1997).
- G. Uhlenbeck and L. Ornstein, *Phys. Rev.* **36**, 823 (1930).
- R. Metzler, J. H. Jeon, A. G. Cherstvy, and E. Barkai, *Phys. Chem. Chem. Phys.* **16**, 24128 (2014).
- R. Metzler and J. Klafter, *Phys. Rep.* **339**, 1 (2000).
- Mittag-Leffler Functions, Related Topics and Applications, Springer Monographs in Mathematics*, edited by R. Gorenflo, A. A. Kilbas, F. Mainardi, and S. V. Rogosin (Springer, Heidelberg, 2014).
- W. G. Glöckle and T. F. Nonnenmacher, *Biophys. J.* **68**, 46 (1995).
- H. Frauenfelder, S. G. Sligar, and P. G. Wolynes, *Science* **254**, 1598 (1991).
- R. Zwanzig, *Nonequilibrium Statistical Mechanics* (Oxford University Press, Oxford; New York, 2001).
- O. Arnold, *Nucl. Instrum. Methods Phys. Res., Sect. A* **764**, 156 (2014).
- WolfRAM Research, Inc., Mathematica, Version 13.2, Champaign, IL, 2021.
- A. N. Hassani, A. M. Stadler, and G. R. Kneller, *J. Chem. Phys.* **157**, 134103 (2022).

## Bibliographie

- [1] Charles Tanford and Jacqueline Reynolds. *Nature's Robots : A History of Proteins*. Oxford Univ. Press, New York, NY Oxford, 1. issued in paperback edition, 2003.
- [2] Christian B. Anfinsen. Principles that Govern the Folding of Protein Chains. *Science*, 181(4096) :223–230, 1973.
- [3] C. B. Anfinsen, E. Haber, M. Sela, and F. H. White. THE KINETICS OF FORMATION OF NATIVE RIBONUCLEASE DURING OXIDATION OF THE REDUCED POLYPEPTIDE CHAIN. *Proceedings of the National Academy of Sciences*, 47(9) :1309–1314, 1961.
- [4] H. Frauenfelder, B. H. McMahon, and P. W. Fenimore. Myoglobin : The Hydrogen Atom of Biology and a Paradigm of Complexity. *Proceedings of the National Academy of Sciences*, 100(15) :8615–8617, 2003.
- [5] H Frauenfelder, F Parak, and R D Young. Conformational Substates in Proteins. *Annual Review of Biophysics and Biophysical Chemistry*, 17(1) :451–479, 1988.
- [6] Véronique Hamon. *Influence de la pression sur la dynamique du lysozyme en solution, par simulation numérique et diffusion de neutrons*. PhD thesis, Université d'Orléans, 2004.
- [7] J. C. Kendrew, R. E. Dickerson, B. E. Strandberg, R. G. Hart, D. R. Davies, D. C. Phillips, and V. C. Shore. Structure of Myoglobin : A Three-Dimensional Fourier Synthesis at 2 Å. Resolution. *Nature*, 185(4711) :422–427, 1960.
- [8] H Hartmann, F Parak, W Steigemann, GA Petsko, DR Ponzi, and H Frauenfelder. Conformational Substates in a Protein : Structure and Dynamics of Metmyoglobin at 80 K. *P Natl Acad Sci Usa*, 79(16) :4967, 1982.
- [9] SC Kou and XS Xie. Generalized Langevin Equation with Fractional Gaussian Noise : Subdiffusion within a Single Protein Molecule. *Phys Rev Lett*, 93(18) :180603, 2004.

- [10] H Yang, G Luo, P Karnchanaphanurach, TM Louie, I Rech, S Cova, L Xun, and XS Xie. Protein Conformational Dynamics Probed by Single-Molecule Electron Transfer. *Science*, 302(5643) :262, 2003.
- [11] A. L. Agapov, A. I. Kolesnikov, V. N. Novikov, R. Richert, and A. P. Sokolov. Quantum Effects in the Dynamics of Deeply Supercooled Water. *Physical Review E*, 91(2), 2015.
- [12] Vania Calandrini, Daniel Abergel, and Gerald R. Kneller. Protein dynamics from a NMR perspective : Networks of coupled rotators and fractional brownian dynamics. *The Journal of Chemical Physics*, 128(14) :145102, 2008.
- [13] Paolo Calligari, Vania Calandrini, Gerald R Kneller, and Daniel Abergel. From NMR Relaxation to Fractional Brownian Dynamics in Proteins : Results from a Virtual Experiment. *J Phys Chem B*, 115(43) :12370–12379, 2011.
- [14] Rana Ashkar, Hassina Z. Bilheux, Heliosa Bordallo, Robert Briber, David J. E. Callaway, Xiaolin Cheng, Xiang-Qiang Chu, Joseph E. Curtis, Mark Dadmun, Paul Fenimore, David Fushman, Frank Gabel, Kushol Gupta, Frederick Herberle, Frank Heinrich, Liang Hong, John Katsaras, Zvi Kelman, Eugenia Kharlampieva, Gerald R. Kneller, Andrey Kovalevsky, Susan Krueger, Paul Langan, Raquel Lieberman, Yun Liu, Mathias Losche, Edward Lyman, Yimin Mao, John Marino, Carla Mattos, Flora Meilleur, Peter Moody, Jonathan D. Nickels, William B. O'Dell, Hugh O'Neill, Ursula Perez-Salas, Judith Peters, Loukas Petridis, Alexei P. Sokolov, Christopher Stanley, Norman Wagner, Michael Weinrich, Kevin Weiss, Troy Wymore, Yang Zhang, and Jeremy C. Smith. Neutron scattering in the biological sciences : Progress and prospects. *Acta Crystallographica Section D Structural Biology*, 74(12) :1129–1168, 2018.
- [15] M. Bée. *Quasielastic Neutron Scattering : Principles and Applications in Solid State Chemistry, Biology and Materials Science*. Adam Hilger, Bristol, 1988.
- [16] W Bothe. Künstliche Erregung von Kernstrahlen. page 18.
- [17] Irène Curie and Frédéric Joliot. Etude du rayonnement absorbable accompagnant les rayons  $\alpha$  du polonium. *Journal de Physique et le Radium*, 2(1) :20–28, 1931.
- [18] J Chadwick. The Existence of a Neutron. *Proceedings of the Royal Society A : Mathematical, Physical and Engineering Sciences*, 136(830) :692–708, 1932.

- [19] S.W. Lovesey. *Theory of Neutron Scattering from Condensed Matter*, volume I. Clarendon Press, Oxford, 1984.
- [20] Gerald R Kneller. *Quasielastic Neutron Scattering*, 2004.
- [21] B. A. Lippmann and Julian Schwinger. Variational Principles for Scattering Processes. I. *Physical Review*, 79(3) :469–480, 1950.
- [22] Wolfgang Doster, Carlos Granja, Claude Leroy, and Ivan Stekl. The New Munich Neutron Source FRM II : Overview and Uses for Biological Studies. In *AIP Conference Proceedings*, volume 958, pages 16–19. AIP, 2007.
- [23] Maths Karlsson. Proton dynamics in oxides : Insight into the mechanics of proton conduction from quasielastic neutron scattering. *Physical Chemistry Chemical Physics*, 17(1) :26–38, 2015.
- [24] Léon Van Hove. Correlations in Space and Time and Born Approximation Scattering in Systems of Interacting Particles. *Physical Review*, 95(1) :249, 1954.
- [25] Léon Van Hove. A Remark on the Time-Dependent Pair Distribution. *Physica*, 24(1) :404–408, 1958.
- [26] Tasso Springer. Quasielastic Neutron Scattering for the Investigation of Diffusive Motions in Solids and Liquids. In *Springer Tracts in Modern Physics, Volume 64*, volume 64, pages 1–100. Springer Berlin Heidelberg, Berlin, Heidelberg, 1972.
- [27] R Campos-González and J R Glenney. Tyrosine phosphorylation of mitogen-activated protein kinase in cells with tyrosine kinase-negative epidermal growth factor receptors. *Journal of Biological Chemistry*, 267(21) :14535–14538, 1992.
- [28] Christian B. Anfinsen and Edgar Haber. Studies on the Reduction and Re-formation of Protein Disulfide Bonds. *Journal of Biological Chemistry*, 236(5) :1361–1363, 1961.
- [29] Christian B. Anfinsen. Principles that Govern the Folding of Protein Chains. *Science*, 181(4096) :223–230, 1973.
- [30] Vladimir N. Uversky. Intrinsically disordered proteins from A to Z. *The International Journal of Biochemistry & Cell Biology*, 43(8) :1090–1103, 2011.
- [31] Peter Tompa. Intrinsically disordered proteins : A 10-year recap. *Trends in Biochemical Sciences*, 37(12) :509–516, 2012.

- [32] Peter E. Wright and H. Jane Dyson. Intrinsically disordered proteins in cellular signalling and regulation. *Nature Reviews Molecular Cell Biology*, 16(1) :18–29, 2015.
- [33] Peter Tompa. Intrinsically unstructured proteins. *Trends in Biochemical Sciences*, 27(10) :527–533, 2002.
- [34] H. Jane Dyson and Peter E. Wright. Intrinsically unstructured proteins and their functions. *Nature Reviews Molecular Cell Biology*, 6(3) :197–208, 2005.
- [35] V. N. Uversky. Natively Unfolded Proteins : A Point Where Biology Waits for Physics. *Protein Science*, 11(4) :739–756, 2002.
- [36] Dennis Kurzbach, Thomas C. Schwarz, Gerald Platzler, Simone Höfler, Dariush Hinderberger, and Robert Konrat. Kompensatorische Anpassungen der strukturellen Dynamik eines intrinsisch unstrukturierten Protein-Komplexes. *Angewandte Chemie*, 126(15) :3919–3922, 2014.
- [37] Gerald R. Kneller. Franck–condon picture of incoherent neutron scattering. *Proceedings of the National Academy of Sciences USA*, 115(38) :9450–9455, 2018.
- [38] *The Enzymes. 8 : Group Transfer. Part A : Nucleotidyl Transfer, Nucleosidyl Transfer, Acyl Transfer, Phosphoryl Transfer*. Acad. Pr, New York, 3. ed edition, 1973.
- [39] Markus Appel, Bernhard Frick, and Andreas Magerl. A flexible high speed pulse chopper system for an inverted neutron time-of-flight option on backscattering spectrometers. *Scientific Reports*, 8(1) :13580, 2018.
- [40] Markus Appel, Bernhard Frick, and Andreas Magerl. First results with the neutron backscattering and TOF spectrometer option BATS on IN16B. *Physica B : Condensed Matter*, 562 :6–8, 2019.
- [41] J. Ollivier, H. Casalta, H. Schober, J.C. Cook, P. Malbert, M. Locatelli, C. Gomez, S. Jenkins, I.J. Sutton, and M. Thomas. New perspectives on the IN5 time of flight spectrometer. *Applied Physics A : Materials Science & Processing*, 74(0) :s305–s307, 2002.
- [42] O. Arnold, J.C. Bilheux, J.M. Borreguero, A. Buts, S.I. Campbell, L. Chapon, M. Doucet, N. Draper, R. Ferraz Leal, M.A. Gigg, V.E. Lynch, A. Markvardsen, D.J. Mikkelsen, R.L. Mikkelsen, R. Miller, K. Palmen, P. Parker, G. Passos, T.G. Perring, P.F. Peterson, S. Ren, M.A. Reuter, A.T. Savici, J.W. Taylor, R.J. Taylor, R. Tolchenov, W. Zhou, and J. Zikovsky.

- Mantid—Data analysis and visualization package for neutron scattering and  $\mu$  SR experiments. *Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment*, 764 :156–166, 2014.
- [43] Andreas M. Stadler, Franz Demmel, Jacques Ollivier, and Tilo Seydel. Picosecond to Nanosecond Dynamics Provide a Source of Conformational Entropy for Protein Folding. *Phys. Chem. Chem. Phys.*, 18(31) :21527–21538, 2016.
- [44] R. H. Austin, K. W. Beeson, L. Eisenstein, H. Frauenfelder, and I. C. Gunsalus. Dynamics of ligand binding to myoglobin. *Biochemistry*, 14(24) :5355–5373, 1975.
- [45] N. Alberding, R. H. Austin, Shirley S. Chan, L. Eisenstein, H. Frauenfelder, I. C. Gunsalus, and T. M. Nordlund. Dynamics of carbon monoxide binding to protoheme. *The Journal of Chemical Physics*, 65(11) :4701–4711, 1976.
- [46] H Frauenfelder, S G Sligar, and P G Wolynes. The Energy Landscapes and Motions of Proteins. *Science*, 254(5038) :1598–1603, 1991.
- [47] Rv H Austin, K W Beeson, L Eisenstein, H Frauenfelder, and I C Gunsalus. Dynamics of ligand binding to myoglobin. *Biochemistry*, 14(24) :5355–5373, 1975.
- [48] WG Glöckle and TF Nonnenmacher. A fractional calculus approach to self-similar protein dynamics. *Biophysical Journal*, 68(1) :46–53, 1995.
- [49] IE Iben, D Braunstein, W Doster, H Frauenfelder, MK Hong, JB Johnson, S Luck, P Ormos, A Schulte, PJ Steinbach, AH Xie, and RD Young. Glassy behavior of a protein. *Physical Review Letters*, 62(16) :1916–1919, 1989.
- [50] P. Schofield. Space-Time Correlation Function Formalism for Slow Neutron Scattering. *Physical Review Letters*, 4(5) :239–240, 1960.
- [51] N. Alberding, R. H. Austin, Shirley S. Chan, L. Eisenstein, H. Frauenfelder, I. C. Gunsalus, and T. M. Nordlund. Dynamics of carbon monoxide binding to protoheme. *The Journal of Chemical Physics*, 65(11) :4701–4711, 1976.
- [52] W Doster, S Cusack, and W Petry. Dynamical Transition of Myoglobin Revealed by Inelastic Neutron Scattering. *Nature*, 337(6209) :754–756, 1989.
- [53] H Yang and XS Xie. Probing single-molecule dynamics photon by photon. *The Journal of Chemical Physics*, 117 :10965, 2002.

- [54] Abir N. Hassani, Andreas M. Stadler, and Gerald R. Kneller. Quasi-analytical resolution-correction of elastic neutron scattering from proteins. *The Journal of Chemical Physics*, 157(13) :134103, 2022.
- [55] Rudolf Gorenflo, Anatoly A. Kilbas, Francesco Mainardi, and Sergei V. Rogosin, editors. *Mittag-Leffler Functions, Related Topics and Applications*. Springer Monographs in Mathematics. Springer, Heidelberg, 2014.
- [56] G R Kneller. Quasielastic Neutron Scattering and Relaxation Processes in Proteins : Analytical and Simulation-Based Models. *Phys Chem Chem Phys*, 7(13) :2641–2655, 2005.
- [57] R. Zwanzig. Diffusion in a Rough Potential. *Proceedings of the National Academy of Sciences*, 85(7) :2029–2030, 1988.
- [58] Melek Saouessi, Judith Peters, and Gerald R. Kneller. Asymptotic analysis of quasielastic neutron scattering data from human acetylcholinesterase reveals subtle dynamical changes upon ligand binding. *The Journal of Chemical Physics*, 150(16) :161104, 2019.
- [59] Abir N. Hassani, Luman Haris, Markus Appel, Tilo Seydel, Andreas M. Stadler, and Gerald R. Kneller. Multiscale relaxation dynamics and diffusion of myelin basic protein in solution studied by quasielastic neutron scattering. *The Journal of Chemical Physics*, 156(2) :025102, 2022.
- [60] Stephen W. Provencher. CONTIN : A general purpose constrained regularization program for inverting noisy linear algebraic and integral equations. *Computer Physics Communications*, 27(3) :229–242, 1982.
- [61] András Micsonai, Frank Wien, Linda Kernya, Young-Ho Lee, Yuji Goto, Matthieu Réfrégiers, and József Kardos. Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. *Proceedings of the National Academy of Sciences*, 112(24) :E3095–E3103, 2015.
- [62] Guillaume Chevrot, Paolo Calligari, Konrad Hinsén, and Gerald R Kneller. Least constraint approach to the extraction of internal motions from molecular dynamics trajectories of flexible macromolecules. *The Journal of Chemical Physics*, 135(8) :084110, 2011.
- [63] Satoru Fujiwara, Tatsuhito Matsuo, Yasunobu Sugimoto, and Kaoru Shibata. Segmental Motions of Proteins under Non-native States Evaluated Using Quasielastic Neutron Scattering. *The Journal of Physical Chemistry Letters*, 10(23) :7505–7509, 2019.

- [64] Krzysztof Sozański, Agnieszka Wiśniewska, Tomasz Kalwarczyk, and Robert Hołyst. Activation Energy for Mobility of Dyes and Proteins in Polymer Solutions : From Diffusion of Single Particles to Macroscale Flow. *Physical Review Letters*, 111(22) :228301, 2013.
- [65] Irina V. Nesselova, Daria L. Melnikova, Venkatesh Ranjan, and Vladimir D. Skirda. Translational diffusion of unfolded and intrinsically disordered proteins. In *Progress in Molecular Biology and Translational Science*, volume 166, pages 85–108. Elsevier, 2019.
- [66] David W. McCall, Dean C. Douglass, and Ernest W. Anderson. Diffusion in Liquids. *The Journal of Chemical Physics*, 31(6) :1555–1557, 1959.
- [67] R. Mills. Self-diffusion in normal and heavy water in the range 1-45.deg. *The Journal of Physical Chemistry*, 77(5) :685–688, 1973.
- [68] Magnus Kjaergaard, Ann-Beth Nørholm, Ruth Hendus-Altenburger, Stine F. Pedersen, Flemming M. Poulsen, and Birthe B. Kragelund. Temperature-dependent structural changes in intrinsically disordered proteins : Formation of  $\alpha$ -helices or loss of polyproline II? : Temperature Dependence of Protein Disorder. *Protein Science*, 19(8) :1555–1564, 2010.
- [69] R. Wuttke, H. Hofmann, D. Nettels, M. B. Borgia, J. Mittal, R. B. Best, and B. Schuler. Temperature-dependent solvation modulates the dimensions of disordered proteins. *Proceedings of the National Academy of Sciences*, 111(14) :5213–5218, 2014.
- [70] Melek Saouessi, Judith Peters, and Gerald R. Kneller. Communication : Frequency domain modeling of quasielastic neutron scattering from hydrated protein powders : Application to free and inhibited human acetylcholinesterase. *The Journal of Chemical Physics*, 151(12) :125103, 2019.
- [71] R Metzler, J H Jeon, and A G Cherstvy. Anomalous Diffusion Models and Their Properties : Non-Stationarity, Non-Ergodicity, and Ageing at the Centenary of Single Particle Tracking. *Phys Chem Chem Phys*, 16 :24128–24164, 2014.
- [72] R Metzler and J Klafter. The Random Walk’s Guide to Anomalous Diffusion : A Fractional Dynamics Approach. *Phys Rep*, 339(1) :1–77, 2000.
- [73] Frank W J Olver, Daniel W Lozier, and Ronald F Boisvert. *NIST Handbook of Mathematical Functions*. Cambridge Univ Pr, 2010.
- [74] A. Rahman, K. S. Singwi, and A. Sjölander. Theory of Slow Neutron Scattering by Liquids. I. *Physical Review*, 126(3) :986–996, 1962.

- [75] Gerald R. Kneller. Asymptotic Neutron Scattering Laws for Anomally Diffusing Quantum Particles. *The Journal of Chemical Physics*, 145(4) :044103, 2016.
- [76] Wolfram Research, Inc. Mathematica, Version 13.0. Champaign, IL, 2019.
- [77] Valentina Riveros-Moreno and Jonathan B. Wittenberg. The Self-Diffusion Coefficients of Myoglobin and Hemoglobin in Concentrated Solutions. *Journal of Biological Chemistry*, 247(3) :895–901, 1972.
- [78] Mattia Rocco and Olwyn Byron. Computing translational diffusion and sedimentation coefficients : An evaluation of experimental data and programs. *European Biophysics Journal*, 44(6) :417–431, 2015.
- [79] M. L. Huber, R. A. Perkins, A. Laesecke, D. G. Friend, J. V. Sengers, M. J. Assael, I. N. Metaxa, E. Vogel, R. Mareš, and K. Miyagawa. New International Formulation for the Viscosity of H<sub>2</sub>O. *Journal of Physical and Chemical Reference Data*, 38(2) :101–125, 2009.
- [80] Giuseppe Zaccai. How Soft Is a Protein? A Protein Dynamics Force Constant Measured by Neutron Scattering. *Science*, 288(5471) :1604–1607, 2000.
- [81] Paul Tanswell, Edward W. Westhead, and Robert J. P. Williams. Nuclear-Magnetic-Resonance Study of the Active-Site Structure of Yeast Phosphoglycerate Kinase. *European Journal of Biochemistry*, 63(1) :249–262, 1976.
- [82] R. D. Banks, C. C. F. Blake, P. R. Evans, R. Haser, D. W. Rice, G. W. Hardy, M. Merrett, and A. W. Phillips. Sequence, structure and activity of phosphoglycerate kinase : A possible hinge-bending enzyme. *Nature*, 279(5716) :773–777, 1979.
- [83] Michael A. Sinev, Oleg I. Razgulyaev, Maria Vas, Alexander A. Timchenko, and Oleg B. Ptitsyn. Correlation between enzyme activity and hinge-bending domain displacement in 3-phosphoglycerate kinase. *European Journal of Biochemistry*, 180(1) :61–66, 1989.
- [84] Bradley E. Bernstein, Paul A. M. Michels, and Wim G. J. Hol. Synergistic effects of substrate-induced conformational changes in phosphoglycerate kinase activation. *Nature*, 385(6613) :275–278, 1997.
- [85] Erika Balog, Torsten Becker, Martin Oettl, Ruediger Lechner, Roy Daniel, John Finney, and Jeremy C. Smith. Direct Determination of Vibrational Density of States Change on Ligand Binding to a Protein. *Physical Review Letters*, 93(2) :028103, 2004.

- [86] Zoltan Palmai, Laurent Chaloin, Corinne Lionne, Judit Fidy, David Perahia, and Erika Balog. Substrate binding modifies the hinge bending characteristics of human 3-phosphoglycerate kinase : A molecular dynamics study. *Proteins : Structure, Function, and Bioinformatics*, 77(2) :319–329, 2009.
- [87] R. Inoue, R. Biehl, T. Rosenkranz, J. Fitter, M. Monkenbusch, A. Radulescu, B. Farago, and D. Richter. Large Domain Fluctuations on 50-ns Timescale Enable Catalytic Activity in Phosphoglycerate Kinase. *Biophysical Journal*, 99(7) :2309–2317, 2010.
- [88] Erika Balog, David Perahia, Jeremy C Smith, and Franci Merzel. Vibrational Softening of a Protein on Ligand Binding. *J Phys Chem B*, 115(21) :110510094236054, 2011.
- [89] N Smolin, R Biehl, G R Kneller, D Richter, and J C Smith. Functional Domain Motions in Proteins on the 1– 100 Ns Timescale : Comparison of Neutron Spin-Echo Spectroscopy of Phosphoglycerate Kinase with Molecular-Dynamics Simulation. *Biophysical Journal*, 102(5) :1108–1117, 2012.
- [90] Xiaohu Hu, Liang Hong, Micholas Dean Smith, Thomas Neusius, Xiaolin Cheng, and Jeremy C. Smith. The Dynamics of Single Protein Molecules Is Non-Equilibrium and Self-Similar over Thirteen Decades in Time. *Nature Phys*, 12(2) :171–174, 2016.
- [91] Abir N Hassani, Luman Haris, Gerald R Kneller, Markus Appel, Tilo Seydel, and Andreas M Stadler. Signature of functional enzyme dynamics in quasielastic neutron scattering spectra : The case of phosphoglycerate kinase. *The Journal*, 2023.
- [92] G.E. Uhlenbeck and L.S. Ornstein. On the Theory of the Brownian Motion. *Physical Review*, 36(5) :823, 1930.
- [93] Andreas M. Stadler, Laura Stingaciu, Aurel Radulescu, Olaf Holderer, Michael Monkenbusch, Ralf Biehl, and Dieter Richter. Internal Nanosecond Dynamics in the Intrinsically Disordered Myelin Basic Protein. *Journal of the American Chemical Society*, 136(19) :6987–6994, 2014.