



**HAL**  
open science

# Etude expérimentale de la dynamique moléculaire des protéines : nouvelle méthode de criblages de conformations protéiques

Laëtitia Bourgeat

► **To cite this version:**

Laëtitia Bourgeat. Etude expérimentale de la dynamique moléculaire des protéines : nouvelle méthode de criblages de conformations protéiques. Biophysique [physics.bio-ph]. Université de Lyon, 2020. Français. NNT : 2020LYSE1260 . tel-03434994

**HAL Id: tel-03434994**

**<https://theses.hal.science/tel-03434994>**

Submitted on 18 Nov 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

N°d'ordre NNT : 2020LYSE1260



**THESE de DOCTORAT DE L'UNIVERSITE DE LYON**  
opérée au sein de  
**l'Université Claude Bernard Lyon 1**

**Ecole Doctorale N° 160**  
**(Electronique, Electrotechnique, Automatique de Lyon)**

**Spécialité de doctorat** : Biophysique

Soutenue publiquement le 09/12/2020, par :  
**Laëtitia Bourgeat**

---

**Etude expérimentale de la dynamique  
moléculaire des protéines : Nouvelle  
méthode de criblages de conformations  
protéiques**

---

Devant le jury composé de :

Joly, Laurent	Professeur des Universités	Université de Lyon, UMR 5306	Président
Chomilier, Jacques Prévost, Chantal	Directeur de recherche CNRS Chargé de recherche	Sorbonne Université, UMR 7590 Université de Paris, UPR 9080	Rapporteur Rapporteuse
Barbi, Maria Correia, Natalia Rahbe, Yvan	Professeur des Universités Maître de conférences Directeur de recherche	Sorbonne Université, UMR 7600 Université de Lille, UMR 8207 INSA de Lyon, UMR 5240	Examinatrice Examinatrice Examineur
Lesieur, Claire Serghei, Anatoli	Chargé de recherche Chargé de recherche	Université de Lyon, UMR 5005 Université de Lyon, UMR 5223	Directrice de thèse Co-directeur de thèse



# **Université Claude Bernard – LYON 1**

Administrateur provisoire de l'Université	M. Frédéric FLEURY
Président du Conseil Académique	M. Hamda BEN HADID
Vice-Président du Conseil d'Administration	M. Didier REVEL
Vice-Président du Conseil des Etudes et de la Vie Universitaire	M. Philippe CHEVALLIER
Vice-Président de la Commission de Recherche	M. Jean-François MORNEX
Directeur Général des Services	M. Pierre ROLLAND

## **COMPOSANTES SANTE**

Département de Formation et Centre de Recherche en Biologie Humaine	Directrice : Mme Anne-Marie SCHOTT
Faculté d'Odontologie	Doyenne : Mme Dominique SEUX
Faculté de Médecine et Maïeutique Lyon Sud - Charles Mérieux	Doyenne : Mme Carole BURILLON
Faculté de Médecine Lyon-Est	Doyen : M. Gilles RODE
Institut des Sciences et Techniques de la Réadaptation (ISTR)	Directeur : M. Xavier PERROT
Institut des Sciences Pharmaceutiques et Biologiques (ISBP)	Directrice : Mme Christine VINCIGUERRA

## **COMPOSANTES & DEPARTEMENTS DE SCIENCES & TECHNOLOGIE**

Département Génie Electrique et des Procédés (GEP)	Directrice : Mme Rosaria FERRIGNO
Département Informatique	Directeur : M. Behzad SHARIAT
Département Mécanique	Directeur M. Marc BUFFAT
Ecole Supérieure de Chimie, Physique, Electronique (CPE Lyon)	Directeur : Gérard PIGNAULT
Institut de Science Financière et d'Assurances (ISFA)	Directeur : M. Nicolas LEBOISNE
Institut National du Professorat et de l'Education	Administrateur Provisoire : M. Pierre CHAREYRON
Institut Universitaire de Technologie de Lyon 1	Directeur : M. Christophe VITON
Observatoire de Lyon	Directrice : Mme Isabelle DANIEL
Polytechnique Lyon	Directeur : Emmanuel PERRIN
UFR Biosciences	Administratrice provisoire : Mme Kathrin GIESELER
UFR des Sciences et Techniques des Activités Physiques et Sportives (STAPS)	Directeur : M. Yannick VANPOULLE
UFR Faculté des Sciences	Directeur : M. Bruno ANDRIOLETTI





# Remerciements

Je tiens à remercier dans un premier temps tout particulièrement mes deux directeurs de thèse : Madame Claire LESIEUR et Monsieur Anatoli SERGHEI d'avoir cru en moi et de m'avoir offert la possibilité de travailler sur un sujet aussi innovant que passionnant, merci pour votre soutien et vos conseils.

Un grand merci aux laboratoires AMPERE (UMR 5005) et Ingénierie des Matériaux Polymères (UMR 5223) de m'avoir accueilli durant ces trois années de thèse.

Je remercie Monsieur Jacques CHOMILIER et Madame Chantal PREVOST d'avoir accepté d'être les rapporteurs de ce travail.

Je remercie Monsieur Laurent JOLY d'avoir présidé le jury de soutenance.

Je remercie Madame Maria BARBI, Madame Natalia CORREIA et Monsieur Yvan RAHBE d'avoir examiné ce travail.

Je remercie tous mes collègues Sofia, Rania, Etienne, Lorenza, Mathilde, Laroussi... pour leurs bonnes humeurs et leurs amitiés.

Enfin je remercie mes proches en particulier ma sœur Audrey, mes parents Jocelyne et Gérard qui m'apporte amour, soutien et sans lesquels je n'en serais pas là aujourd'hui.

---

## **Etude expérimentale de la dynamique moléculaire des protéines : Nouvelle méthode de criblages de conformations protéiques**

---

Les protéines sont des molécules fortement étudiées à cause de leurs différents rôles et de l'impact de leurs dysfonctionnements sur l'état de santé des organismes vivants. Ces dernières années, des études ont montré qu'une protéine ne possède pas une seule conformation (structure 3D) pour répondre à sa fonction, mais plusieurs. Au jour d'aujourd'hui, différentes techniques existent pour détecter et déterminer la conformation protéique, toutefois aucune technique ne permet de reconstituer l'ensemble des conformations adoptées par une protéine en criblant individuellement chacune des conformations qui le constitue.

Le travail de ce mémoire est de déterminer si la combinaison du nanoconfinement et de la spectroscopie diélectrique à large bande est utilisable pour cribler différentes conformations protéiques et ainsi essayer de reconstituer l'ensemble conformationnel d'une protéine. Le nanoconfinement est utilisé pour réduire l'hétérogénéité conformationnelle qui existe au sein d'une population à l'échelle macroscopique. La spectroscopie diélectrique à large bande caractérise la dynamique moléculaire qui est étroitement liée à la conformation protéique. Cette méthode peut donc être utilisée comme outil pour étudier l'impact d'un changement de conformation sur la dynamique moléculaire.

Mon travail de thèse a démontré que la combinaison du nanoconfinement et de la spectroscopie diélectrique à large bande pouvait être utilisée pour cribler différentes conformations protéiques en modifiant les dimensions du confinement. Différentes conformations des sous-unités B pentamériques de la toxine du choléra (CtxB<sub>5</sub>), de l'enterotoxine thermolabile (LTB<sub>5</sub>), et du lysozyme ont été observées. Les différences entre les signaux diélectriques des deux toxines révèlent des structures locales de dynamiques distinctes, alors que les différences avec le lysozyme révèlent des structures globales de dynamiques distinctes. Ces résultats ouvrent des perspectives pour diagnostiquer des impacts structuraux et/ou dynamiques liés à des mutations, dont l'implication dans le développement de maladies ou la conception de médicaments pourra être explorées dans le cadre de thérapies personnalisées.

Mots clés : protéine, nanoconfinement, spectroscopie diélectrique à large bande, conformation, dynamique moléculaire, crible

---

## **Experimental investigation of protein molecular dynamics : new method of protein conformation screening**

---

Proteins are highly studied molecules because of their different roles and the impact of their dysfunction on the state of health of living organisms. In recent years, studies have shown that a protein doesn't have a single conformation (3D structure) to respond to its function, but several. As of today, different techniques exist to detect and determine the protein conformation, no technique can reconstitute all the conformations studied by a protein by individually screening the conformations that constitute it.

The work of this memory is to determine if the combination of nanoconfinement and broadband dielectric spectroscopy can be used to screen different protein conformations and thus try to reconstruct the conformational set of a protein. Nanoconfinement is used to reduce the conformational heterogeneity that exists within a population on a macroscopic scale. Broadband dielectric spectroscopy characterizes molecular dynamics which are related to protein conformation. This method can therefore be used as a tool to study the impact of a change in conformation on molecular dynamics.

My thesis work demonstrated that the combination of nanoconfinement and broadband dielectric spectroscopy could be used to screen different protein conformations by modifying the dimensions of the confinement. Different conformations of pentameric B subunits of cholera toxin (CtxB<sub>5</sub>), heat labile enterotoxin (LTB<sub>5</sub>), and lysozyme have been observed. The differences between the dielectric signals of the two toxins reveal local structures with distinct dynamics, while the differences with lysozyme reveal global structures with distinct dynamics. These results open perspectives for diagnosing structural and / or dynamic impacts linked to mutations, the implication of which in the development of diseases or the design of drugs can be explored within the framework of personalized therapies.

Keywords: protein, nanoconfinement, broadband dielectric spectroscopy, conformation, molecular dynamics, screen

# Table des matières

Liste des abréviations.....	11
Listes des figures.....	12
Listes des tables.....	16
Chapitre 1 Introduction.....	17
Chapitre 2 Etat de l'art .....	20
2.1 Protéines.....	20
2.2 Détermination de la conformation protéique.....	34
2.2.1 Cristallographie par diffraction de rayon X.....	34
2.2.2 Résonance magnétique nucléaire.....	35
2.3 Technique pour détecter une molécule unique.....	38
2.3.1 Pince optique.....	38
2.3.2 Microscopie à force atomique .....	39
2.3.3 Single-molecule fluorescence Resonance Energy : SmFRET .....	39
2.4 Nanoconfinement et la spectroscopie diélectrique à large bande .....	41
2.4.1 Nanoconfinement .....	41
2.4.2 Spectroscopie diélectrique à large bande (BDS) .....	44
2.4.3 Combinaison nanoconfinement et BDS .....	45
Chapitre 3 Protéines, Méthodes, Méthodologie .....	47
3.1 Protéines étudiées.....	47
3.1.1 Sous-unité B pentamérique de la toxine du choléra.....	47
3.1.2 Sous-unité B pentamérique de l'entérotoxine thermolabile.....	48
3.1.3 Lysozyme .....	49
3.1.4 Choix des protéines .....	49
3.2 Méthodes.....	50
3.2.1 Membrane AAO.....	50
3.2.2 Spectroscopie diélectrique à large bande.....	51
3.2.3 Analyse des données de pertes diélectriques.....	58

3.3	Méthodologie .....	59
3.3.1	Membrane avec des nanopores.....	59
3.3.2	Mesure diélectrique par spectroscopie diélectrique à large bande .....	60
3.3.3	Echantillon protéique .....	61
3.3.4	Essais diélectriques.....	64
3.3.5	Analyse du spectre diélectrique .....	66
Chapitre 4 Combinaison du nanoconfinement et de la spectroscopie diélectrique à la large bande pour l'étude de la dynamique des structures protéiques .....		68
4.1	Un signal diélectrique des protéines est-il détecté ? .....	70
4.1.1	Test nanomembrane vide .....	70
4.1.2	Test nanomembrane avec le solvant .....	70
4.1.3	Test nanomembrane avec CtxB <sub>5</sub> .....	74
4.2	Analyse de CtxB <sub>5</sub> à différentes fréquences .....	78
4.3	Analyse du comportement de CtxB <sub>5</sub> en fonction de la température.....	80
4.4	Analyse du comportement de CtxB <sub>5</sub> en fonction de la fréquence.....	85
4.5	Interprétation du signal diélectrique.....	92
Chapitre 5 Impact du confinement : Variation de la longueur .....		99
5.1	Confinement à 40 nm de diamètre et 5 µm de long.....	101
5.1.1	Contrôle de l'échantillon protéique .....	101
5.1.2	Analyse de CtxB <sub>5</sub> à différentes fréquences .....	103
5.1.3	Analyse du comportement de CtxB <sub>5</sub> en fonction de la température .....	104
5.1.4	Analyse du comportement de CtxB <sub>5</sub> en fonction de la fréquence .....	106
5.2	Confinement à 40 nm de diamètre et 2 µm de long.....	110
5.2.1	Contrôle de l'échantillon protéique .....	110
5.2.2	Analyse de CtxB <sub>5</sub> à différentes fréquences .....	111
5.2.3	Analyse du comportement de CtxB <sub>5</sub> en fonction de la température .....	113
5.2.4	Analyse du comportement de CtxB <sub>5</sub> en fonction de la fréquence .....	115
5.3	Comparaison.....	119
5.3.1	Effet de la déshydratation .....	119
5.3.2	Effet de la variation de longueur du confinement.....	120

Chapitre 6 Impact du confinement : Variation du diamètre.....	127
6.1 Confinement à 120 nm de diamètre et 2 $\mu\text{m}$ de long.....	127
6.1.1 Contrôle de l'échantillon protéique.....	127
6.1.2 Analyse de CtxB <sub>5</sub> à différentes fréquences.....	128
6.1.3 Analyse du comportement de CtxB <sub>5</sub> en fonction de la température.....	129
6.1.4 Analyse du comportement de CtxB <sub>5</sub> en fonction de la fréquence.....	131
6.2 Confinement à 20 nm de diamètre et 2 $\mu\text{m}$ de long.....	133
6.2.1 Contrôle de l'échantillon protéique.....	133
6.3 Comparaison.....	136
6.3.1 Effet de la déshydratation.....	136
6.3.2 Effet de la variation du diamètre du confinement.....	137
 Chapitre 7 Etude sur d'autres protéines.....	 146
7.1 Etude d'un variant de séquence : protéine LTB <sub>5</sub> .....	147
7.1.1 Contrôle de l'échantillon protéique.....	148
7.1.2 Analyse du comportement de LTB <sub>5</sub> en fonction de la température.....	151
7.1.3 Analyse du comportement de LTB <sub>5</sub> en fonction de la fréquence.....	152
7.1.4 Comparaison de deux variants de séquence : LTB <sub>5</sub> et CtxB <sub>5</sub> .....	153
7.2 Etude du lysozyme.....	166
7.2.1 Contrôle de l'échantillon protéique.....	166
 Chapitre 8 Discussion, Conclusions, Perspectives.....	 175
 Références.....	 181
Annexes.....	193
Annexe A. Experimental Protein Molecular Dynamics : Broadband Dielectric Spectroscopy coupled with nanoconfinement.....	193
Annexe B. Analysis of Nanoconfined Protein Dielectric Signals Using Charged Amino Acid Network Models.....	205

# Liste d'abréviations

ADN	: acide désoxyribonucléique
ARNm	: acide ribonucléique messenger
BDS	: broadband dielectric spectroscopy : spectroscopie diélectrique à large bande
RMS	: root mean square : moyenne quadratique
DSC	: differential scanning calorimetry : calorimétrie différentielle à balayage
RMN	: résonance magnétique nucléaire
PDB	: protein data bank : banque de donnée des protéines
AFM	: atomic force microscopy : microscopie à force atomique
FRET	: fluorescence resonance energy transfert : transfert d'énergie par résonance de fluorescence
SmFRET	: single-molecule fluorescence resonance energy transfert : transfert d'énergie par résonance de fluorescence d'une molécule unique
Membrane AAO	: anodized aluminium oxide membranes
CT	: toxine du choléra
CtxB	: sous-unité B de la choléra toxine
CtxA	: sous-unité A de la choléra toxine
CtxB <sub>5</sub>	: sous-unité B pentamérique de la choléra toxine
GM <sub>1</sub>	: monosialotetrahexosylganglioside
AMPC	: adénosine monophosphate cyclique
LT	: entérotoxine thermolabile
LTB	: sous-unité B de l'entérotoxine thermolabile
LTB <sub>5</sub>	: sous-unité B pentamérique de l'entérotoxine thermolabile
PBS	: phosphate buffered saline : solution saline dans un tampon phosphate
MEB	: microscopie électronique à balayage
Trp	: tryptophane
R1 60 °C	: courbe de refroidissement de 60 °C à -80 °C pour le premier cycle à 60 °C
R2 60 °C	: courbe de refroidissement de 60 °C à -80 °C pour le deuxième cycle à 60 °C
R 80 °C	: courbe de refroidissement de 80 °C à -80 °C
R 100 °C	: courbe de refroidissement de 100 °C à -80 °C
R 140 °C	: courbe de refroidissement de 140 °C à -80 °C
R 180 °C	: courbe de refroidissement de 180 °C à -80 °C
MDb	: dynamique moléculaire du solvant
MD <sub>1</sub>	: pentamère de type natif
MD <sub>2</sub>	: intermédiaires d'assemblage et/ou monomères repliés
MD <sub>3</sub>	: pentamère partiellement replié ou déplié
$\theta_{max}$	: température où l'intensité du pic de relaxation est maximale : position du pic de relaxation en température



# Liste des figures

Figure 2-1. Représentation de la structure développée d'un acide aminé. ....	20
Figure 2-2. Représentation de l'ionisation des groupements amine et acide carboxylique d'un acide aminé. ....	22
Figure 2-3. Formation d'une chaîne polypeptidique d'acide aminé.....	26
Figure 2-4. Représentation de la formation de structures secondaires.....	28
Figure 2-5. Représentation de la structure tertiaire de la chaîne D de la sous-unité B pentamérique de la choléra toxine via un programme de visualisation (swiss-pdb).....	29
Figure 2-6. Représentation de la structure quaternaire d'une protéine grâce au programme de visualisation swiss-pdb.....	30
Figure 2-7. Représentation de l'organisation temporelle d'une conformation protéique. ....	30
Figure 2-8. Représentation de différentes positions de la chaîne latérale du tryptophane ...	31
Figure 2-9. Représentation de l'impact d'une mutation sur les interactions atomiques. ....	33
Figure 2-10. Enrichissement ou sélection des conformations protéiques via le confinement. ....	42
Figure 3-1. Sous-unité B pentamérique de la choléra toxine (code PDB : 1EEI).....	48
Figure 3-2. Sous-unité B pentamérique de l'entérotoxine thermolabile (PDB : 1LTR).....	48
Figure 3-3. Lysozyme (code PDB : 1DPX).....	49
Figure 3-4. Membrane avec des nanopores cylindriques, parallèles, identiques et ordonnés .....	50
Figure 3-5. Formation d'une couche d'oxyde entre la plaque d'aluminium et la solution d'électrolyte. ....	51
Figure 3-6. Représentation de la polarisation électronique. ....	53
Figure 3-7. Représentation de la polarisation atomique. ....	53
Figure 3-8. Représentation des moments dipolaires.....	54
Figure 3-9. Représentation de la polarisation d'orientation.....	54
Figure 3-10. Détection des polarisations dans le temps suite à l'application d'un champ électrique. ....	55
Figure 3-11. Représentation du complexe de permittivité en fonction de la fréquence. ....	56
Figure 3-12. Représentation des données de pertes diélectriques en fonction de la fréquence et de la température. ....	57
Figure 3-13. Représentation de la dépendance en température d'un processus de relaxation en fonction des temps de relaxation. ....	58
Figure 3-14. Nanomembrane achetée chez InRedox59	
Figure 3-15. Photos de l'appareil de Spectroscopie Diélectrique à Large Bande. ....	61
Figure 3-16. Rampe de température utilisée dans les expériences.....	65
Figure 3-17. Variation de l'intensité du pic de relaxation en augmentant la quantité de copolymères dans les nanopores.....	66
Figure 3-18. Variation de la position du pic de relaxation suite à un changement de conformation.....	67

Figure 4-1. Nanomembrane vide avec un confinement à 40 nm de diamètre et 10 $\mu\text{m}$ de long .....	70
Figure 4-2. Superposition des courbes de refroidissement du solvant à différentes fréquences. ....	71
Figure 4-3. Superpositions des courbes de refroidissement et de chauffages mesurées à différentes fréquences après différents traitements thermiques pour le contrôle du solvant. ....	73
Figure 4-4. Superposition des courbes de refroidissement et de chauffage mesurées à la fréquence de $10^3$ Hz après le premier traitement thermique à 60 $^{\circ}\text{C}$ .....	75
Figure 4-5. Superposition des courbes de refroidissement et de chauffage mesurées à la fréquence de $10^3$ Hz après des traitements thermiques à 60 $^{\circ}\text{C}$ .....	76
Figure 4-6. Reproduction de l'expérience avec un échantillon protéique à la fréquence de $10^3$ Hz.....	77
Figure 4-7. CtxB <sub>5</sub> dans le confinement à 40 nm de diamètre et 10 $\mu\text{m}$ de long avec superposition des courbes de refroidissement à la fréquence de $10^3$ Hz de la protéine, du solvant et de la nanomembrane vide .....	77
Figure 4-8. Analyse de CtxB <sub>5</sub> à différentes fréquences après le traitement thermique à 60 $^{\circ}\text{C}$ . ....	79
Figure 4-9. Etude de CtxB <sub>5</sub> à la fréquence de $10^3$ Hz après différents traitements thermiques .....	81
Figure 4-10. Etude du comportement de CtxB <sub>5</sub> après différents traitements thermiques.....	82
Figure 4-11. Test de réversibilité pour CtxB <sub>5</sub> .....	85
Figure 4-12. Analyse de la dépendance en fréquence des processus de relaxation après différents traitements thermiques. ....	87
Figure 4-13. Dépendance en température des temps de relaxation après différents traitements thermiques. ....	89
Figure 4-14. Représentation des interfaces dans CtxB .....	95
Figure 4-15. Modèle de la dénaturation thermique de CtxB <sub>5</sub> .....	96
Figure 4-16. Résumé des paramètres pour analyser le signal diélectrique détecté.....	98
Figure 5-1. CtxB <sub>5</sub> dans le confinement à 40 nm de diamètre et 5 $\mu\text{m}$ de long.....	101
Figure 5-2. Analyse du comportement du solvant à la fréquence de 1 Hz après différents traitements thermiques. ....	102
Figure 5-3. Comparaison du signal diélectrique pour les courbes de refroidissement du solvant et de l'échantillon protéique à la fréquence de 1 Hz après différents traitements thermiques .....	103
Figure 5-4. Analyse de CtxB <sub>5</sub> à différentes fréquences après le deuxième traitement thermique à 60 $^{\circ}\text{C}$ .....	104
Figure 5-5. Etude du comportement de CtxB <sub>5</sub> après différents traitements thermiques.....	105
Figure 5-6. Analyse de la dépendance en fréquence des processus de relaxation pour le confinement à 40 nm de diamètre et 5 $\mu\text{m}$ de long après différents traitements thermiques. ....	107
Figure 5-7. Dépendance en température des temps de relaxation dans le confinement à 40 nm de diamètre et 5 $\mu\text{m}$ de long après différents traitements thermiques .....	109

Figure 5-8. CtxB <sub>5</sub> dans le confinement à 40 nm de diamètre et 2 µm de long à la fréquence de 10 <sup>3</sup> Hz.....	111
Figure 5-9. Analyse de CtxB <sub>5</sub> à différentes fréquences.....	112
Figure 5-10. Etude du comportement de CtxB <sub>5</sub> après différents traitements thermiques...	114
Figure 5-11. Analyse de la dépendance en fréquence des processus de relaxation dans le confinement à 40 nm de diamètre et 2 µm de long après différents traitements thermiques .....	117
Figure 5-12. Dépendance en température des temps de relaxation dans le confinement à 40 nm de diamètre et 2 µm de long après différents traitements thermiques .....	118
Figure 5-13. Superposition des courbes de refroidissement et de chauffage à la fréquence de 10 <sup>3</sup> Hz après deux traitements thermiques à 60 °C .....	120
Figure 5-14. Comparaison des processus de relaxation mesurés à la fréquence de 10 <sup>3</sup> Hz en faisant varier la longueur du confinement après le deuxième traitement thermique à 60 °C .....	121
Figure 5-15. Comparaison des processus de relaxation détectés pour différentes longueurs de confinement après les traitements thermiques à 140 °C et 180 °C.....	122
Figure 5-16. Reproduction des expériences des confinements de longueurs différentes à la fréquence de 10 <sup>3</sup> Hz.....	126
Figure 6-1. CtxB <sub>5</sub> dans le confinement à 120 nm de diamètre et 2 µm de long.....	128
Figure 6-2. Analyse de CtxB <sub>5</sub> à différentes fréquences après le deuxième traitement thermique à 60 °C.....	129
Figure 6-3. Etude du comportement de CxtB <sub>5</sub> après différents traitements thermiques.....	131
Figure 6-4. Analyse de la dépendance en fréquence des processus de relaxation détectés dans le confinement à 120 nm de diamètre et 2 µm de long après différents traitements thermiques .....	133
Figure 6-5. Analyse de CtxB <sub>5</sub> dans le confinement à 20 nm de diamètre et 2 µm de long à différentes fréquences après le deuxième traitement thermique à 60 °C.....	134
Figure 6-6. Reproduction du signal diélectrique de CtxB <sub>5</sub> à la fréquence de 10 <sup>3</sup> Hz dans le confinement à 20 nm de diamètre et 2 µm de long après le traitement thermique à 60 °C . .....	134
Figure 6-7. Superposition des courbes de refroidissement et de chauffage à la fréquence de 1 Hz après deux traitements thermiques à 60 °C .....	136
Figure 6-8. Comparaison des processus de relaxation obtenus à la fréquence de 1 Hz en faisant varier le diamètre du confinement après le deuxième traitement thermique à 60 °C.....	137
Figure 6-9. Reproduction des expériences des confinements de diamètres différents à la fréquence de 1 Hz .....	141
Figure 6-10. Analyse du comportement de CtxB <sub>5</sub> dans le confinement à 15 nm de diamètre et 5 µm de long.....	142
Figure 6-11. Comparaison du signal de CtxB <sub>5</sub> à la fréquence de 10 <sup>3</sup> Hz en diminuant le diamètre de confinement après le traitement thermique à 60 °C.....	143
Figure 6-12. Etude de CtxB <sub>5</sub> avec différents pourcentages de remplissage dans le confinement à 15 nm de diamètre et 5 µm de long à la fréquence de 10 <sup>3</sup> Hz après le traitement thermique à 60 °C.....	144

Figure 7-1. Alignement de séquence des toxines : CtxB <sub>5</sub> (code PDB :1EEI) et LTB <sub>5</sub> (code PDB :1LTR).....	147
Figure 7-2. Alignement des conformations protéiques de CtxB <sub>5</sub> (code PDB : 1EEI) et LTB <sub>5</sub> (code PDB : 1LTR) avec l'application de visualisation de structure macromoléculaire 3D Cn3D....	148
Figure 7-3. LTB <sub>5</sub> dans le confinement à 40 nm de diamètre et 10 µm de long à la fréquence de 10 <sup>3</sup> Hz après un deuxième traitement thermique à 60 °C.....	149
Figure 7-4. Analyse de LTB <sub>5</sub> dans le confinement à 40 nm de diamètre et 10 µm de long à la fréquence de 10 <sup>3</sup> Hz. ....	150
Figure 7-5. Analyse de LTB <sub>5</sub> à différentes fréquences après le deuxième traitement thermique à 60 °C.....	150
Figure 7-6. Analyse de LTB <sub>5</sub> à différentes fréquences après le traitement thermique à 80 °C. ....	151
Figure 7-7. Etude du comportement de LTB <sub>5</sub> à la fréquence de 10 <sup>3</sup> Hz après différents traitements thermiques dans le confinement de 40 nm de diamètre et 10 µm de long.....	152
Figure 7-8. Analyse de la dépendance en fréquence des processus de relaxation de LTB <sub>5</sub> dans le confinement de 40 nm de diamètre et 10 µm de long après différents traitements thermiques .....	153
Figure 7-9. Comparaison de la réponse diélectrique pour CtxB <sub>5</sub> et LTB <sub>5</sub> dans le confinement à 40 nm de diamètre et 10 µm de long après le traitement thermique à 60 °C. ....	154
Figure 7-10. Comparaison de la réponse diélectrique pour CtxB <sub>5</sub> et LTB <sub>5</sub> dans le confinement à 40 nm de diamètre et 10 µm de long après le traitement thermique à 80 °C .....	155
Figure 7-11. Comparaison de la réponse diélectrique pour CtxB <sub>5</sub> et LTB <sub>5</sub> dans le confinement à 40 nm de diamètre et 10 µm de long après le traitement thermique à 140 °C. ....	156
Figure 7-12. Analyse de LTB <sub>5</sub> dans le confinement à 40 nm de diamètre et 2 µm de long à la fréquence de 1 Hz après le deuxième traitement thermique à 60 °C.....	163
Figure 7-13. Etude de LTB <sub>5</sub> à la fréquence de 1 Hz après différents traitements thermiques dans le confinement à 40 nm de diamètre et 2 µm de long.....	164
Figure 7-14. Analyse de la dépendance en fréquence des processus de relaxation de LTB <sub>5</sub> dans le confinement à 40 nm de diamètre et 2 µm de long après différents traitements thermiques .....	165
Figure 7-15. Reproduction de l'expérience LTB <sub>5</sub> dans le confinement à 40 nm de diamètre et 2 µm de long à la fréquence de 1 Hz après le traitement thermique à 60 °C. ....	166
Figure 7-16. Lysozyme dans le confinement à 40 nm de diamètre et 5 µm de long à différentes fréquences.....	167
Figure 7-17. Lysozyme dans le confinement à 40 nm de diamètre et 5 µm de long à la fréquence de 10 <sup>3</sup> Hz après le traitement thermique à 60 °C .....	168
Figure 7-18. Comparaison des signaux diélectriques du solvant et du lysozyme après le deuxième traitement thermique à 60 °C dans le confinement à 40 nm de diamètre et 5 µm de long.....	169
Figure 7-19. Comparaison des processus de relaxation du solvant et du lysozyme à différentes fréquences pour le confinement à 40 nm de diamètre et 5 µm de long.....	170
Figure 7-20. Analyse du lysozyme dans le confinement à 40 nm de diamètre et 5 µm de long à la fréquence de 10 <sup>3</sup> Hz. ....	173

# Liste des tables :

Table 2-1. Liste des vingt acides aminés naturels présent dans une protéine.....	25
Table 2-2. Récapitulatif des techniques de cristallographie et de RMN.....	37
Table 3-1. Valeur d'électronégativité des atomes présents dans une protéine déterminée avec l'échelle de Pauling.....	53
Table 3-2. Dimensions des confinements utilisées.....	60
Table 5-1. Intensité et position des pics de relaxation en température associée au processus de relaxation MD <sub>1</sub> à la fréquence de 10 <sup>3</sup> Hz après différents traitements thermiques pour différents confinements en variant la longueur .....	122
Table 5-2. Intensité et position des pics de relaxation en température associée au processus de relaxation MD <sub>3</sub> à la fréquence de 10 <sup>3</sup> Hz après différents traitements thermiques pour différents confinements en variant la longueur .....	123
Table 5-3. Intensité et position des pics de relaxation en température des trois processus de relaxation détectés dans le confinement à 40 nm de diamètre et 5 µm de long à la fréquence de 1 Hz.....	125
Table 6-1. Intensité et position des pics de relaxation à la fréquence de 1 Hz associée au processus de relaxation MD <sub>1</sub> après différents traitements thermiques pour différents confinements en variant le diamètre.....	138
Table 6-2. Intensité et position du processus de relaxation en température associée au processus de relaxation MD <sub>3</sub> à la fréquence de 1 Hz après différents traitements thermiques pour différents confinements en variant le diamètre.....	139
Table 7-1. Représentation des dipôles ioniques détectés dans les interfaces de LTB <sub>5</sub> et CtxB <sub>5</sub> .....	159
Table 7-2. Représentation des dipôles ioniques intramoléculaires détectés dans LTB <sub>5</sub> et CtxB <sub>5</sub> .....	160
Table 7-3. Table récapitulative de la position du premier processus de relaxation en température détecté pour l'échantillon de lysozyme et pour l'échantillon du solvant.....	171
Table 7-4. Table récapitulative de la position du deuxième processus de relaxation en température détecté pour l'échantillon du lysozyme et pour l'échantillon du solvant.....	172

# Chapitre 1

## Introduction

Les protéines sont définies par leurs séquences, leurs conformations, leurs dynamiques et leurs fonctions. Les conformations protéiques, les dynamiques moléculaires et les fonctions protéiques peuvent-être altérées suite à une ou des mutations dans la séquence protéique. Pour comprendre l'origine d'une mutation, il est nécessaire de remonter au niveau de l'ADN (acide désoxyribonucléique).

L'ADN est composé de deux brins, chacun constitués de quatre types de nucléotides. Ces deux brins d'ADN sont complémentaires, l'un avec l'autre. Un brin d'ADN est transcrit en ARNm (acide ribonucléique messager) complémentaire au brin d'ADN transcrit et de séquence identique au brin d'ADN non transcrit. Trois nucléotides successifs (codon) dans l'ARNm sont associés à un acide aminé permettant la traduction de l'ARNm en séquence protéique. Un codon correspond à un seul acide aminé, mais un acide aminé peut-être codé par plusieurs codons (redondance). Lorsque les trois nucléotides forment un codon stop (codon associé à aucun acide aminé), la traduction de la protéine s'arrête. La séquence protéique résulte de l'information génétique.

La séquence d'ADN peut subir différents types de mutations : des mutations silencieuses où le changement de nucléotide n'impacte pas la séquence protéique à cause de la redondance du code génétique, des mutations non-sens où le changement de nucléotide peut remplacer le codon spécifique d'un acide aminé par un codon stop et les mutations faux-sens où un nucléotide est remplacé par un autre nucléotide entraînant un changement d'acide aminé dans la séquence protéique. La diversité des mutations peut fournir des séquences protéiques alternatives fonctionnelles. D'autre part, la diversité des mutations peut avoir des impacts négatifs conduisant à la perte de la fonction protéique, à la résistance aux médicaments etc...

La présence de mutations dans une séquence protéique est détectable grâce aux méthodes de séquençage, l'impact des mutations sur la conformation protéique est détectable grâce aux techniques de cristallographie par rayon X et la résonance magnétique nucléaire et l'impact sur la fonction est détectable via des tests fonctionnels comme les tests Elisa.

Toutefois, la fonction protéique n'est pas définie pas une conformation unique mais par un ensemble de conformations à partir d'une séquence protéique. Cette diversité conformationnelle explique pourquoi une mutation n'a pas forcément d'impact sur la fonction

protéique. Si la conformation protéique issue de la séquence protéique mutée correspond à une conformation de l'ensemble conformationnel de la protéine sauvage (issue de la séquence protéique non mutée), alors la fonction protéique ne devrait pas être impactée malgré la ou les mutations dans la séquence. Inversement, si la conformation protéique issue de la séquence protéique mutée n'appartient pas à l'ensemble de conformations de la protéine sauvage alors la ou les mutations devraient avoir un impact sur la fonction protéique. Cet impact peut être positif (nouvelle fonction) ou négatif (perte de fonction). Il est donc important de caractériser l'ensemble des conformations associées à une protéine de séquence donnée pour pouvoir comprendre/ mesurer les conséquences structurales d'une ou plusieurs mutations et ainsi anticiper son impact sur la fonction.

Au jour d'aujourd'hui différentes techniques permettent d'étudier la conformation d'une protéine, toutefois, aucune technique ne permet de cribler l'ensemble conformationnel d'une protéine. Ce mémoire a pour objectif de déterminer si la combinaison du nanoconfinement et de la spectroscopie diélectrique permet de cribler différentes conformations protéiques. Le nanoconfinement permet d'avoir des dimensions à l'échelle de la protéine et devrait donc enrichir et/ou sélectionner une population de conformation protéique. Les dimensions du nanoconfinement dépendent de la longueur et du diamètre des nanopores. La modification des dimensions du confinement devrait ainsi enrichir et/ou sélectionner différentes populations de conformations et donc réaliser un criblage des conformations d'une protéine. La spectroscopie diélectrique à large bande mesure des fluctuations dipolaires. Une protéine possède une organisation multi-échelle dans le temps, avec des mouvements de chaînes latérales, de structures secondaires, de dépliement et de repliement. Ces mouvements dépendent de la conformation protéique. De plus, la conformation protéique repose sur l'interaction d'un ensemble d'atomes entre eux. Par conséquent, la caractérisation de fluctuations dipolaires devrait permettre d'obtenir des informations sur la conformation de la population conformationnelle enrichie et/ou sélectionnée.

Cette approche combinant le nanoconfinement et la spectroscopie diélectrique à large bande n'a jamais été utilisée sur des protéines. Ainsi, plusieurs questions se posent pour établir l'utilité de cette approche pour l'étude des protéines :

- Cette approche est-elle applicable aux protéines ? Les protéines possèdent-elles un signal diélectrique caractéristique de leur conformation ?
- Différentes conformations protéiques peuvent-elles être sélectionnées en modifiant les dimensions du confinement ?
- Des variants de séquence (séquences mutées) peuvent-ils être détectés et distingués des conformations de la protéine sauvage ?

Pour répondre à ces questions ce mémoire est divisé en plusieurs chapitres :

- Le chapitre 2 intitulé « Etat de l'art » explique les propriétés essentielles des protéines, de l'échelle des acides aminés à l'échelle des conformations protéiques. Ce chapitre se consacrera aussi aux différentes techniques généralement utilisées pour déterminer la conformation protéique, et aux techniques utilisées pour étudier des molécules



uniques. Enfin, le principe du nanoconfinement et de la spectroscopie diélectrique à large bande sera énoncé.

- Le chapitre 3 : Protéines, Méthodes, Méthodologie explique le choix des protéines, les méthodes utilisées et la méthodologie pour préparer les échantillons protéiques.
- Le chapitre 4 étudie une protéine (sous-unité B pentamérique de la choléra toxine : CtxB<sub>5</sub>) dans un confinement donné pour déterminer si la combinaison du nanoconfinement et de la spectroscopie diélectrique à large bande détecte un signal caractéristique de la conformation protéique.
- Les chapitres 5 et 6 analysent la protéine CtxB<sub>5</sub> dans différents confinements en variant la longueur (chapitre 5) et le diamètre (chapitre 6) du confinement pour déterminer si le nanoconfinement permet d'enrichir et/ou sélectionner différentes conformations protéiques et si ces différentes conformations sont détectées par spectroscopie diélectrique à large bande.
- Le chapitre 7 détermine si cette approche permet de détecter l'impact des mutations sur la conformation protéique. Dans un premier temps, une protéine possédant une conformation et une fonction protéique similaire à CtxB<sub>5</sub> sera étudiée pour déterminer si des variants de séquence peuvent être détectés et distingués (modification locale des conformations protéiques), dans un deuxième temps une protéine avec une conformation et une fonction protéique différente sera analysée pour déterminer si des modifications globales sont aussi détectées et distinguables.



# Chapitre 2

## Etat de l'art

### 2.1 Protéines

Les protéines sont des molécules qui jouent des rôles majeurs dans les organismes vivants. Les protéines exercent différentes fonctions comme :

- Fonction de structure : le collagène, la kératine ... permettant aux cellules et aux tissus de maintenir leur organisation dans l'espace.
- Fonction de transport : l'hémoglobine permet de transporter l'oxygène dans le sang.
- Fonction de défense immunitaire : les anticorps permettent à l'organisme de se défendre contre les pathogènes (virus, bactéries, microbes ...).
- Fonction de mouvement : la contraction des muscles est due à l'interaction entre deux protéines : l'actine et la myosine.
- Fonction enzymatique : la pepsine est une protéine enzymatique présente naturellement dans les sucs gastriques qui aident à la digestion des aliments en dégradant les protéines dans l'estomac.
- Fonction hormonale : l'insuline est une protéine qui participe à la régulation de la glycémie en stimulant la synthèse de glycogène.

La fonction protéique est assurée via une forme (conformation) encodée dans une séquence d'acide aminé.

Une protéine est composée de différents acides aminés. Il existe plus d'une centaine d'acides aminés, toutefois seul vingt sont présents dans les protéines. Chaque acide aminé est composé d'un atome de carbone, appelé carbone alpha ( $C\alpha$ ), relié à un atome d'hydrogène, un groupement amine ( $NH_3^+$ ), un acide carboxylique ( $COOH$ ) et une chaîne latérale (R) (Fig. 2-1). Le carbone alpha des acides aminés est donc asymétrique (1).

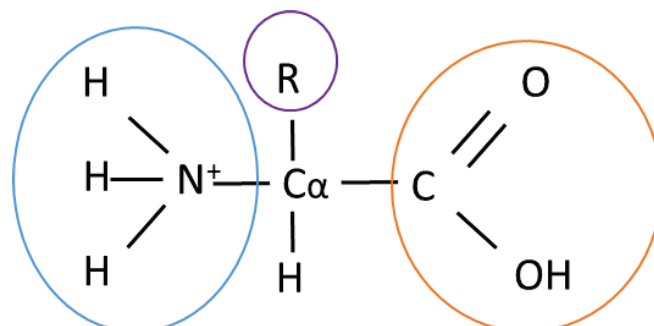


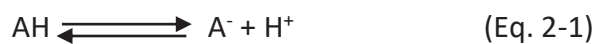
Figure 2-1. Représentation de la structure développée d'un acide aminé. Dans le cercle bleu, représentation du groupement amine, dans le cercle orange le groupement acide carboxylique et le cercle violet représente la chaîne latérale R.

Les acides aminés se distinguent par la composition de leur chaîne latérale (nommé R), c'est donc cette chaîne qui définit le type d'acide aminé. Les autres atomes présents dans tous les acides aminés, participent au squelette peptidique (backbone). Chaque acide aminé est désigné par un code à trois lettres et un code à une lettre. Par exemple, le code à trois lettres de l'acide aminé glycine est Gly et son code à une lettre est G (Table. 2-1).

La chaîne latérale R de chaque acide aminé est variable en taille mais aussi en chimie (polarité, hydrophobie et charge). La variation de taille de la chaîne latérale des différents acides aminés peut être illustrée par la glycine et le tryptophane. En effet, la chaîne latérale de la glycine est composée d'un seul atome d'hydrogène, alors que la chaîne latérale du tryptophane est composée d'un noyau indole (noyau benzène et noyau pyrrole). La glycine est le plus petit (75 Da soit 75 g/mol) acide aminé présent dans une protéine, alors que le tryptophane (204 Da) est l'acide aminé le plus gros. Il est nécessaire de noter que la glycine est le seul acide aminé ne possédant pas de carbone alpha asymétrique.

Les groupements des chaînes latérales impactent la polarité et l'hydrophobie des acides aminés. Les acides aminés avec des chaînes latérales composées uniquement d'hydrocarbures (carbone et hydrogène) comme la glycine, leucine, phénylalanine sont des acides aminés hydrophobes. Les acides aminés possédant des chaînes latérales avec un groupement hydroxyle (OH) comme la sérine, thréonine et tyrosine sont des acides aminés polaires (du moins avec une partie polaire).

La chaîne latérale des acides aminés peut être chargée. La charge des acides aminés dépend du pH du milieu et du pKa des chaînes latérales. Les chaînes latérales chargées possèdent un pKa, soit une valeur de pH où la chaîne latérale est à 50 % sous forme acide (AH) et à 50 % sous forme basique (A<sup>-</sup>) (Eq. 2-1).



Si le pH du milieu est inférieur au pKa de la chaîne latérale, alors la chaîne latérale est majoritairement sous forme acide. Inversement, si le pH du milieu est supérieur à la valeur de pKa, la chaîne latérale est majoritairement sous forme basique. L'acide aspartique, l'acide glutamique, l'histidine, la lysine, l'arginine, la cystéine et la tyrosine sont les seuls acides aminés avec une chaîne latérale ionisable. La chaîne latérale des acides aminés peut être chargée négativement, comme par exemple pour l'acide glutamique. La chaîne latérale de l'acide glutamique possède une fonction acide carboxylique avec un pKa de 4,1. Par conséquent, si le pH du milieu est inférieur à 4,1, la chaîne latérale sera majoritairement sous forme acide (COOH) et inversement, si le pH du milieu est supérieur à 4,1, la chaîne latérale sera majoritairement sous forme basique (COO<sup>-</sup>). D'autres acides aminés comme la lysine ont des chaînes latérales positives.

La chaîne latérale de la lysine possède une fonction amine avec un pKa de 10,8, ainsi si le pH du milieu est inférieur au pKa, la chaîne latérale sera majoritairement sous forme NH<sub>3</sub><sup>+</sup> et inversement, si le pH du milieu est supérieur au pKa, la chaîne latérale sera majoritairement sous forme NH<sub>2</sub>. A pH 6,9, l'acide aspartique et l'acide glutamique sont chargés négativement (pKa : 4,1), la lysine (pKa : 10,8) et l'arginine (pKa : 12,5) sont chargés positivement, l'histidine (pKa : 6), cystéine (pKa : 8,3) et la tyrosine (pKa : 10,9) ne sont pas chargés (1). Comme le montre la figure 2-2, en plus des chaînes latérales, les fonctions amine (NH<sub>3</sub><sup>+</sup>) et acide

carboxylique (COOH) des acides aminés sont aussi ionisables à des pKa de 8 et 3,1 respectivement.

Les interactions entre deux acides aminés chargés proches spatialement (distance inférieure ou égale à 5 Å) formeront des dipôles ioniques, qui seront les principaux contributeurs des signaux diélectriques. Les dipôles ioniques dépendent donc du pH de l'échantillon protéique et de la position des acides aminés chargés dans la conformation protéique.

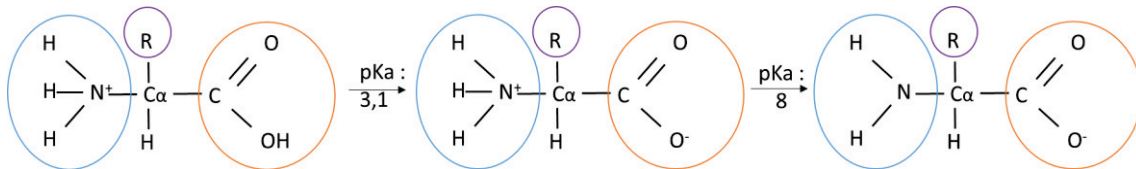
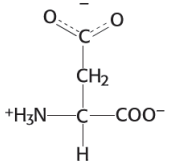
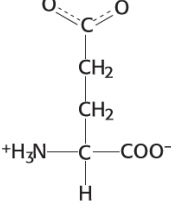
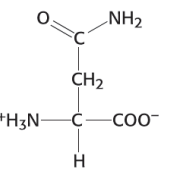
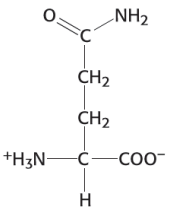
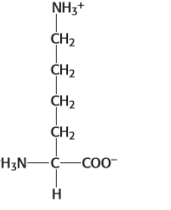
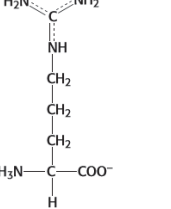
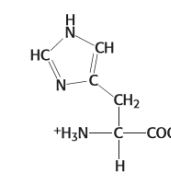


Figure 2-2. Représentation de l'ionisation des groupements amine et acide carboxylique d'un acide aminé. Représentation des trois états d'ionisation pour un acide aminé en fonction du pH si la chaîne latérale de l'acide aminé n'est pas ionisable.

Les caractéristiques des vingt acides aminés pouvant être présents dans une protéine sont représentées dans la table 2-1. Les figures 2-1,2-2 et la table 2-1 révèlent que les acides aminés et par conséquent les protéines sont composés des atomes de carbone, hydrogène, oxygène et azote. Certains acides aminés comme la cystéine et la méthionine possèdent en plus un atome de soufre. Les protéines et les acides aminés correspondent à un ensemble d'atomes.

Nom acide aminé Nom (code trois lettres, code une lettre)	Structure	Caractéristique de la chaîne latérale
Glycine (Gly, G)	$\begin{array}{c} \text{H} \\   \\ ^+\text{H}_3\text{N}-\text{C}-\text{COO}^- \\   \\ \text{H} \end{array}$	Hydrophobe Non chargée  Poids acide aminé : 75 Da
Alanine (Ala, A)	$\begin{array}{c} \text{CH}_3 \\   \\ ^+\text{H}_3\text{N}-\text{C}-\text{COO}^- \\   \\ \text{H} \end{array}$	Hydrophobe Non chargée Groupement méthyl  Poids acide aminé : 89 Da
Proline (Pro, P)	$\begin{array}{c} \text{H}_2 \\   \\ \text{H}_2\text{C}-\text{C}-\text{CH}_2 \\   \quad   \\ \text{N}^+-\text{C}-\text{COO}^- \\   \\ \text{H} \end{array}$	Hydrophobe Non chargée Fonction : amine secondaire  Poids acide aminé : 115 Da

Valine (Val, V)	$  \begin{array}{c}  \text{CH}_3 \\    \\  \text{H}-\text{C}-\text{CH}_3 \\    \\  ^+\text{H}_3\text{N}-\text{C}-\text{COO}^- \\    \\  \text{H}  \end{array}  $	Hydrophobe Non chargée Groupement isopropyle  Poids acide aminé : 117 Da
Leucine (Leu, L)	$  \begin{array}{c}  \text{CH}_3 \\    \\  \text{H}-\text{C}-\text{CH}_3 \\    \\  \text{CH}_2 \\    \\  ^+\text{H}_3\text{N}-\text{C}-\text{COO}^- \\    \\  \text{H}  \end{array}  $	Hydrophobe Non chargée Groupement isobutyle  Poids acide aminé : 131 Da
Isoleucine (Ile, I)	$  \begin{array}{c}  \text{CH}_3 \\    \\  \text{CH}_2 \\    \\  \text{H}-\text{C}-\text{CH}_3 \\    \\  ^+\text{H}_3\text{N}-\text{C}-\text{COO}^- \\    \\  \text{H}  \end{array}  $	Hydrophobe Non chargée Groupement butyle secondaire  Poids acide aminé : 131 Da
Sérine (Ser, S)	$  \begin{array}{c}  \text{OH} \\    \\  \text{CH}_2 \\    \\  ^+\text{H}_3\text{N}-\text{C}-\text{COO}^- \\    \\  \text{H}  \end{array}  $	Polaire Non chargée Fonction : alcool primaire  Poids acide aminé : 105 Da
Thréonine (Thr, T)	$  \begin{array}{c}  \text{OH} \\    \\  \text{H}-\text{C}-\text{CH}_3 \\    \\  ^+\text{H}_3\text{N}-\text{C}-\text{COO}^- \\    \\  \text{H}  \end{array}  $	Polaire Non chargée Fonction : alcool secondaire  Poids acide aminé : 119 Da
Cystéine (Cys, C)	$  \begin{array}{c}  \text{SH} \\    \\  \text{CH}_2 \\    \\  ^+\text{H}_3\text{N}-\text{C}-\text{COO}^- \\    \\  \text{H}  \end{array}  $	Polaire  Ionisation : fonction thiol (pKa : 8,3)  Poids acide aminé : 121 Da
Méthionine (Met, M)	$  \begin{array}{c}  \text{CH}_3 \\    \\  \text{S} \\    \\  \text{CH}_2 \\    \\  \text{CH}_2 \\    \\  ^+\text{H}_3\text{N}-\text{C}-\text{COO}^- \\    \\  \text{H}  \end{array}  $	Hydrophobe Non chargée Fonction : thioether  Poids acide aminé : 149 Da

Acide aspartique (Asp, D)		<p>Polaire</p> <p>Ionisation : fonction acide carboxylique (pKa : 4,1)</p> <p>Poids acide aminé : 133 Da</p>
Acide glutamique (Glu, E)		<p>Polaire</p> <p>Ionisation : fonction acide carboxylique (pKa : 4,1)</p> <p>Poids acide aminé : 147 Da</p>
Asparagine (Asn, N)		<p>Polaire</p> <p>Non chargée</p> <p>Fonction amide</p> <p>Poids acide aminé : 132 Da</p>
Glutamine (Gln, Q)		<p>Polaire</p> <p>Non chargée</p> <p>Fonction amide</p> <p>Poids acide aminé : 146 Da</p>
Lysine (Lys, K)		<p>Polaire</p> <p>Ionisation : fonction amine (pKa : 10,8)</p> <p>Poids acide aminé : 146 Da</p>
Arginine (Arg, R)		<p>Polaire</p> <p>Ionisation : groupement guanidyl (pKa : 12,5)</p> <p>Poids acide aminé : 174 Da</p>
Histidine (His, H)		<p>Polaire</p> <p>Ionisation : un azote du cycle imidazole (pKa : 6)</p> <p>Poids acide aminé : 155 Da</p>

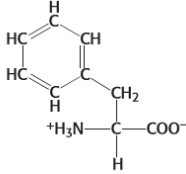
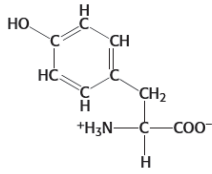
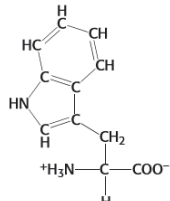
Phénylalanine (Phe, F)		Hydrophobe Non chargée Groupement phényle Acide aminé aromatique  Poids acide aminé : 165 Da
Tyrosine (Tyr, Y)		Polaire Groupement phénol  Ionisation : fonction hydroxyle (pKa : 10,9) Acide aminé aromatique  Poids acide aminé : 181 Da
Tryptophane (Trp, W)		Hydrophobe Non chargé Groupement indole Acide aminé aromatique  Poids acide aminé : 204 Da

Table 2-1. Liste des vingt acides aminés naturels présent dans une protéine.

Il est intéressant de noter que parmi les vingt acides aminés présents dans les protéines, neuf acides aminés sont dit essentiels car ils ne sont pas synthétisés par nos organismes. Il s'agit de la valine, leucine, isoleucine, thréonine, méthionine, lysine, histidine, phénylalanine et le tryptophane, ces acides aminés essentiels sont fournis par l'alimentation (exemple : poisson, œuf...).

Les acides aminés sont reliés les uns à la suite des autres grâce à des liaisons peptidiques et forment une chaîne d'acide aminé (chaîne polypeptidique) correspondant à la structure primaire (aussi appelée séquence protéique). La figure 2-3a, permet de visualiser la formation d'une liaison peptidique entre deux acides aminés suite à une réaction de condensation entre le groupement carboxylique d'un acide aminé (n) et le groupement amine d'un autre acide aminé ( $n_{+1}$ ). Dans la structure primaire, chaque acide aminé (n) possède deux voisins ( $n_{+1}$  et  $n_{-1}$ ), sauf le premier et le dernier acide aminé de la chaîne polypeptidique. En effet, le groupement amine du premier acide aminé est libre, ce résidu se nomme N-terminal ; de même que le groupement carboxylique du dernier acide aminé de la chaîne polypeptidique nommé C-terminal (Fig. 2-3b). L'ordre des acides aminés de la chaîne polypeptidique est encodé dans l'ADN. Pour information, les acides aminés sont aussi appelés résidus dans une chaîne polypeptidique.

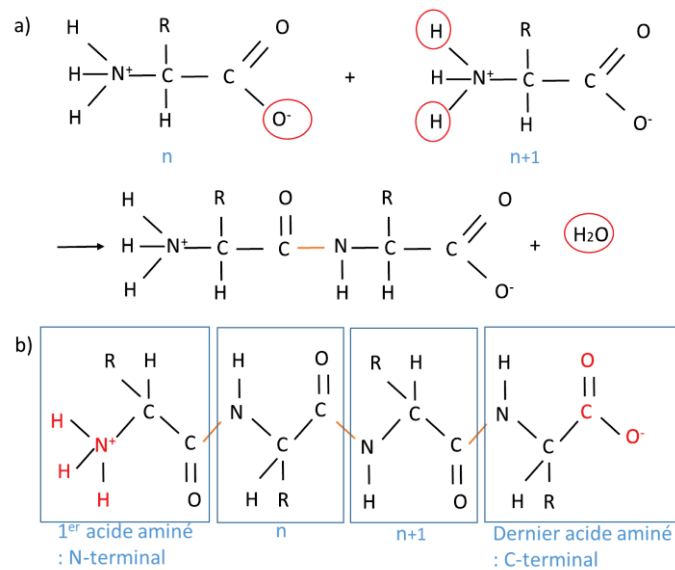


Figure 2-3. Formation d'une chaîne polypeptidique d'acide aminé. a) Représentation de la formation d'une liaison peptidique entre deux acides aminés. b) Exemple d'une chaîne polypeptidique de plusieurs acides aminés. R : correspond à la chaîne latérale de chaque acide aminé. Les traits orange représentent une liaison peptidique.

Les atomes O, N et H du squelette peptidique des acides aminés proches les uns des autres dans la structure primaire forment des liaisons hydrogènes entre eux menant à la formation de structures dites secondaires. Les structures secondaires correspondent au premier niveau d'organisation d'une chaîne polypeptidique. Il existe différents types de structures secondaires : les hélices, les feuilletts  $\beta$ , les coudes  $\beta$  et les boucles.

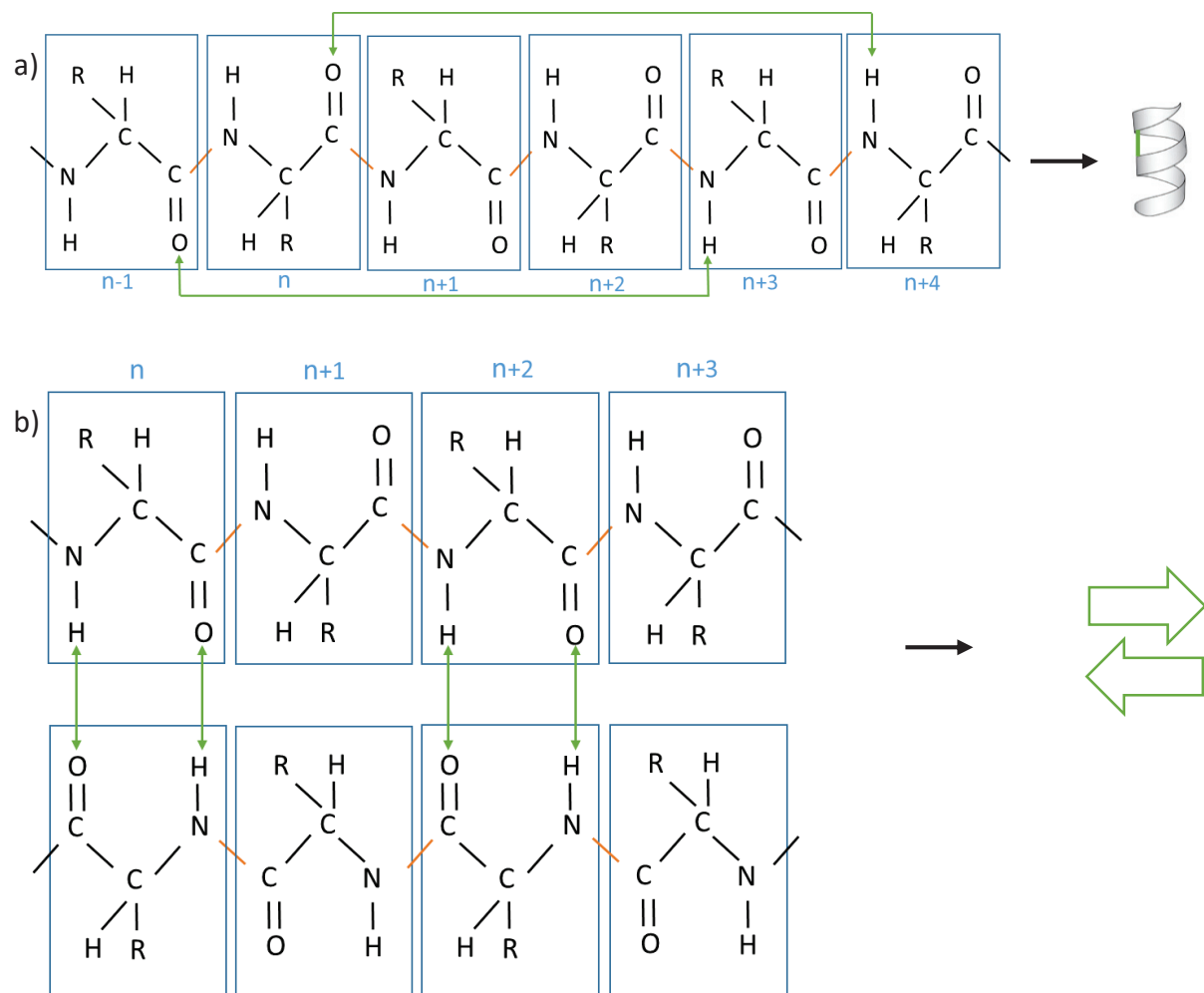
Lorsque les atomes forment des liaisons hydrogène entre des acides aminés de distance 3 à 5 dans la chaîne polypeptidique, une structure secondaire de type hélice est formée. Il existe différents types d'hélices (hélice  $\alpha$ , hélice  $3_{10}$ , hélice  $\pi$ ) suivant la distance entre les deux acides aminés formant la liaison hydrogène. L'hélice  $\pi$ , très rare dans les protéines, est formée suite à la liaison hydrogène entre le groupe amine (NH) d'un acide aminé et le groupement carbonyle (C=O) d'un acide aminé situé 5 résidus plus loin dans la séquence primaire. La formation d'une liaison hydrogène entre NH d'un acide aminé et C=O d'un acide aminé trois résidus plus loin entraîne la formation d'une hélice  $3_{10}$ . Enfin, la formation d'une liaison hydrogène entre le C=O d'un acide aminé et le NH d'un acide aminé quatre résidus plus loin entraîne la formation d'une hélice  $\alpha$  (Fig. 2-4a).

La structure secondaire feuillet  $\beta$  est constituée de plusieurs brins  $\beta$ . Les feuilletts  $\beta$  sont stabilisés par des liaisons hydrogène entre des brins  $\beta$  différents. Si les brins  $\beta$  adjacents sont dans la même direction, un feuillet  $\beta$  parallèle est formé, alors que si les brins beta sont de sens opposés, un feuillet  $\beta$  antiparallèle est créé. Un feuillet  $\beta$  est composé de plusieurs brins  $\beta$ , par conséquent, des feuilletts  $\beta$  mixtes peuvent exister. Dans un feuillet  $\beta$  antiparallèle, les groupements CO et NH de chaque acide aminé forment des liaisons hydrogènes avec les groupements CO et NH d'un acide aminé sur le brin adjacent. Chaque acide aminé d'un brin forme donc deux liaisons hydrogènes avec un acide aminé du brin adjacent (Fig. 2-4b). Dans les feuilletts  $\beta$  parallèles, le groupement NH est lié au groupement CO d'un acide aminé sur le brin adjacent. Toutefois, le groupe CO forme une liaison hydrogène avec le groupe NH d'un

autre acide aminé sur le brin adjacent. Ainsi un acide aminé d'un brin est relié avec deux acides aminés du brin adjacent (Fig. 2-4c). Les coudes  $\beta$  permettent de connecter deux brins  $\beta$  adjacents d'un feuillet antiparallèle. Les coudes  $\beta$  donnent un changement de direction. La liaison hydrogène entre le groupe NH d'un acide aminé et le groupe C=O d'un acide aminé situé trois résidus plus loin permet la formation du coude.

La distance entre deux acides aminés adjacents est approximativement de 3,5 Å pour un brin et de 1,5 Å dans une hélice, par conséquent les brins  $\beta$  sont des structures secondaires plus 'étendues', moins compactes que les hélices  $\alpha$ . Les hélices et brins  $\beta$  sont reliés entre eux via une autre structure secondaire : les boucles. Les boucles sont constituées de plus de quatre acides aminés et permettent de connecter les hélices entre elles, les hélices avec les brins et les différents brins  $\beta$  n'appartenant pas à un même feuillet entre eux. Contrairement aux autres structures secondaires citées précédemment, les boucles ne sont pas des structures régulières, spécifiques.

En conclusion, les atomes du squelette peptidique correspondant aux groupements amines et acides carboxyliques présents dans tous les acides aminés peuvent interagir entre eux et former des liaisons hydrogènes permettant la formation de structures secondaires régulières (hélices, feuilletts  $\beta$ , coudes  $\beta$ ) ou irrégulières (boucles).





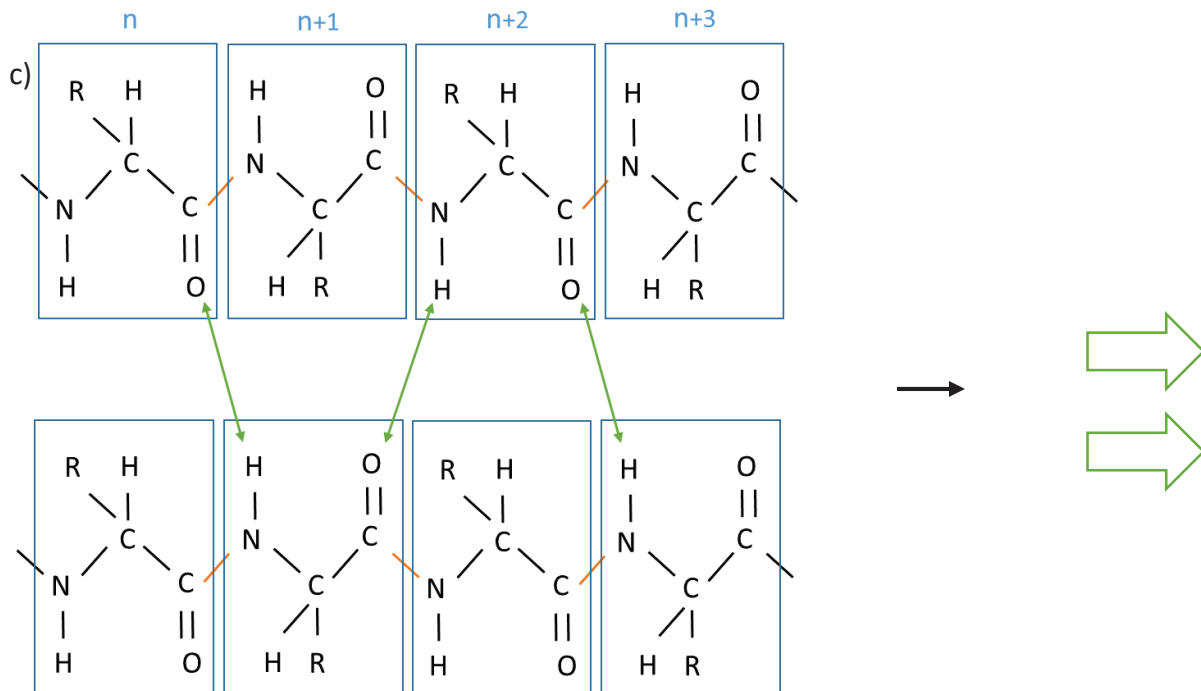


Figure 2-4. Représentation de la formation de structures secondaires. a) Formation d'une hélice  $\alpha$ . b) Formation d'un feuillet  $\beta$  antiparallèle. c) Formation d'un feuillet  $\beta$  parallèle.

Dans les programmes de visualisation des conformations protéiques, comme Swiss-PDB, les hélices sont représentées par des cylindres et les brins  $\beta$  par des flèches indiquant la direction des brins.

L'organisation et le repliement dans l'espace des différentes structures secondaires entre elles permettent la formation d'une structure tertiaire (Fig. 2-5). La structure tertiaire correspond à la conformation repliée de la protéine. Cette conformation est maintenue grâce à des interactions non covalentes comme les liaisons hydrogènes et ioniques, les interactions hydrophobes et Van der Waals entre les atomes des différents acides aminés. Toutefois des liaisons covalentes de types ponts disulfures peuvent se former entre deux cystéines suite à l'oxydation du groupement thiol de leurs chaînes latérales et participer au maintien de la structure tertiaire. La distance entre deux atomes de carbone reliés par une liaison covalente est de 1,54 Å. Les liaisons non-covalentes sont des liaisons de plus longues distances que les liaisons covalentes. La distance entre deux atomes impliqués dans une liaison ionique est d'environ 2,5 Å, d'environ 3 Å pour une liaison hydrogène et entre 3 Å et 6 Å pour une interaction de Van de Waals (2). Il est intéressant de noter que des acides aminés séparés par plusieurs acides aminés dans la structure primaire peuvent être proches (à moins de 5 Å) dans la structure tertiaire.



Figure 2-5. Représentation de la structure tertiaire de la chaîne D de la sous-unité B pentamérique de la choléra toxine via un programme de visualisation (swiss-pdb). En violet, représentation des hélices  $\alpha$ , en jaune des brins  $\beta$  et en bleu des boucles et des coudes  $\beta$ .

La conformation protéique possède une organisation structurale multi-échelle dans l'espace. La plus petite échelle est l'atome (1 à 3 Å soit 0,1 à 0,3 nm). Différents atomes s'associent pour former un acide aminé. Les acides aminés possèdent toujours une taille de l'ordre de l'angström, toutefois la taille des acides aminés varie suivant la longueur de la chaîne latérale. Par exemple, l'alanine possède une longueur de chaîne latérale de moins de 2 Å alors que la chaîne latérale de la lysine est de l'ordre de 6 Å. Les atomes du squelette peptidique des acides aminés proches spatialement interagissent entre eux pour donner des structures secondaires. Une hélice alpha composée de sept acides aminés à une longueur d'environ 11 Å, alors qu'un brin beta de sept acides aminés à une longueur de 21 Å. Enfin, ces structures secondaires interagissent entre elles (via leurs atomes) pour donner la conformation de la protéine dans l'espace. Une protéine globulaire possède une taille entre 3 nm et 10 nm (30 – 100 Å).

La protéine évolue au cours du temps pour former la conformation protéique (structure tertiaire). Par conséquent, en plus d'une organisation multi-échelle dans l'espace, la conformation protéique possède aussi une organisation multi-échelle dans le temps, se traduisant par des fluctuations moléculaires, on parlera donc de la dynamique moléculaire. Pour rappel, la protéine est composée d'acides aminés, de structures secondaires et d'une structure tertiaire. Les chaînes latérales de chaque acide aminé forment 'peu' d'interactions. Les structures secondaires (plusieurs acides aminés) forment un ensemble d'interactions créant un domaine structural avec une réponse collective sur une échelle de temps. La conformation protéique (structure tertiaire) possède un ensemble d'interactions plus important et sur une surface plus grande que les structures secondaires. Par conséquent, les chaînes latérales des acides aminés présentent des mouvements (réponse collective d'atomes qui interagissent entre eux via des liaisons hydrogènes et ioniques et des interactions hydrophobes, de Van Der Waals) plus rapides que les éléments de structures secondaires, eux-mêmes plus rapides que ceux de la structure tertiaire. D'après Muñoz et Cerminara en 2016, les chaînes latérales des acides aminés possèdent des mouvements de l'ordre de la nanoseconde alors que les éléments de structures secondaires présentent des mouvements de l'ordre de la microseconde à la milliseconde (3). Les mouvements de structure tertiaire

(réponse collective d'atomes entraînant le dépliement, repliement de domaine et/ ou de la chaîne polypeptidique) sont donc des mouvements encore plus lents que la milliseconde. Un domaine correspond à une partie d'une chaîne polypeptidique qui se replie de façon indépendante pour former une structure stable.

Si une protéine possède plusieurs chaînes polypeptidiques, l'assemblage de ces chaînes va former une structure quaternaire (Fig. 2-6) (1). Dans une structure quaternaire, les interactions entre les chaînes polypeptidiques correspondent à une interface. L'assemblage protéique possède aussi des mouvements de repliement, dépliement pour chaque chaîne polypeptidique et entre les chaînes (interfaces).

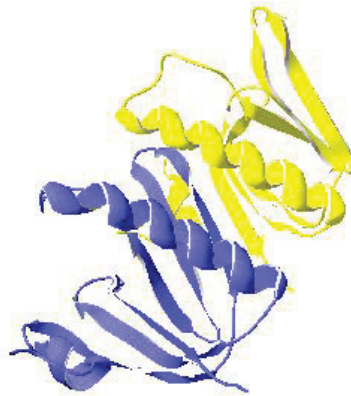


Figure 2-6. Représentation de la structure quaternaire d'une protéine grâce au programme de visualisation swiss-pdb. Une chaîne polypeptidique est représentée en bleu et une autre en jaune.

La figure 2-7 permet de résumer l'organisation multi-échelle d'une conformation protéique (structure tertiaire ou quaternaire) dans le temps.

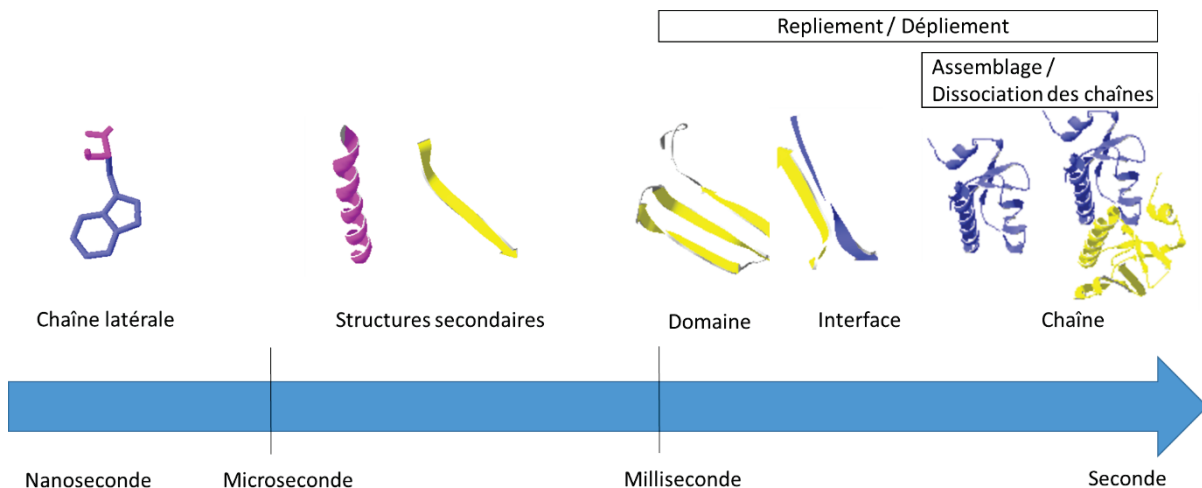


Figure 2-7. Représentation de l'organisation temporelle d'une conformation protéique.

En conclusion, une protéine est définie par sa séquence, sa conformation, sa dynamique moléculaire et sa fonction. La fonction d'une protéine repose donc sur sa conformation protéique et sa dynamique moléculaire.

Dans chaque acide aminé, la position du carbone alpha (carbone lié par une liaison covalente à la chaîne latérale), de l'azote du groupe amine et du carbone du groupe

carboxylique sont repérés par des angles de rotations phi et psi. Les contraintes de ces angles de rotations s'expliquent par la présence de plusieurs acides aminés dans une chaîne polypeptidique. Les atomes de la chaîne latérale ne possèdent pas les mêmes contraintes spécifiques, les chaînes latérales ont une extrémité libre. Ainsi, chaque atome d'un acide aminé et plus particulièrement ceux de la chaîne latérale possèdent un degré de liberté (possibilité de mouvement dans l'espace) plus ou moins important. Cette variation dans le degré de liberté donne la possibilité à chaque acide aminé et plus particulièrement à la chaîne latérale de se mettre dans différentes positions et donc de former différentes interactions non-covalentes avec les atomes des autres acides aminés, comme illustré dans la figure 2-8. En accord avec l'étude de Best et al en 2006 où un faible nombre de structures (généralement moins de 5) est nécessaire pour capturer l'hétérogénéité du squelette peptidique contre un plus grand nombre de structures (entre 20 et 40) pour représenter l'hétérogénéité de la chaîne latérale dans un ensemble de structures avec plus de 90% d'identité de séquence (4).

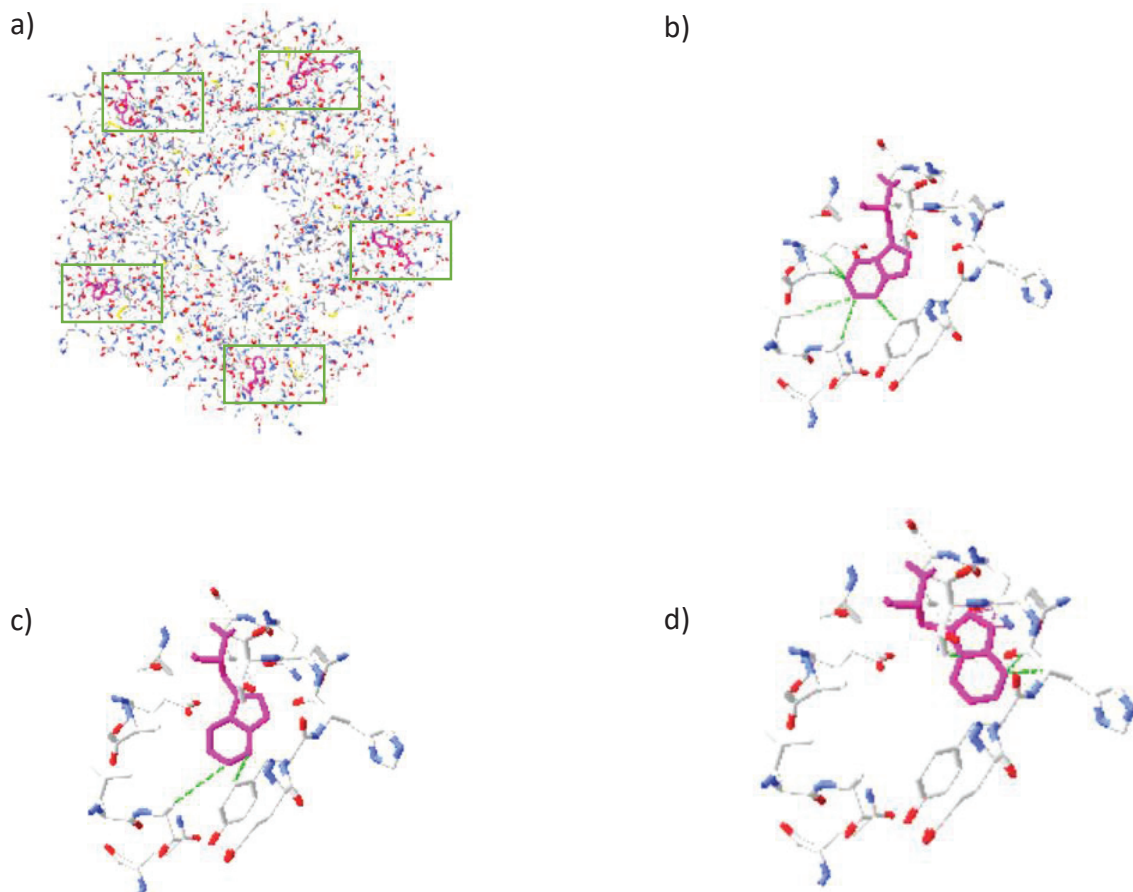
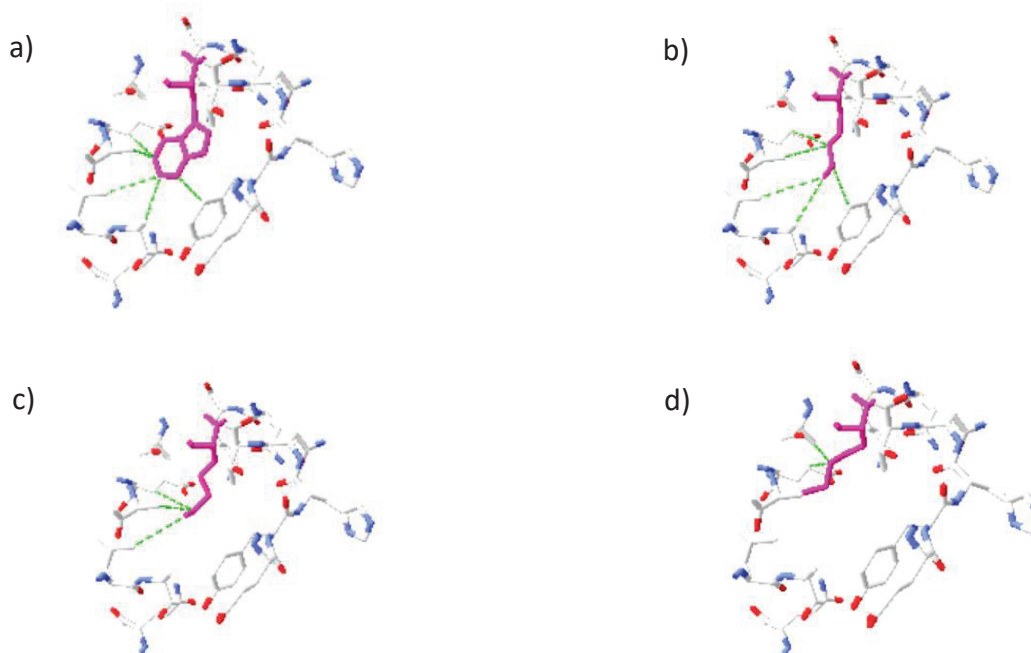


Figure 2-8. Représentation de différentes positions de la chaîne latérale du tryptophane. a) Visualisation de la sous-unité B pentamérique de la choléra toxine (code PDB : 1EEI). En blanc : atome de carbone, en bleu : atome d'azote et en rouge : atome d'oxygène. Le tryptophane 88 représenté en rose est encadré par un carré vert pour chaque monomère (chaîne) de la toxine. b) Zoom sur le tryptophane 88 colorié en rose. Les pointillés en vert représentent les interactions atomiques possibles entre les atomes du tryptophane et les atomes des acides aminés proches spatialement (seuil de 5 Å). c) Le changement de position de la chaîne latérale de Trp 88 entraîne la disparition d'interactions

*moléculaires. d) Le changement de position de la chaîne latérale de Trp 88 entraîne l'apparition de nouvelles interactions moléculaires.*

Comme le montre la figure 2-8, suivant la position de l'acide aminé dans l'espace, des interactions alternatives peuvent apparaître (Fig. 2-8d) et d'autres peuvent disparaître (Fig. 2-8c et Fig. 2-8d). Ces interactions alternatives pourraient changer la structure secondaire et/ou la structure tertiaire de la protéine et donner un ensemble de conformations (set de conformation) pour une même séquence et une même fonction protéique (5). Ainsi une protéine est représentée par un ensemble conformationnel plutôt que par une conformation unique (5–9).

Les mutations faux-sens dans la séquence d'ADN entraînent des changements d'acides aminés dans la séquence protéique. Ces mutations peuvent impacter la structure secondaire et/ou la conformation protéique (Fig. 2-9). C'est le cas par exemple de la protéine aérolysine, où la mutation de la thréonine 253 en cystéine (T253C) provoque la perte d'une liaison hydrogène augmentant ainsi la flexibilité d'une structure secondaire (boucle) sans impacter la conformation protéique. Toutefois, la double mutation T253C et A300C provoque la formation d'un pont disulfure entre une cystéine d'une boucle (T253C) et une cystéine d'un feuillet  $\beta$  (A300C) impactant ainsi la conformation protéique en empêchant l'oligomérisation de la protéine en heptamère (10). Ces changements pourraient potentiellement expliquer pourquoi les individus réagissent différemment à un même traitement. En effet, l'interaction entre la protéine cible et le médicament peut être associée à de petits ajustements locaux comme le mouvement de chaînes latérales mais aussi à des mouvements plus grands (11). Or si des interactions alternatives sont présentes au sein de l'ensemble conformationnel de la protéine ou à cause de mutation, les ajustements suite à l'interaction de la protéine et du médicament pourraient être perturbés, limités et donc impacter l'effet du médicament.



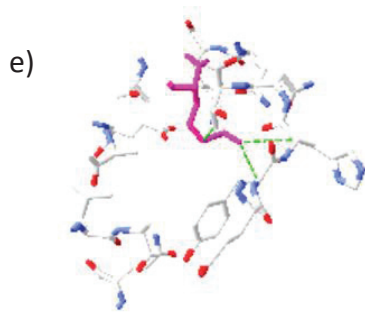


Figure 2-9. Représentation de l'impact d'une mutation sur les interactions atomiques. a) Représentation des interactions atomiques du tryptophane 88 (colorié en rose) avec les atomes des autres acides aminés proches spatialement (distance seuil de 5Å) en pointillés vert, de la sous-unité B pentamérique de la choléra toxine (code PDB : 1EEI). b) à e) Représentation des interactions atomiques suite à la mutation du tryptophane 88 en lysine 88, avec tous les cas de figures possibles : b) pas de changement d'interactions. c) Conservation de certaines interactions atomiques mais d'autres disparaissent. d) Conservation d'interactions atomiques et apparition d'une nouvelle interaction. e) Apparition de nouvelles interactions atomiques.

L'hétérogénéité conformationnelle est un paramètre essentiel pour expliquer la robustesse des protéines contre les mutations. En effet, une protéine peut subir une ou plusieurs mutations sans que celle-ci n'affecte la fonction de la protéine. Alors qu'une autre ou plusieurs autres mutations pourront entraîner une perte de fonction (pathologie). L'absence de perte de fonction pour chaque changement conformationnel, s'explique par l'hétérogénéité conformationnelle de la protéine native, soit l'ensemble des conformations protéiques possibles (set de conformation). Si la mutation entraîne un changement conformationnel mais que cette 'nouvelle' conformation fait partie du set conformationnel de la protéine native, alors la mutation n'aura aucun impact sur la fonction. Ce qui ne sera pas le cas, si la mutation entraîne une 'nouvelle' conformation qui n'appartient pas au set conformationnel de la protéine native. Ponzoni et Bahar en 2018, suggèrent qu'en plus de l'étude des séquences, et des conformations, l'étude de la dynamique des protéines est importante pour déterminer l'impact de mutation sur la fonction protéique (12).

De plus, il est démontré que l'ensemble conformationnel est essentiel pour comprendre les aspects biologiques des fonctions protéiques comme (9) : les processus catalytiques des enzymes (13–15), la reconnaissance protéine-protéine (16,17), la transduction du signal (18), la capacité des protéines à développer de nouvelles fonctions (19).

Cette première partie met en évidence, l'importance de caractériser une conformation protéique individuelle à l'échelle de l'atome et/ou des interactions atomiques et/ou de l'acide aminé et de cribler l'ensemble des conformations d'une protéine pour mesurer l'impact des mutations. Pour établir quand les mutations affectent la structure et la dynamique de la protéine pour mieux diagnostiquer les mutations et leurs relations aux maladies ou à la résistance des médicaments.

Au jour d'aujourd'hui il existe différentes techniques pour étudier la conformation protéique. Les techniques les plus couramment utilisées sont la cristallographie par diffraction de rayon X et la résonance magnétique nucléaire (RMN : nuclear magnetic resonance). Les



principes généraux de ces deux techniques seront discutés dans la deuxième partie de ce chapitre.

## 2.2 Détermination de la conformation protéique

La cristallographie rayon X et la résonance magnétique nucléaire (RMN) sont généralement utilisées pour déterminer expérimentalement une conformation protéique. Ces techniques permettent de connaître la position de chaque atome dans l'espace. Une fois, la conformation protéique déterminée, celle-ci est déposée dans la PDB (protein data bank). Dans la PDB, environ 80 % des conformations protéiques proviennent de l'étude cristallographique, et 20 % des études RMN. Le site web de la PDB est : <https://www.rcsb.org/>.

### 2.2.1 Cristallographie par diffraction de rayon X

L'utilisation de la technique de cristallographie pour déterminer la conformation protéique nécessite :

- D'obtenir des cristaux de protéines.
- D'utiliser les rayons X pour déterminer la position des atomes.

Le signal obtenu est donné par les électrons, la résolution de la cristallographie est généralement supérieure à 1 Å, par conséquent, tous les atomes d'une protéine (carbone, azote, oxygène, soufre) peuvent être visualisés sauf l'hydrogène.

La technique de cristallographie ne possède pas de limite de taille en acide aminé, toutefois il est nécessaire de cristalliser la protéine. Pour obtenir des cristaux, la concentration protéique initiale doit être supérieure au mg/ml. Cette quantité de protéine peut poser des problèmes, notamment pour l'étude des protéines membranaires. En effet, pour avoir une telle concentration, les protéines membranaires doivent être solubilisées, généralement via des détergents. Toutefois, l'utilisation de détergent risque de dénaturer les protéines membranaires, ce qui rendra impossible l'utilisation de la cristallographie.

Pour permettre la cristallisation des protéines, un agent cristallisant est ajouté. Via la diffusion de vapeur, la méthode de la goutte suspendue est utilisée pour former des cristaux. A une certaine concentration protéique et sous certaines conditions (température, pH) les protéines peuvent s'associer en cristal. Un cristal de protéines est donc un solide ordonné constitué d'un grand nombre de protéines se répétant périodiquement dans les trois dimensions (xyz) de l'espace. Cependant, les protéines fortement hydrophobes et flexibles cristallisent mal, celles-ci sont donc généralement caractérisées par domaines (partie d'une chaîne polypeptidique qui se replie de façon indépendante).

Pour pouvoir être exploitable aux rayons X, les cristaux obtenus doivent généralement avoir une taille supérieure ou égale à 50 micron. Les cristaux avec une taille inférieure à 10 micron sont inutilisables en rayon X. De plus, ce n'est pas parce qu'un cristal est obtenu que la structure protéique peut être déterminée.

Pour minimiser les dégâts de radiation des rayons X, les cristaux sont généralement congelés. Il est nécessaire de procéder à une congélation rapide (vitrification) en présence d'agent cryoprotecteur pour éviter la formation de cristaux de glace. Les cristaux de glace sont plus

gros que l'eau liquide, ainsi la formation de tels cristaux aura pour conséquence l'explosion des cristaux protéiques.

Dans les premières étapes de la cristallisation par diffraction de rayon X, il est nécessaire de cristalliser une protéine purifiée pour obtenir un cristal qui sera ensuite 'bombardé' par des rayons X. Le cristal et plus particulièrement les électrons des atomes de carbone, d'azote, d'oxygène et de soufre vont réémettre les rayons X. Suite à la réémission des rayons X par le cristal, une carte de densité électronique pourra être obtenue. A partir de ces informations les coordonnées xyz (en Å) et la nature des atomes pourront être déterminés avec précision (erreur possible de 0,3 à 0,1 Å). Grâce à ces informations la structure 3D de la protéine est résolue.

Les données de cristallographie par rayon X ne sont pas toutes réalisées sur des cristaux congelés, des expériences sur des cristaux à température ambiante sont aussi réalisées. Suivant la température utilisée pour l'étude des cristaux, les cartes de densités électroniques peuvent être différentes. En étudiant la protéine H-Ras en cristallographie, Fraser et al en 2011 ont trouvé un réseau allostérique compatible avec les fluctuations détectées en solution par RMN dans les cartes de densité d'électrons à température ambiante mais pas pour les températures cryogéniques. De plus, ils ont émis l'hypothèse que le refroidissement cryogénique remodèle les ensembles conformationnels des chaînes latérales exposées aux solvants et enfouies (20). Ainsi, le refroidissement cryogénique engendre de potentiels changements conformationnels par rapport à la protéine native, relevant ainsi une hétérogénéité conformationnelle des protéines. De même, la modification des conditions de cristallisation (pH, température) peut aussi modifier la stabilité relative des conformères et générer des différences entre les structures cristallographiques d'une même protéine (9). La cristallographie par rayon X pourrait donc permettre potentiellement de cribler différentes conformations protéiques.

### 2.2.2 Résonance magnétique nucléaire

En RMN, la protéine étudiée est en solution, de plus les protéines étudiées doivent avoir une taille inférieure à 30 kDa. En partant du principe qu'un acide aminé fait en moyenne 110 Da (110 g/mol), les protéines étudiées en RMN doivent avoir moins de 273 acides aminés. Les concentrations sont généralement de l'ordre du millimolaire dans les expériences RMN.

En RMN, tous les atomes avec une masse atomique impaire sont visibles. Si la masse atomique est paire, la charge nucléaire de l'atome doit-être regardée. Si celle-ci est impaire, l'atome sera visible, si la charge est paire l'atome sera invisible en RMN. La représentation symbolique du noyau d'un atome est :



avec X le symbole de l'atome, A la masse atomique soit le nombre de nucléons (nombre de protons + nombre de neutrons) et Z la charge nucléaire soit le numéro atomique (nombre de protons). Les atomes  ${}^{12}_6\text{C}$ ,  ${}^{16}_8\text{O}$ ... ne sont pas observables en RMN contrairement aux atomes



$^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ,  $^{14}\text{N}$  ... Certains atomes comme le carbone 13 ( $^{13}\text{C}$ ) observés en RMN sont peu présents naturellement.

Les atomes observables en RMN possèdent un spin. Le spin se comporte comme un moment magnétique. En absence de champ magnétique externe, les spins d'un échantillon sont distribués uniformément, par conséquent, le moment magnétique total de l'échantillon est nul. Pour déterminer la conformation protéique, l'échantillon est soumis à un champ magnétique externe. Les spins vont s'orienter par rapport à ce champ magnétique et le moment magnétique total de l'échantillon ne sera plus nul. Ensuite un second champ magnétique est appliqué pour basculer les spins hors de leurs positions d'équilibres et vers des états énergétiques plus élevés. La perturbation des spins est réalisée pendant quelques microsecondes à quelques millisecondes. Pour finir, le second champ magnétique est coupé et le retour à la normale des spins (processus de relaxation) est observé. C'est la différence entre les deux états énergétiques des spins qui est responsable du moment magnétique mesuré sur l'échantillon pendant l'expérience RMN. Le moment magnétique enregistré est soumis à une transformée de Fourier pour passer du domaine des temps (FID : free induction decay) au domaine des fréquences. Ce signal correspond à la somme des différents spins ayant subi le changement d'état énergétique. De plus, la fréquence de résonance de chaque spin dépend de son environnement électronique.

Pour avoir des informations sur la structure des échantillons protéiques étudiés, la position des pics est utilisée, de même que l'intégration (aire sous la courbe, pour déterminer le nombre de spins impliqués) et la multiplicité des pics (pour avoir des informations sur le voisinage).

En conclusion, la cristallographie par rayon X et la RMN permettent de déterminer avec précision les coordonnées atomiques et donc la conformation protéique à partir d'échantillons protéiques très concentrés. La table 2-2 regroupe les paramètres essentiels de ces deux techniques pour déterminer les conformations protéiques. A cause des fortes concentrations protéiques utilisées dans ces techniques, les protéines peuvent précipiter. De plus, en raison de ces concentrations protéiques, la plupart des expériences RMN mais aussi de cristallographie par rayons X correspondent à une observation moyenne de plusieurs états conformationnels (6). Par conséquent, il est difficile d'utiliser ces techniques pour réaliser un criblage conformationnel des protéines. Pour réaliser un criblage des différentes conformations protéiques, il est nécessaire d'avoir une technique qui détecte une seule protéine à la fois. La troisième partie expliquera les principes généraux de trois techniques à molécule unique.

	Cristallographie par diffraction de rayon X	RMN
Objectif	Détermination conformation protéique	Détermination conformation protéique

Echantillon protéique	Protéine (pas de limite de taille)	Taille protéique inférieure à 30 kDa, soit environ 273 acides aminés
Difficulté échantillon protéique	Etude de protéines membranaires  Protéines fortement hydrophobes et flexibles  Formation et taille des cristaux	Protéine avec une taille supérieur à 30 kDa
Préparation échantillon	Concentration supérieure au mg/ml  Doit trouver les 'bonnes' conditions de cristallisation (pH, température...) pour obtenir des cristaux	Protéine en solution  Concentration millimolaire  Marquage des atomes de carbones
Risques	Dégâts de radiation des rayons X  Durée de l'expérience, le temps pour obtenir des cristaux est très variable (jours à des mois)  Cristaux pas toujours utilisable (Cristaux < 10 µm = inutilisables)	Précipitation des protéines
Résolution	Résolution atomique, supérieur à 1 Å  Ne voit pas les atomes d'hydrogènes	Résolution atomique  Si masse atomique et charge atomique pair → atome invisible ( $^{12}_6\text{C}$ )
Informations obtenues	Coordonnées (xyz) et la nature des atomes	Fréquence de résonance de chaque spin dépendant de leur environnement
Plusieurs conformations dans un échantillon ?	Possible	Généralement
Détermine mutation ?	Oui	Oui

Table 2-2. Récapitulatif des techniques de cristallographie et de RMN.

## 2.3 Technique pour détecter une molécule unique

L'avantage des techniques à molécule unique comme son nom l'indique est l'étude de molécule unique et non d'une population de conformations. De plus, ces techniques permettent généralement d'observer des états transitoires et/ ou de courte durée qui sont habituellement masqués par l'ensemble conformationnel à cause de leurs courtes durée de vie et de leurs faibles nombres (21).

Au jour d'aujourd'hui, il existe plusieurs techniques pour la détection de macromolécule unique, comme :

- Pince Optique (Optical tweezer)
- Microscopie à force atomique (AFM : atomic force microscopy)
- SmFRET (transfert d'énergie par résonance de fluorescence d'une molécule unique : single-molecule fluorescence resonance energy transfert)

### 2.3.1 Pince optique

Un piège optique est formé en focalisant un faisceau laser avec un objectif de microscope à grande ouverture (22). Aux alentours du foyer du laser, un gradient de champ électrique est formé, par conséquent les particules diélectriques proches du faisceau laser subiront des forces qui les piégeront.

Il existe deux types de forces :

- La force de diffusion qui pousse la particule diélectrique dans la direction de propagation de la lumière.
- La force de gradient qui attire la particule diélectrique dans la direction du gradient. C'est l'interaction des dipôles de la particule avec le champ électrique qui crée la force de gradient.

La force de diffusion est dans le même sens que le faisceau laser, alors que la force de gradient est perpendiculaire au faisceau laser. Lorsque que la force de gradient dépasse la force de diffusion, la particule est piégée dans un état stable dans les trois dimensions de l'espace (21,23–25). La position du piège peut être modifiée en déplaçant le faisceau laser, permettant ainsi de sonder de multiples états conformationnels (26).

Cette technique permet de piéger de manière stable une particule dont la taille est comprise entre environ 20 nm et plusieurs micromètres. Cette technique peut aussi mesurer le déplacement tridimensionnel de la particule piégée avec une précision inférieure au nanomètre et une résolution temporelle inférieure à la milliseconde (23).

Les pinces optiques sont principalement utilisées pour l'étude des déplacements des moteurs moléculaires comme la kinésine (protéine enzymatique) le long des microtubules (27), de l'ARN polymérase le long de l'ADN (28), les propriétés mécaniques des protéines (29).

Lors de l'utilisation de cette technique, il est nécessaire que les échantillons soient hautement purifiés, car toutes les particules diélectriques proches du foyer du laser peuvent être piégées. De plus, les dommages optiques induits dans les échantillons piégés ne sont pas totalement compris (23).

En conclusion, cette technique permet de piéger de manière stable et unique différentes particules. Toutefois, pour piéger des particules plus petites que 20 nm, la technique des pinces optiques à double nanotrou a dû être développée (30,31).

### 2.3.2 Microscopie à force atomique

La microscopie à force atomique (AFM) est une version du microscope à sonde de balayage (scanning probe microscope) où la topologie d'une surface est examinée point par point. L'AFM est composée d'un petit levier qui possède une pointe à son extrémité. Après le contact initial entre la pointe du levier et la surface où l'échantillon est déposé, le levier est pressé sur la surface avec une force constante. La pointe se déplace sur la surface du support et permet de cartographier les caractéristiques de la surface à une résolution inférieure au nanomètre (23). Plus précisément, la résolution en X et Y de l'AFM limitée par la taille de la pointe est généralement de l'ordre de quelques nanomètres, alors que la résolution en Z est de l'ordre de l'Angström (32). Cette technique détecte toute variation de hauteur sur la surface.

De plus, l'AFM est capable de déterminer les interactions biomoléculaires : interaction anticorps-antigène (33,34), ligand-récepteur impliquant une protéine (35,36) ... Dans ce cas, l'AFM mesure les interactions attractives et/ou répulsives entre la molécule fixée sur la pointe du levier et les molécules présentes sur la surface. L'AFM est aussi utilisée pour étudier le dépliement, repliement des protéines (37–39). Quand la pointe est à proximité de la surface, les forces d'interactions entre la pointe et l'échantillon entraîneront une déviation du levier. Cette déviation est mesurée à partir du déplacement du faisceau de laser réfléchi à l'arrière du levier et projeté sur un détecteur.

Les avantages de la technique AFM sont : la préparation simple et rapide des échantillons, la possibilité de modifier la surface du support avec n'importe quelle molécule, l'étude de molécules individuelles. Toutefois, le principal inconvénient de la technique AFM est la spécificité des interactions. En effet, il est parfois difficile de distinguer les interactions entre la pointe de l'AFM et la molécule d'intérêt des interactions non spécifiques, ou des interactions entre la pointe de l'AFM et des contaminants adsorbés sur la surface. De plus, les dimensions de la pointe du levier sont plus grandes que celles des molécules, par conséquent plusieurs molécules peuvent se fixer sur le levier (23).

En conclusion, l'AFM est principalement un outil d'imagerie, toutefois des versions modifiées existent comme l'AFM à haute vitesse (40) qui permet de suivre des événements moléculaires en temps réel.

### 2.3.3 Single-molecule fluorescence Resonance Energy : SmFRET

Le FRET (fluorescence resonance energy transfert) est un processus de transfert d'énergie entre une molécule à l'état excité et une molécule à l'état fondamental. Le fluorophore qui transfère son énergie est appelé donneur tandis que la molécule qui accepte l'énergie est appelée accepteur. Le spectre d'émission du donneur et le spectre d'excitation de l'accepteur doivent présenter un chevauchement important pour que le transfert d'énergie

puisse avoir lieu. Ainsi, lorsque le donneur est à l'état excité et à proximité suffisante de l'accepteur, l'énergie d'excitation du donneur peut être transférée à l'accepteur, ce qui entraîne l'émission de l'accepteur. Le transfert d'énergie correspond à une interaction dipôle-dipôle entre l'état excité du donneur et l'état fondamental de l'accepteur. (41,42). On parle de fluorophore quand une molécule est capable d'émettre de la fluorescence après excitation.

Le SmFRET (Single-molecule fluorescence Resonance Energy) est une application du FRET dans laquelle une paire de fluorophore (donneur et accepteur) est sur une seule molécule plutôt que sur un ensemble de molécules comme pour le FRET (21). Les mesures SmFRET sont réalisées soit avec un microscope confocal, soit avec un microscope de fluorescence à réflexion interne totale (TIRF : total-internal reflection). Les molécules étudiées peuvent être immobilisées sur une surface ou diffusées librement. Un inconvénient du SmFRET est l'effet de l'irradiation du laser sur l'échantillon (41).

L'efficacité du FRET et du SmFRET dépend de la distance entre le donneur et l'accepteur, généralement la distance doit être de l'ordre de 10 nm ou moins (42,43). Plus précisément, si la distance entre l'accepteur et le donneur est faible (quelque nm), une faible intensité du spectre d'émission du donneur et une forte intensité du spectre d'excitation de l'accepteur est observée, traduisant une efficacité élevée, soit un transfert d'énergie supérieur à 50 %. Inversement, si la distance entre l'accepteur et le donneur est grande, alors l'intensité du spectre d'émission du donneur est très élevée mais l'intensité du spectre d'excitation de l'accepteur est faible traduisant une faible efficacité du FRET, soit un transfert d'énergie inférieur à 50 %. Ainsi, plus la distance entre l'accepteur et le donneur est faible, plus le pourcentage de transfert d'énergie entre les deux est important et plus le FRET, SmFRET est efficace, et inversement quand la distance entre les deux augmente (41).

Une relation quantitative (Eq. 2-2) relie l'efficacité du transfert d'énergie ( $E_{\text{FRET}}$ ) et la distance  $r$  entre le donneur et l'accepteur :

$$E_{\text{FRET}} = \frac{1}{1 + \left(\frac{r}{R_0}\right)^6} \quad (\text{Eq. 2-2})$$

$R_0$ , est la distance entre le donneur et l'accepteur en Å où le transfert d'énergie est de 50 %. Si  $r = R_0$ , alors les intensités du spectre d'émission du donneur et du spectre d'excitation de l'accepteur sont équivalentes (21,41).

$R_0$  est définie comme la constante de Förster et s'exprime par l'équation 2-3 :

$$R_0 = 0,211 [\kappa^2 n^{-4} Q_D J(\lambda)]^{1/6} \quad (\text{Eq. 2-3})$$

$n$  est l'indice de réfraction du milieu,  $Q_D$  est le rendement quantique du donneur en absence d'accepteur,  $\kappa^2$  est le facteur d'orientation et  $J(\lambda)$  correspond au chevauchement spectral entre l'émission du donneur et l'absorption de l'accepteur.  $J(\lambda)$  est lié à l'intensité de fluorescence du donneur en absence d'accepteur et au coefficient d'extinction de l'accepteur (41).

L'efficacité du FRET reflète les changements de distances entre l'accepteur et le donneur (41) permettant ainsi d'étudier le repliement des protéines (44) et les changements de conformation protéique (45,46). Les échantillons utilisés en SmFRET peuvent être constitués de plusieurs molécules mais une seule molécule est sondée à la fois.

L'avantage des techniques précédemment citées est qu'elles permettent de regarder des molécules uniques et non des moyennes de populations. De plus, le déplacement du faisceau laser dans la technique des pinces optiques permettrait de cribler différents états conformationnels. Toutefois cette technique et la technique de l'AFM ne possèdent pas une résolution à l'échelle de l'acide aminé ou de l'atome. Par conséquent, il sera difficile de comprendre les différences entre les différentes conformations et l'impact d'une ou des mutations sur la conformation protéique. Inversement, le SmFRET est une technique à molécule unique avec une résolution à l'échelle de l'acide aminé car elle détecte le transfert d'énergie entre un donneur et un accepteur. Bien que cette technique pourrait être utilisable pour cribler différentes conformations, une limite de détection peut être présente. Dans une protéine, si celle-ci n'est pas marquée, trois acides aminés sur vingt peuvent être des fluorophores : le tryptophane, la tyrosine et la phénylalanine. Ainsi, uniquement l'environnement autour de ces acides aminés pourra être étudié si ces acides aminés sont suffisamment proches spatialement pour pouvoir permettre le transfert d'énergie. Par conséquent, pour cribler différentes conformations protéiques il est nécessaire d'avoir une technique qui possède une échelle équivalente à la taille des protéines mais aussi une résolution à l'échelle de l'acide aminé ou des atomes sur l'ensemble de la protéine et non uniquement sur certaines parties.

## 2.4 Nanoconfinement et la spectroscopie diélectrique à large bande

Le nanoconfinement est utilisé pour réduire l'hétérogénéité conformationnelle d'un échantillon protéique et la spectroscopie diélectrique à large bande mesure la dynamique moléculaire qui dépend de la conformation protéique.

### 2.4.1 Nanoconfinement

Les membranes AAO (anodic aluminium oxyde) sont des membranes avec des nanopores hautement ordonnés, cylindriques, identiques obtenus par anodisation de l'aluminium pur. Ces nanopores peuvent contenir de faible quantité de matière. La quantité de matière dans les nanopores dépend de la concentration de l'échantillon et du volume du nanopore (47). L'utilisation de nanopores très courts permet de mesurer des attogrammes (un attogramme =  $10^{-18}$  g) et des zeptogrammes (un zeptogramme =  $10^{-21}$  g) de matière et d'étudier le comportement diélectrique de la matière à une échelle de longueur comparable aux dimensions moléculaires (47,48).

Le volume caractéristique d'un attogramme est de  $10 \text{ nm} * 10 \text{ nm} * 10 \text{ nm}$ , se rapprochant ainsi des dimensions caractéristiques d'une protéine. Par conséquent, si les protéines sont incorporées dans des nanopores, les degrés de liberté des atomes des acides aminés associés à différentes conformations protéiques devraient être 'globalement' réduits.

Le nanoconfinement devrait donc réduire l'ensemble conformationnel de la protéine et favoriser l'enrichissement et/ou la sélection d'une conformation protéique spécifique. Dans le cas de l'enrichissement conformationnel, si la conformation protéique ne peut pas rentrer dans les nanopores, les protéines modifient leurs conformations pour s'adapter à la forme du confinement (Fig. 2-10a). Dans ce cas, le confinement enrichit une population conformationnelle moyenne. Attention les conformations enrichies peuvent être des conformations protéiques 'forcées' qui n'existent pas dans l'ensemble conformationnel de la protéine. Pour la sélection conformationnelle, les conformations protéiques ne pouvant pas rentrer dans les nanopores sont exclues (Fig. 2-10b). La sélection conformationnelle est issue de l'enrichissement d'une population conformationnelle au détriment des autres conformations de l'ensemble conformationnel. Si les dimensions des nanopores changent en restant dans le même ordre de grandeur que les dimensions caractéristiques d'une protéine, le confinement changera et différentes conformations protéiques devraient ainsi pouvoir être enrichies et/ou sélectionnées.

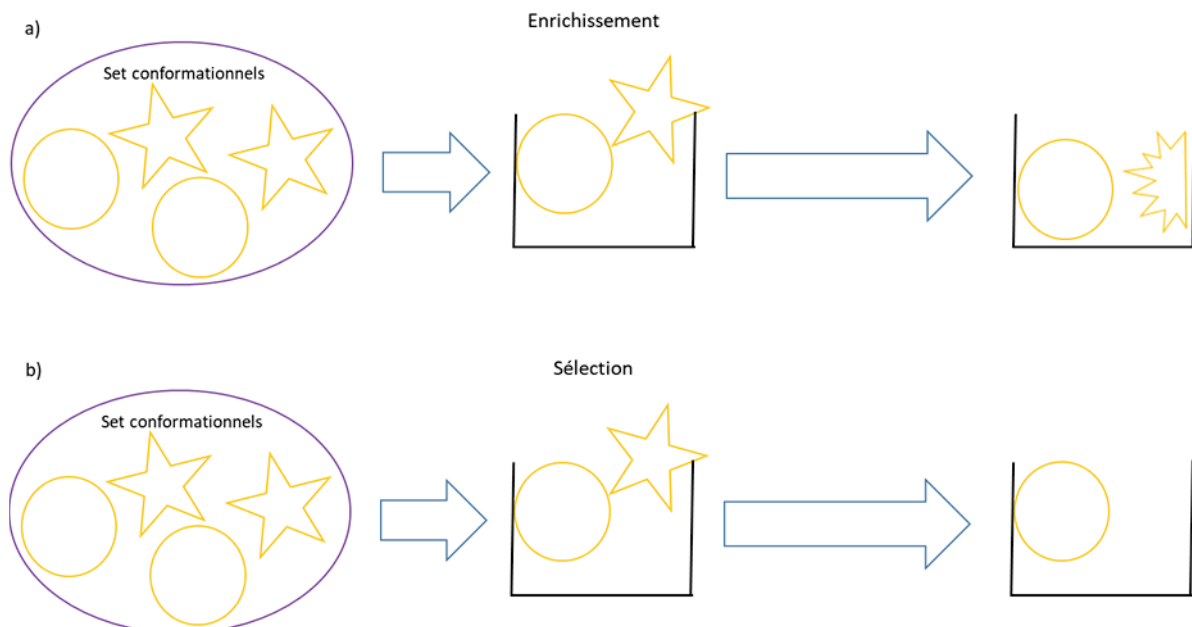


Figure 2-10. Enrichissement ou sélection des conformations protéiques via le confinement. a) Représentation de l'enrichissement conformationnel. b) Représentation de la sélection conformationnelle.

Comme les nanopores sont tous identiques, la conformation protéique enrichie et/ou sélectionnée dans un nanopores devrait être la même dans tous les nanopores de la membrane AAO. Le nombre de nanopores présent sur une membrane AAO dépend des dimensions de la membrane et des dimensions des nanopores. Généralement des centaines de milliers de nanopores sont présents dans une membrane, par conséquent, le signal de la conformation protéique est amplifié et un ensemble statistique est observé.

Les membranes AAO sont appliquées aux protéines pour :

- ✓ La reconstitution de protéine membranaire dans des bicouches lipidiques

Soubias et Gawrish en 2013 utilise des filtres AAO avec des nanopores de 200 nm de diamètre pour reconstituer des bicouches lipidiques. Pour cela, des protéoliposomes unilamellaires constitués de lipides et de protéine membranaire (la rhodopsine) sont passés à travers des



filtres AAO. La première bicouche formée adhéra fortement à la surface du pore, tandis que les autres bicouches seront plus 'lâches'. Dans leur étude, la rhodopsine conserve sa fonction dans les nanopores (49).

✓ La formation de couche protéique

Lazzara et al en 2012 (50) ont démontré que les nanopores cylindriques imposent une contrainte sphérique limitant la quantité de matière pouvant être incorporée. Dans leur étude, des couches de protéines globulaires sont formées. Dans un premier temps, l'avidine est absorbée par voie électrostatique sur la surface AAO puis reconnue par la sérualbumine bovine biotinylée qui sera à son tour reconnue par l'avidine et ainsi de suite. Le nombre de couches de protéines globulaires pouvant être formées augmente avec l'augmentation du diamètre des nanopores. En effet, avant d'arriver à la saturation du nanopore, une seule couche de protéine peut être déposée pour un diamètre de nanopore de 41 nm, contre 3 couches de protéines si le diamètre est de 65 nm et 5 couches si le diamètre est de 80 nm. De plus, la cinétique de dépôt des couches de protéines est significativement plus lente pour des petits diamètres de pores (25 à 35 nm), entraînant même une absence de dépôt de protéines si le diamètre des nanopores se situe entre 25-30 nm. Cette étude met en avant de potentiels problèmes de nanofluidique pour des petits diamètres de pores.

✓ La détection de protéine en créant des biocapteurs sélectifs

Pour la création d'un biocapteur, des anticorps sont généralement immobilisés sur la surface interne des pores.

De la Escosura-Muñiz et al en 2013 (51) utilisent les membranes AAO comme biocapteur de thrombine. Des aptamères de thrombine (séquence d'acides nucléiques synthétiques) sont immobilisés sur les surfaces internes des pores de 200 nm de diamètre et 60  $\mu\text{m}$  de long grâce à des liaisons peptiques entre les aptamères et des groupes carboxyliques générés sur la paroi interne des membranes AAO. La thrombine présente dans un échantillon sanguin se fixera sur les aptamères de thrombine et sera reconnue par des anticorps anti-thrombine marqués avec du AuNPs. Les membranes AAO sont fixées sur des SPCEs (screen-printed carbon electrodes). Cette cellule est ensuite immergée dans 100  $\mu\text{L}$  de  $\text{K}_3[\text{Fe}(\text{CN})_6]$ . Suite à une impulsion voltamétrique différentielle,  $[\text{Fe}(\text{CN})_6]^{4-}$  est oxydé en  $[\text{Fe}(\text{CN})_6]^{3-}$ . Or si la thrombine est fixée sur les aptamères, et que celle-ci est reconnue par l'anticorps anti-thrombine, la diffusion du  $[\text{Fe}(\text{CN})_6]^{4-}$  est ralentie, bloquée principalement à cause de la taille d'AuNPs (80 nm), entraînant une diminution du signal d'oxydation. En sachant que l'apatamère de thrombine agit comme un inhibiteur de l'activité enzymatique de la thrombine et que l'interaction aptamère thrombine-thrombine est très spécifique, ce test semble permettre de vérifier la fonctionnalité de l'enzyme.

En 2018, De la Escosura-Muñiz et al (52), ont aussi utilisé des membranes AAO pour détecter l'hormone parathyroïdienne (PHTLH : parathyroid hormone-like hormone). Comme précédemment, ils utilisent un signal d'oxydation pour quantifier la PHTLH mais cette fois avec des nanoparticules Bleu de Prusse qui sont oxydés en vert de Berlin.

Les études avec des membranes AAO contenant des nanopores ont mis en évidence que suivant la technique de détection :



- Les biocapteurs peuvent détecter des concentrations de l'ordre du ng/ml (51,52).
- La protéine est toujours fonctionnelle malgré le confinement (49).
- Des problèmes de nanofluidique sont présents suivant le diamètre des nanopores (50).

Ces études montrent qu'il est possible de descendre dans de faibles quantités de protéines tout en ayant un signal et en conservant la fonction protéique en utilisant des membranes AAO. Indiquant qu'il est possible de diminuer énormément la quantité de protéine déposée dans les nanopores et donc d'augmenter les chances de mesurer le signal diélectrique d'une ou de quelques protéines par pores sans impacter la fonction protéique, à condition que le signal diélectrique soit détectable.

#### 2.4.2 Spectroscopie diélectrique à large bande (BDS)

La technique de la Spectroscopie Diélectrique à Large Bande (Broadband Dielectric Spectroscopy : BDS) est utilisée pour caractériser la réponse diélectrique d'un matériau par la mesure de différents paramètres comme la perte diélectrique, la permittivité, la conductivité. Par la polarisation d'orientation, cette technique nous donne un accès direct à la dynamique moléculaire associée aux fluctuations de dipôles permanents. Comme les protéines possèdent des dipôles permanents, cette technique a été utilisée pour étudier la dynamique moléculaire des protéines. La méthode de spectroscopie diélectrique à large bande sera expliquée plus en détail dans le Chapitre 3 intitulé « Protéines, Méthodes, Méthodologie ».

Des protéines globulaires à l'état macroscopique ont déjà été étudiées par spectroscopie diélectrique à large bande (55–60). Ces études démontrent que la dynamique moléculaire des protéines globulaires à l'état macroscopique est globalement assez complexe. En effet, la dynamique moléculaire possède généralement plusieurs pics de relaxation et donc plusieurs processus de relaxation. Trois processus de relaxation moléculaire sont généralement détectés dans la gamme de fréquence  $10^{-2}$  Hz à  $10^7$  Hz (100 s à 10  $\mu$ s). Un premier processus de relaxation rapide (haute fréquence:  $10^4$  Hz et plus), attribué à la relaxation collective des molécules d'eau (bulk), un deuxième processus de relaxation intermédiaire, attribué aux interactions entre les molécules d'eau liée et les acides aminés de la protéine en surface, et un dernier processus de relaxation plus lent (basse fréquence :  $10^2$  Hz ou moins), attribué quant à lui aux mouvements internes de la protéine potentiellement couplés avec l'eau liée (55,56,59,60). Compte tenu du couplage étroit des protéines et de l'eau liée, la séparation des contributions de l'eau et des protéines est souvent discutable (55,56). Même si tous les dipôles de la protéine contribuent au signal diélectrique, les auteurs suggèrent que les groupements polaires et chargés de la protéine participent fortement au signal (59,60).

Il existe principalement deux moyens pour simplifier la complexité des spectres diélectriques. Le premier démontré par Khodadadi S et al en 2008 (56), est d'agir sur les niveaux d'hydratations des échantillons protéiques. S'ils sont faibles ( $0.05 < h < 0.1$  avec  $h$  : g de molécule d'eau / g de la protéine), un seul processus de relaxation est présent. Ce processus de relaxation est attribué au couplage protéine-eau en surface, soit le processus intermédiaire cité précédemment. Le deuxième montré par Schirò G et al en 2009 (59) confine

les solutions de protéines dans un gel de silice et conserve un niveau d'hydratation élevée ( $h : 0,3$ ). A l'état macroscopique, soit non confiné, trois processus de relaxation sont détectés alors qu'un seul processus de relaxation est détecté à l'état confiné, processus attribué aux mouvements internes des protéines. Par conséquent, si les protéines sont confinées, aucun processus de relaxation du solvant et des protéines couplées au solvant n'est détecté. Le confinement géométrique semble jouer un rôle crucial sur la détection de processus de relaxation associé 'uniquement' à la conformation protéique, suggérant qu'en changeant le confinement géométrique des nanopores, différents processus de relaxation moléculaire pourraient être détectés. Ce qui tend à confirmer notre hypothèse de départ, où la variation de géométrie des nanopores pourraient permettre d'enrichir et/ ou sélectionner différentes conformations protéiques permettant ensuite de réaliser un criblage de ces conformations.

Ces études suggèrent que pour simplifier la compréhension des spectres diélectriques, nous devons utiliser le confinement mais aussi avoir un faible niveau d'hydratation. Pour maximiser l'évaporation et donc avoir un faible niveau d'hydratation, nos échantillons protéiques resteront trois heures à 60 °C, puis une mesure diélectrique sera réalisée à la fin de ces trois heures. Pour vérifier que la déshydratation de l'échantillon protéique est maximale, un deuxième traitement thermique de trois heures sera ensuite réalisé à la même température que précédemment suivi d'une mesure diélectrique. Si les données de ces deux mesures sont identiques alors le niveau d'hydratation est le même. Le confinement sera utilisé pour enrichir et/ou sélectionner différentes conformations protéiques.

#### 2.4.3 Combinaison nanoconfinement et BDS

Le nanoconfinement possédant des dimensions comparables aux dimensions géométriques des protéines devrait permettre de réduire la quantité de protéine utilisée et d'étudier quelques protéines. De plus, la modification des dimensions du confinement devrait permettre d'enrichir et/ou sélectionner différentes conformations protéiques caractérisées par différentes interactions atomiques. Le BDS caractérise la dynamique moléculaire qui est étroitement liée à la conformation.

Pour ces raisons, la combinaison du nanoconfinement et de la spectroscopie diélectrique à large bande offre la possibilité de mesurer la dynamique moléculaire des protéines à l'échelle des fluctuations moléculaires et de cribler différentes conformations protéiques en changeant les dimensions du confinement. De plus, comme démontré dans la figure 2-9, la mutation d'un acide aminé peut entraîner un changement des fluctuations moléculaires qui sera facilement détectée par le BDS.

L'étude de la dynamique moléculaire est intéressante car les propriétés dynamiques associées à la mobilité du squelette peptidique et des chaînes latérales jouent un rôle essentiel dans de nombreux aspects du comportement protéique (stabilité, repliement, fonction) (61–64).

La combinaison de ces deux techniques a déjà été utilisée sur des polymères inertes (47,48). Toutefois à notre connaissance, l'étude de la dynamique moléculaire des protéines par nanoconfinement et BDS n'a jamais été réalisée auparavant. Par conséquent, le premier chapitre de résultat de ce mémoire (chapitre 4) se consacrera à l'étude de la membrane AAO,

du solvant et d'une protéine (la sous-unité B pentamérique de la choléra toxine : CtxB<sub>5</sub>) dans un confinement donné pour vérifier qu'une dynamique moléculaire associée à une conformation protéique soit bien détectée.

# Chapitre 3

## Protéines, Méthodes, Méthodologie

Les échantillons protéiques seront étudiés dans des conditions nanoconfinés et faiblement hydratés.

### 3.1 Protéines étudiées

#### 3.1.1 Sous-unité B pentamérique de la toxine du choléra

La principale protéine étudiée durant la thèse est la sous-unité B pentamérique de la toxine du choléra (CtxB<sub>5</sub>). La toxine du choléra (CT) est un facteur de virulence du choléra produit par *Vibrio cholerae* possédant une masse moléculaire d'environ 84 kDa (65).

CT possède deux sous-unités :

- Une sous-unité A (CtxA) qui possède une activité ADP-ribose transférase, il s'agit de la sous-unité avec l'activité toxique.
- Une sous-unité B (CtxB) qui est pentamérique (CtxB<sub>5</sub>) et qui permet l'internalisation de CtxA dans les cellules intestinales en se liant au récepteur ganglioside GM<sub>1</sub> (monosialotetrahexosylganglioside) présent sur la surface apicale des cellules épithéliales de l'intestin (66,67).

Plus précisément, la sous-unité A est constituée de deux fragments A<sub>1</sub> et A<sub>2</sub> reliés par un pont disulfure. CtxA est reliée à CtxB<sub>5</sub> via A<sub>2</sub> (67,68). L'interaction entre la sous-unité B pentamérique et GM<sub>1</sub> permet l'internalisation de la toxine cholérique dans les cellules intestinales. Celle-ci est ensuite transloquée dans le réticulum endoplasmique, où la réduction de la liaison disulfure entre A<sub>1</sub> et A<sub>2</sub>, permettra à A<sub>1</sub> de stimuler l'adénylate cyclase entraînant une augmentation de la production d'AMP cyclique (AMPC) (69). Les protéines kinases AMPC seront activées de façon permanentes entraînant un défaut dans la balance ionique de la cellule caractérisé par une accumulation excessive de sel (sodium) et d'eau dans la lumière intestinale (65,68). Les symptômes se manifestent par des crampes sévères et de très fortes diarrhées (65). La conformation pentamérique de CtxB (CtxB<sub>5</sub>) est nécessaire pour permettre l'interaction avec le récepteur GM<sub>1</sub> (65).

CtxB<sub>5</sub> est constituée de cinq monomères de 11 700 Dalton (65,70) comme représenté dans la figure 3-1. Chaque monomère possède 103 acides aminés. Ces acides aminés s'organisent en deux hélices alpha et six brins beta. Dans les conditions physiologiques, CtxB<sub>5</sub> est sous forme de pentamère natif, fonctionnel (71). Chaque monomère de CtxB<sub>5</sub> (M) est relié aux monomères adjacents (M<sub>-1</sub> et M<sub>+1</sub>) par différents types d'interactions non covalentes (liaisons hydrogènes, interactions hydrophobes ...). Par conséquent, chaque monomère de CtxB<sub>5</sub> possède deux domaines capables d'interagir avec ses monomères adjacents (72). CtxB<sub>5</sub>

possède trois interfaces entre deux chaînes (monomères) : une interface dans le domaine N-terminal (interface  $I_1$ ), une interface entre brins beta (interface  $I_2$ ) et une interface entre les hélices alpha (interface  $I_3$ ).

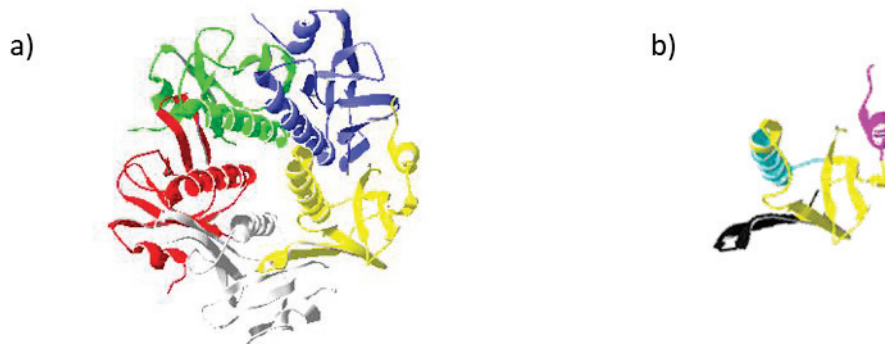


Figure 3-1. Sous-unité B pentamérique de la choléra toxine (code PDB : 1EEI). a) Représentation de CtxB<sub>5</sub>. Les couleurs jaune, blanche, rouge, verte et bleu représente les cinq monomères de la toxine. b) Représentation des différentes zones du monomère (chaîne jaune) impliquées dans les interfaces. La zone du monomère impliquée dans l'interface  $I_1$  est coloriée en violet, celle de  $I_2$  en noir et celle de  $I_3$  en bleu.

### 3.1.2 Sous-unité B pentamérique de l'entérotoxine thermolabile

L'entérotoxine thermolabile (LT) est produite par des souches entérotoxigènes d'*Escherichia coli* (73,74) et possède une masse d'environ 84 kDa (75). LT est constituée de deux sous-unités, une sous-unité A (LTA) et une sous-unité B pentamérique (LTB<sub>5</sub>). Comme pour la toxine cholérique (CT), l'interaction entre la sous-unité pentamérique (LTB<sub>5</sub>) et le ganglioside GM<sub>1</sub> des cellules épithéliales de l'intestin permet l'internalisation de la sous-unité A dans les cellules (65,73). LT est responsable des mêmes symptômes que CT, mais de manière plus bénigne avec la 'diarrhée du voyageur' (65).

LTB<sub>5</sub> est constituée de 5 monomères de 11 700 Dalton (Fig. 3-2). Chaque monomère de LTB est composé de 103 acides aminés organisés en deux hélices alpha et six brins beta (76). De plus, chaque monomère de LTB interagit avec ces monomères adjacents via des liaisons hydrogènes, interactions hydrophobes... (73). Comme pour CtxB<sub>5</sub>, LTB<sub>5</sub> possède trois interfaces entre deux chaînes : une interface  $I_1$  dans le domaine N-terminal, une interface  $I_2$  entre les brins beta et une interface  $I_3$  entre les hélices alpha.

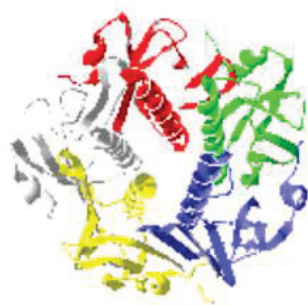


Figure 3-2. Sous-unité B pentamérique de l'entérotoxine thermolabile (PDB : 1LTR).

En conclusion, CT et LT font toutes les deux parties de la famille des toxines AB<sub>5</sub>, soit des toxines hétéro-hexamérique composées d'une sous-unité A toxique et d'une sous-unité B

pentamérique (77). Les deux toxines (CT et LT) possèdent des structures 3D mais aussi des fonctions similaires (65). De plus, les sous-unité B pentamérique de CT et LT possèdent une identité de séquence élevée (supérieure à 80 %), des structures 3D et des fonctions similaires (77).

### 3.1.3 Lysozyme

Le lysozyme est une protéine enzymatique dégradant la paroi des cellules microbiennes en hydrolysant les liaisons glycosidiques  $\beta$  (1-4) entre l'acide N-acétyl muramique et l'acide N-acétyl glucosamine du peptidoglycane (78). Le peptidoglycane (ou muréine) est le principal composant de la paroi cellulaire bactérienne, son hydrolyse provoque la lyse cellulaire. Par conséquent, le lysozyme possède une fonction bactéricide (79).

Le lysozyme que nous avons utilisé est le lysozyme de poule, une protéine monomérique de 129 acides aminés (Fig. 3-3) avec une masse moléculaire de 14 300 Dalton (79,80). La protéine est divisée en deux domaines : un domaine alpha et un domaine beta. Le domaine alpha est composé de quatre hélices alpha et d'une hélice  $3_{10}$  tandis que le domaine beta est constitué d'un feuillet beta antiparallèle de trois brins et d'une hélice  $3_{10}$  (80–82). Un feuillet beta double brins antiparallèle relie les deux domaines (80).



Figure 3-3. Lysozyme (code PDB : 1DPX).

### 3.1.4 Choix des protéines

Ces trois protéines (CtxB<sub>5</sub>, LTB<sub>5</sub> et le lysozyme) possèdent une résistance fonctionnelle à la lyophilisation. Par conséquent, la déshydratation de ces protéines ne devrait pas entraîner de changements conformationnels irréversibles. Cela est une propriété importante car les mesures diélectriques sont effectuées dans des conditions faiblement hydratée où l'échantillon protéique devrait contenir uniquement du PBS (solution saline dans un tampon phosphate) et de l'eau liée. Pour rappel, en plus du faible niveau d'hydratation les échantillons protéiques utilisés sont aussi nanoconfinés. D'après l'étude de Soubias et Gawrish en 2013, le confinement ne devrait pas impacter la fonction protéique (49).

La sous-unité B pentamérique de la choléra toxine est choisie en raison de sa stabilité élevée. CtxB<sub>5</sub> possède un fort pourcentage d'identité de séquence, une structure tridimensionnelle superposable et une fonction similaire à la sous-unité B pentamérique de l'entérotoxine thermolabile. De plus, les mécanismes d'assemblages et de désassemblages macroscopique de ces deux toxines sont connus et présentent une différence dans les intermédiaires de dépliement (73,77). Par conséquent, ces deux protéines sont adaptées pour

déterminer si des variants de séquence sont détectables et distinguables en utilisant le nanoconfinement et la spectroscopie diélectrique à large bande.

Le lysozyme est choisi car il s'agit d'une protéine globalement différente des deux autres protéines étudiées (monomère) possédant une stabilité thermique élevée (80,83,84). Les différences de structure et de fonction entre le lysozyme et les deux toxines permettront de déterminer si des mutations pathologiques sont identifiables en combinant le nanoconfinement et la spectroscopie diélectrique à large bande.

## 3.2 Méthodes

### 3.2.1 Membrane AAO

Les membranes AAO (anodic aluminium oxide) sont des membranes inorganiques fabriquées à partir d'aluminium oxyde ( $\text{Al}_2\text{O}_3$ ). L'oxydation de l'aluminium à lieu pendant l'anodisation et permet de changer la surface de l'échantillon. Deux étapes d'anodisation sont nécessaires pour créer des nanopores cylindriques, parallèles, identiques et ordonnés (Fig. 3-4).

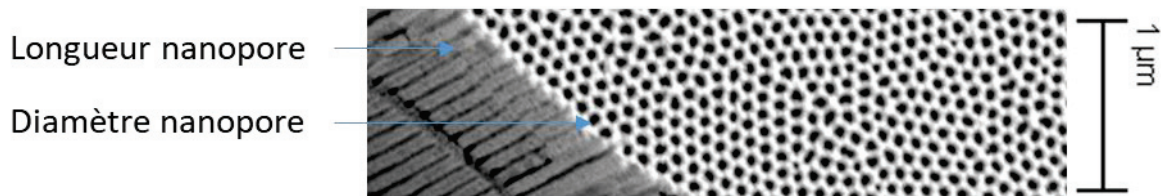
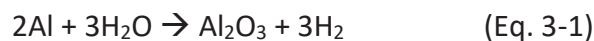


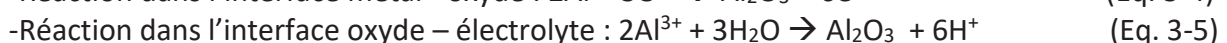
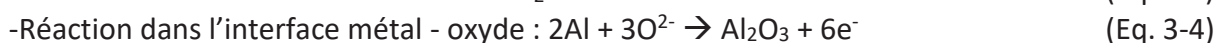
Figure 3-4. Membrane avec des nanopores cylindriques, parallèles, identiques et ordonnés.

Les ions  $\text{Al}^{3+}$  libérés par le substrat d'aluminium dans l'électrolyte permettent la formation d'une couche d'oxyde à l'interface aluminium-électrolyte. Suite à la formation de cette couche d'oxyde ( $\text{Al}_2\text{O}_3$ ), deux interfaces apparaissent : l'interface métal-oxyde et l'interface oxyde-électrolyte. Pendant l'anodisation, la couche d'oxyde est régénérée avec une croissance continue d'oxyde (Fig. 3-5) et transformée en une couche avec des structures semi-sphériques d'épaisseur constante conduisant à la formation de pore. La formation des pores est due à la migration des dérivés de l'oxygène ( $\text{O}^{2-}$  ou  $\text{OH}^-$ ) de la solution d'électrolyte vers la couche d'oxyde (85,86).

L'équation générale du processus d'anodisation de l'aluminium est définie par l'équation 3-1 :



Cette équation est obtenue à partir des équations suivantes (Eq.3-2 à Eq. 3-5) :





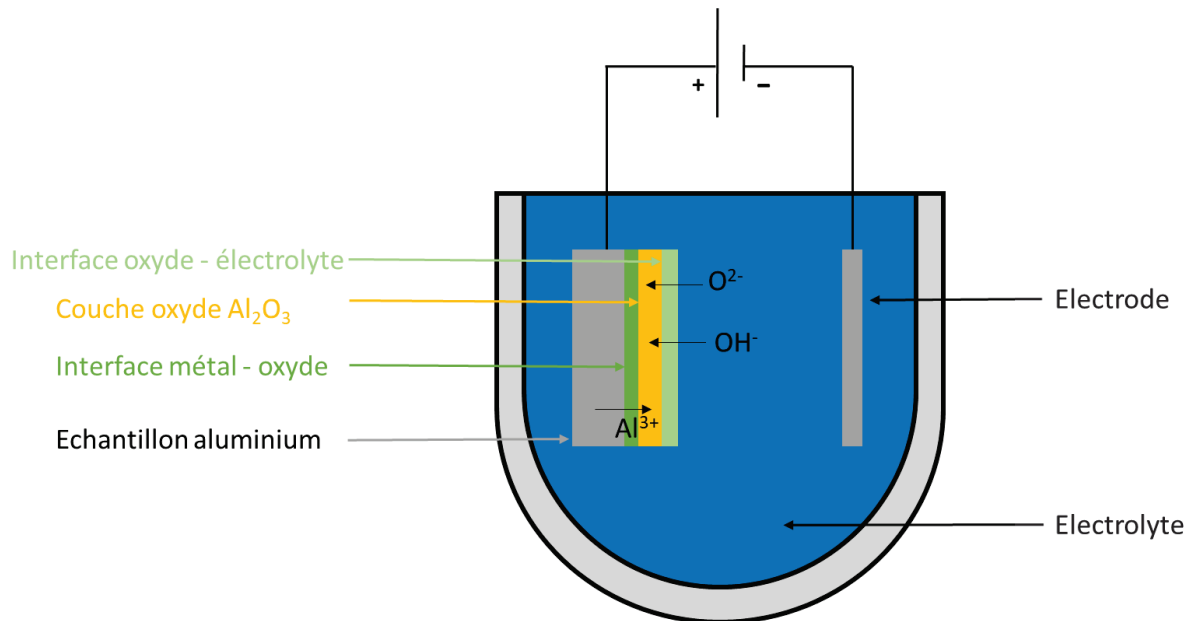


Figure 3-5. Formation d'une couche d'oxyde entre la plaque d'aluminium et la solution d'électrolyte.

Après le premier processus d'anodisation, des pores de différentes tailles, formes et orientations sont obtenus dans la couche supérieure de la couche d'oxyde mais des pores ordonnés commencent à apparaître dans la couche d'oxyde inférieure. Un traitement chimique est réalisé pour éliminer la couche supérieure de la couche d'oxyde possédant les pores non ordonnés puis un second processus d'anodisation est réalisé dans les mêmes conditions que le premier pour allonger les pores ordonnés de la couche d'oxyde inférieure.

Les dimensions des pores (longueur, diamètre et distance interpore) sont contrôlées par différents paramètres pendant l'anodisation :

- La nature et la concentration de l'électrolyte
- La température de la solution d'électrolyte
- Le voltage
- La durée de l'anodisation

Le diamètre des pores varie suivant la nature de l'électrolyte utilisé. Le diamètre des pores augmente avec l'augmentation de la température de la solution d'électrolyte, le voltage et la durée d'anodisation. Le diamètre des pores diminue avec l'augmentation de la concentration d'électrolyte. La longueur des pores augmente avec l'augmentation de la concentration de l'électrolyte et avec l'augmentation de la durée d'anodisation. La distance interpore correspond à la distance entre les centres de deux pores. Cette distance augmente avec l'augmentation du voltage (85–89). La porosité dépend du diamètre des pores et de la distance interpore, par conséquent elle est influencée par le voltage, la température et la durée d'anodisation (88).

### 3.2.2 Spectroscopie diélectrique à large bande

L'application d'un champ électrique induit le mouvement des charges électriques entraînant la création d'un courant électrique. Il existe deux types de charges électriques : charges libres et charges liées. Les charges libres correspondent aux électrons et ions alors que les charges liées correspondent aux moments de dipôles.



Les équations de Maxwell, peuvent être résumés par l'équation 3-6. Dans cette équation le courant correspond à la somme d'un courant continu ( $\vec{J}_f$ ) et d'un courant de déplacement ( $\vec{J}_d$ ) (Eq. 3-6) :

$$\vec{J}_{\text{total}} = (\vec{J}_f) + (\vec{J}_d) \quad (\text{Eq. 3-6})$$

Le courant continu correspond au déplacement de charges libres et est défini par l'équation 3-7 :

$$(\vec{J}_f) = \sigma \vec{E} \quad (\text{Eq. 3-7})$$

où  $\sigma$  est la conductivité et  $\vec{E}$  le champ électrique.

Le courant de déplacement correspond au déplacement de charges liées et est défini par l'équation 3-8 :

$$(\vec{J}_d) = \frac{\partial \vec{D}}{\partial t} \quad (\text{Eq. 3-8})$$

$\vec{D}$  correspond au déplacement diélectrique et est défini par l'équation 3-9 :

$$\vec{D} = \epsilon_0 \vec{E} + \vec{P} \quad (\text{Eq. 3-9})$$

où  $\epsilon_0$  est la permittivité de l'air.

$\vec{P}$  correspond à la polarisation électrique et est définie par l'équation 3-10 :

$$\vec{P} = \epsilon_0 \chi_e \vec{E} \quad (\text{Eq. 3-10})$$

où  $\chi_e$  est la susceptibilité électrique.

Par conséquent, le déplacement diélectrique est défini par l'équation 3-11 :

$$\begin{aligned} \vec{D} &= \epsilon_0 \vec{E} + \epsilon_0 \chi_e \vec{E} \\ &= \epsilon_0 (1 + \chi_e) \vec{E} = \epsilon_0 \epsilon_r \vec{E} = \epsilon \vec{E} \end{aligned} \quad (\text{Eq. 3-11})$$

où  $\epsilon_r$  correspond à la permittivité relative ( $1 + \chi_e$ ) et  $\epsilon$  à la permittivité absolue.

La permittivité est caractérisée par une fonction complexe de  $\epsilon^*$  définie par :

$$\epsilon^* = \epsilon' - i\epsilon'' \quad (\text{Eq. 3-12})$$

où  $\epsilon'$  est la partie réelle et  $\epsilon''$  la partie imaginaire de la fonction complexe.

Par conséquent, le courant continu est associé à la conductivité alors que le courant de déplacement est associé à la permittivité.

Le courant de déplacement est associé à la polarisation électrique. Dans un matériau, trois types de polarisations existent : la polarisation électronique, la polarisation atomique et la polarisation d'orientation. La somme de ces trois contributions donne la polarisation totale (Eq. 3-13).

$$\vec{P}_{\text{totale}} = \vec{P}_{\text{électronique}} + \vec{P}_{\text{atomique}} + \vec{P}_{\text{orientation}} \quad (\text{Eq. 3-13})$$

La polarisation électronique correspond au déplacement d'électrons suite à l'application d'un champ électrique et se produit à une échelle de  $10^{-15}$  s. Cette polarisation se produit aussi dans les atomes neutres, lorsque le champ électrique externe déplace le noyau par rapport à la charge négative (Fig. 3-6). La polarisation électronique est décrite par l'équation 3-14 :

$$\vec{P}_{\text{électronique}} = n \alpha_e \vec{E} \quad (\text{Eq. 3-14})$$

avec  $n$  le nombre de densité,  $\alpha_e$  la polarisabilité électronique et  $\vec{E}$  le champ électrique.

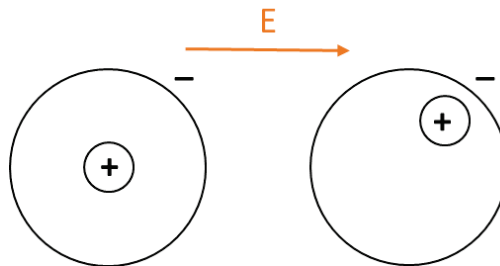


Figure 3-6. Représentation de la polarisation électronique.

La polarisation atomique correspond au déplacement des atomes constituant une molécule (Fig. 3-7) et dépend de la nature des atomes. Cette polarisation se produit à une échelle de  $10^{-12}$  s et est décrite par l'équation 3-15 :

$$\vec{P}_{\text{atomique}} = n \alpha_a \vec{E} \quad (\text{Eq. 3-15})$$

avec  $n$  le nombre de densité,  $\alpha_a$  la polarisabilité atomique et  $\vec{E}$  le champ électrique.



Figure 3-7. Représentation de la polarisation atomique.

Chaque atome possède une électronégativité qui représente la capacité de cet atome à attirer les électrons d'un autre atome pour former une liaison chimique (liaison covalente ou liaison ionique). La table 3-1, représente les valeurs d'électronégativité des atomes présents dans les protéines.

Atomes	Electronégativité
Hydrogène (H)	2,20
Carbone (C)	2,55
Souffre (S)	2,58
Azote (N)	3,04
Oxygène (O)	3,44

Table 3-1. Valeur d'électronégativité des atomes présents dans une protéine déterminée avec l'échelle de Pauling.

Deux atomes possédant une électronégativité différente et liés dans une molécule entraînent la création d'un dipôle permanent. Le dipôle permanent résulte de la distribution asymétrique des charges dans une molécule (Fig. 3-8a). Son moment dipolaire  $\mu$ , défini par l'interaction de deux atomes asymétriques séparés par une distance  $d$  et possédant chacun une charge électrique (charge  $+q$  et  $-q$ ) est exprimé par l'équation 3-16 :

$$\mu = q * d \quad (\text{Eq. 3-16})$$

L'interaction entre deux atomes symétriques (identiques mais aussi avec le même environnement) ne possède pas de moment dipolaire permanent (Fig. 3-8b) mais l'application d'un champ électrique peut donner lieu à une polarisation induite.



Figure 3-8. Représentation des moments dipolaires. a) Moment dipolaire permanent présent entre deux atomes asymétriques. b) Création d'un moment dipolaire induit lors de l'application d'un champ électrique entre deux atomes symétriques possédant le même environnement.

Sous l'application d'un champ électrique, les dipôles permanents ont la tendance de s'aligner dans la direction du champ électrique (Fig. 3-9) donnant lieu à une polarisation d'orientation. La polarisation d'orientation est définie par l'équation 3-17 :

$$\vec{P}_{\text{orientation}} = \frac{n \mu^2}{3 k_B T} \vec{E} \quad (\text{Eq. 3-17})$$

avec  $n$  le nombre de densité des dipôles,  $\mu$  le moment dipolaire,  $k_B$  la constante de Boltzmann,  $T$  la température absolue et  $\vec{E}$  le champ électrique.

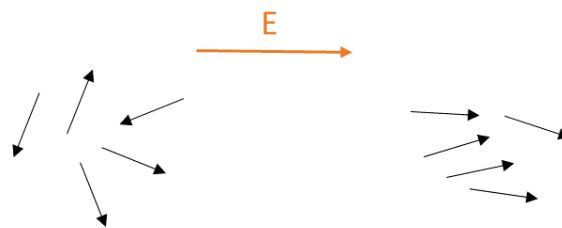


Figure 3-9. Représentation de la polarisation d'orientation.

En conclusion, dans les hautes fréquences ( $>10^{12}$  Hz), la fonction complexe de permittivité est dominée par la polarisation électronique et par la polarisation atomique, tandis que, dans les fréquences plus faibles, la contribution dominante reste la polarisation d'orientation.

La figure 3-10 montre la formation de la polarisation électronique et atomique et de la polarisation d'orientation, suite à l'application d'un champ électrique.

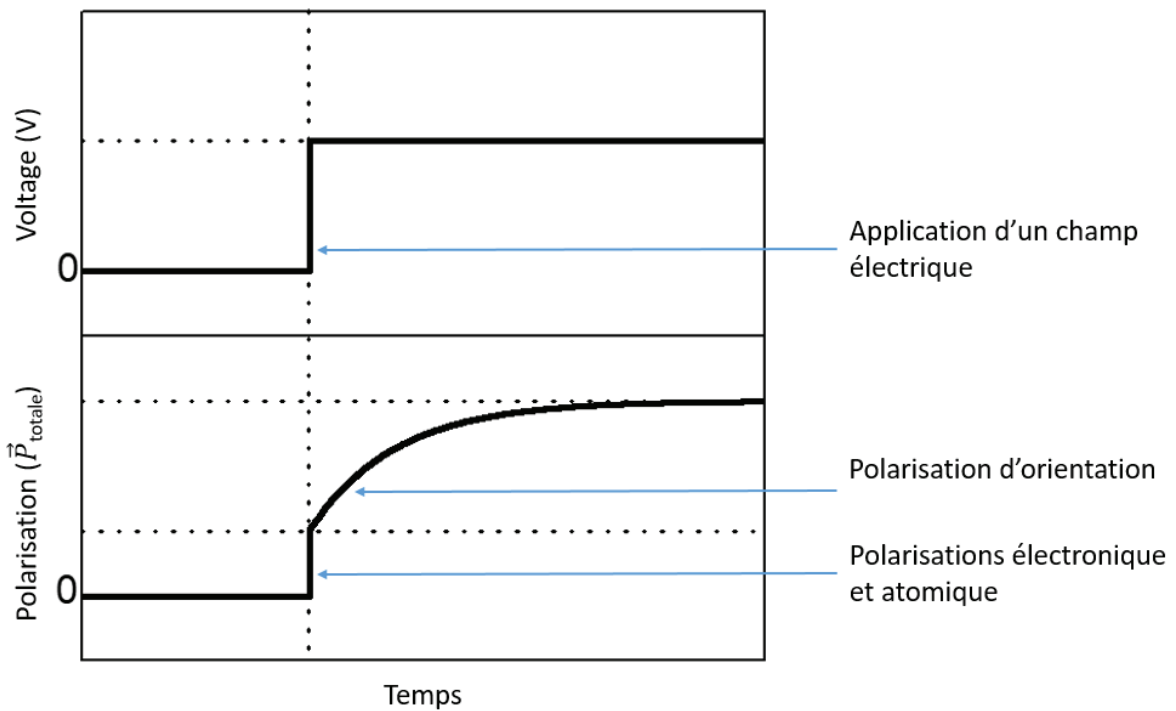


Figure 3-10. Détection des polarisations dans le temps suite à l'application d'un champ électrique.

Le processus d'orientation des moments dipolaires permanents se produit par des fluctuations moléculaires donnant lieu à un processus de relaxation diélectrique, avec un temps de relaxation caractéristique déterminé par l'équilibre thermodynamique à une température donnée. L'application de la Transformée de Fourier (Eq. 3-18) sur les données en fonction de temps (Fig. 3-10) permet la détection de processus de relaxation en fonction de la fréquence (Fig. 3-11). Pour une température donnée, quand la fréquence augmente, la partie réelle de la permittivité complexe diminue tandis que les données de pertes diélectriques ( $\epsilon''$ ) mesurées dans la partie imaginaire présentent un pic de relaxation (Fig. 3-11).

$$X(\omega) = \int_{t=-\infty}^{\infty} x(t)e^{-j\omega t} dt$$

$$x(t) = \frac{1}{2\pi} \int_{\omega=-\infty}^{\infty} X(\omega)e^{j\omega t} d\omega \quad (\text{Eq. 3-18})$$

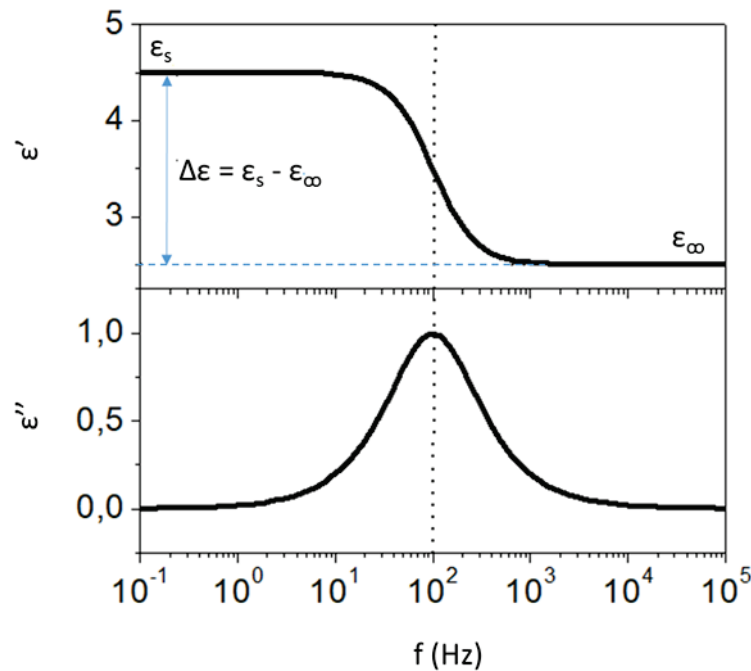


Figure 3-11. Représentation du complexe de permittivité en fonction de la fréquence.

En 1912, Debye a développé un modèle (modèle de Debye représenté dans la figure 3-11) qui démontre que les pics de relaxation diélectrique sont produits par des fluctuations des moments dipolaires permanents. Le processus de relaxation des dipôles lors de l'application d'un champ électrique dans le modèle de Debye suit l'équation 3-19 :

$$\varepsilon^*(\omega) = \varepsilon_{\infty} + \frac{\varepsilon_s - \varepsilon_{\infty}}{1 + i\omega\tau} \quad (\text{Eq. 3-19})$$

où  $\varepsilon_{\infty}$  correspond à la permittivité à haute fréquence,  $\varepsilon_s$  la permittivité statique à basse fréquence et  $\tau$  le temps de relaxation de la molécule.

La partie réelle et la partie imaginaire du complexe de permittivité sont séparées comme suit (Eq. 3-20) :

$$\varepsilon'(\omega) = \varepsilon_{\infty} + \frac{\Delta\varepsilon}{1 + \tau_D^2\omega^2} \quad \text{et} \quad \varepsilon''(\omega) = \frac{\omega\tau_D\Delta\varepsilon}{1 + \tau_D^2\omega^2} \quad (\text{Eq. 3-20})$$

$\tau_D$  correspond au taux de relaxation du moment dipolaire fluctuant.

La position du pic de relaxation (température où l'intensité du pic est maximale :  $\theta_{\max}$  ou fréquence où l'intensité du pic est maximale :  $\omega_{\max}$ ) varie en fonction de la fréquence et de la température. La fréquence du pic de relaxation augmente avec l'augmentation de la température (Fig. 3-12a). Par conséquent la température doit augmenter pour détecter un processus de relaxation à plus haute fréquence. De même, la température du pic de relaxation augmente avec l'augmentation de la fréquence (Fig. 3-12b).

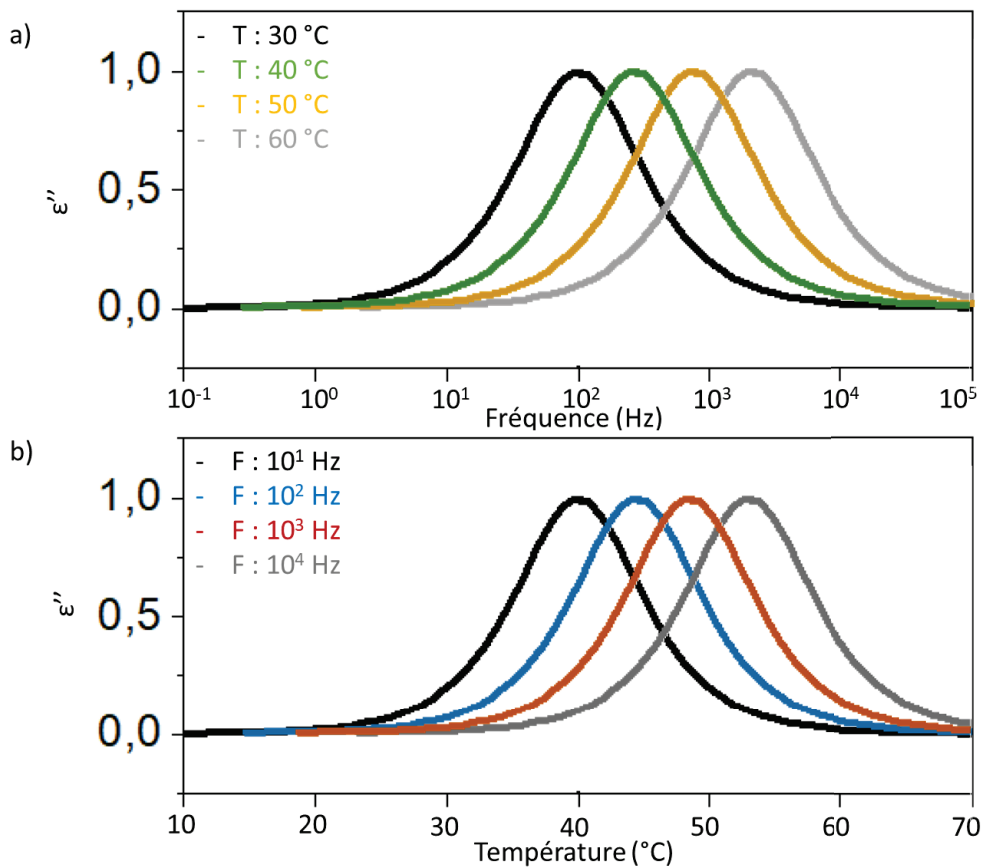


Figure 3-12. Représentation des données de pertes diélectriques en fonction de la fréquence et de la température. a) Représentation des pics de relaxation en fréquence à différentes températures. b) Représentation des pics de relaxation en température à différentes fréquences.

La position où l'intensité du pic de relaxation est maximale (Fig. 3-11 et 3-12) donne le temps de relaxation (Eq. 3-21) :

$$\tau = \frac{1}{f} \quad (\text{Eq. 3-21})$$

où  $\tau$  correspond au temps de relaxation et  $f$  à la fréquence.

La dépendance du temps de relaxation en fonction de la température peut être modélisée par une dépendance linéaire de type Arrhenius (Eq. 3-22) ou non linéaire de type Vogel-Fulcher-Tammann (VFT) (Eq. 3-23) comme représenté dans la figure 3-13.

$$\text{Equation d'Arrhenius : } \tau = \tau_0 \exp (E_A / k_B T) \quad (\text{Eq. 3-22})$$

$$\text{Equation VFT : } \tau = \tau_0 \exp \left( \frac{D T_0}{T - T_0} \right) \quad (\text{Eq. 3-23})$$

Où  $\tau_0$  est le temps de relaxation extrapolé à une température infinie,  $T$  la température,  $E_A$  l'énergie d'activation,  $k_B$  la constante de Boltzmann,  $D$  le paramètre de fragilité et  $T_0$  la température où  $\tau$  irait à l'infini (58).

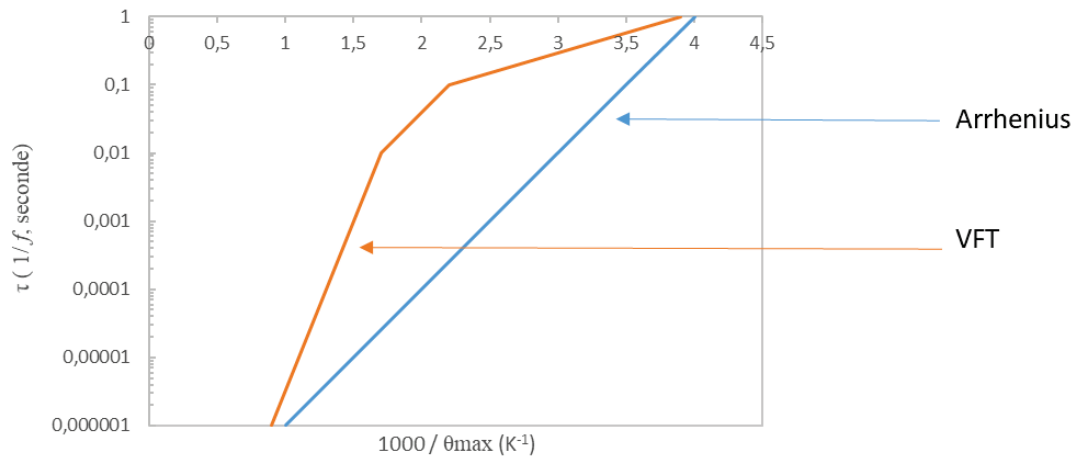


Figure 3-13. Représentation de la dépendance en température d'un processus de relaxation en fonction des temps de relaxation.

La spectroscopie diélectrique à large bande est la méthode expérimentale qui nous permet de mesurer la dynamique de relaxation diélectrique (ou tout simplement la dynamique moléculaire) dans une large gamme de fréquence ( $10^{-6}$  Hz à  $10^{12}$  Hz). Elle est donc adaptée pour caractériser la dynamique moléculaire des protéines. Les protéines présentant des mouvements de chaînes latérales, de structures secondaires, de repliement, dépliement à l'échelle de la nanoseconde à la seconde (1 Hz à  $10^9$  Hz) (Fig. 2-7). Les informations sur la dynamique moléculaire sont obtenues par l'analyse de la fonction de permittivité (partie réelle ou partie imaginaire) (86,90). Dans les figures 3-11 et 3-12, la dynamique moléculaire de l'échantillon est caractérisée par un seul pic de relaxation, toutefois un échantillon peut être plus complexe et caractérisé par plusieurs processus de relaxation et donc montrer plusieurs pics. Chaque pic de relaxation est défini par sa position en température ou en fréquence, son intensité et sa largeur.

### 3.2.3 Analyse des données de pertes diélectriques

Pour déterminer les dipôles contribuant au signal diélectrique dans le ou les processus de relaxations détectés, le logiciel Swiss-PDB Viewer est utilisé avec les données des protéines étudiées issues de la banque de donnée PDB (protein data bank). Le logiciel Swiss-PDB Viewer permet de visualiser la conformation protéique d'une protéine (structure tertiaire, quaternaire) et de déterminer pour chaque acide aminé (chaîne latérale, squelette peptidique) les acides aminés proches spatialement. Le voisinage de chaque acide aminé (seuil 5 Å) et plus particulièrement des acides aminés chargés est analysé pour de déterminer les dipôles impliqués dans les processus de relaxation détectés (91). La structure native tridimensionnelle de la protéine est donc utilisée pour proposer l'ensemble des dipôles existant dans la protéine.

Les acides aminés chargés sont considérés comme contributeurs du signal diélectrique car le signal diélectrique du BDS est sensible aux groupements polaires et chargés (59). Chaque acide aminé est composé de différents atomes qui peuvent interagir avec les atomes d'un acide aminé voisin. L'interaction entre des atomes différents issus d'acides aminés différents forme un dipôle ionique si l'interaction a lieu entre deux acides aminés chargés et un dipôle

induit si l'interaction a lieu entre un acide aminé chargé et un acide aminé non chargé. En considérant que les dipôles ioniques et induits participent au signal diélectrique, les données diélectriques seront analysées en termes structurales.

Les dipôles ioniques peuvent avoir des forces électrostatiques attractives si la charge des deux acides aminés est différente ou des forces électrostatiques répulsives si la charge des deux acides aminés est identique.

Les codes PDB 1EEI, 1LTR et 1DPX sont utilisés pour étudier CtxB<sub>5</sub>, LTB<sub>5</sub> et le lysozyme, respectivement.

### 3.3 Méthodologie

#### 3.3.1 Membrane avec des nanopores

Les films (membranes) d'oxyde aluminium sur aluminium (AAO) avec des nanopores aux dimensions souhaités sont achetés chez le fournisseur InRedox. La dimension des membranes est de 10 mm \* 10 mm (Fig. 3-14a). Les dimensions du confinement (nanopore) dans les membranes sont vérifiées par microscopie électronique à balayage (MEB) pour une membrane de chaque lot commandé (Fig. 3-14b). Les membranes AAO contenant des nanopores sont appelées nanomembranes. Au vue des dimensions des nanomembranes, chaque nanomembrane possède des centaines de milliers de nanopores, par conséquent les mesures diélectriques détiennent une grande statistique des états structuraux.

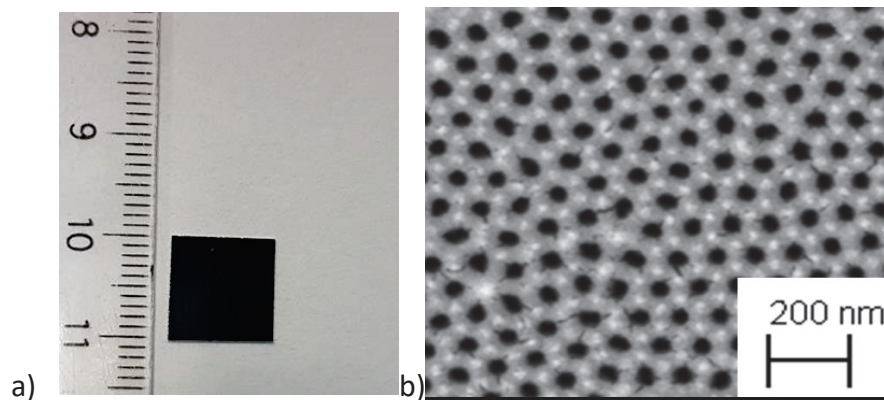


Figure 3-14. Nanomembrane achatée chez InRedox. a) Nanomembrane de dimension 10 mm \* 10mm. b) Nanomembrane avec des nanopores de 40 nm de diamètre observée au MEB.

Cinq types de nanomembranes avec les confinements suivant sont achetées chez InRedox :

- 40 nm de diamètre et 10  $\mu$ m de long
- 40 nm de diamètre et 5  $\mu$ m de long
- 40 nm de diamètre et 2  $\mu$ m de long
- 120 nm de diamètre et 2  $\mu$ m de long
- 20 nm de diamètre et 2  $\mu$ m de long

Comme les confinements sont de l'ordre du nm\* $\mu$ m, nous parlerons d'une population de conformation enrichie et/ou sélectionnée et non d'une conformation protéique unique.



Dans la table 3-2, les dimensions des confinements fournies par InRedox sont représentées :

Longueur des nanopores : $10 \pm 0.2 \mu\text{m}$ Diamètre des nanopores : $40 \pm 4 \text{ nm}$ Porosité : $12 \pm 2 \%$
Longueur des nanopores : $2 \pm 0.2 \mu\text{m}$ Diamètre des nanopores : $40 \pm 4 \text{ nm}$ Porosité : $12 \pm 2 \%$
Longueur des nanopores : $2 \pm 0.1 \mu\text{m}$ Diamètre des nanopores : $120 \pm 14 \text{ nm}$ Porosité : $16 \pm 2 \%$
Longueur des nanopores : $2 \pm 0.1 \mu\text{m}$ Diamètre des nanopores : $20 \pm 3 \text{ nm}$ Porosité : $12 \pm 2 \%$
Longueur des nanopores : $5 \pm 0.3 \mu\text{m}$ Diamètre des nanopores : $40 \pm 4 \text{ nm}$ Porosité : $12 \pm 2 \%$

*Table 3- 2. Dimensions des confinements utilisées.*

Avant de pouvoir être utilisées, les nanomembranes subissent un traitement au plasma pendant vingt minutes pour éliminer toutes les impuretés organiques encore présentes. Pour réaliser les mesures diélectriques, la nanomembrane (Fig. 3-14a) est ensuite intercalée entre deux électrodes de 7 mm et 40 mm de diamètre puis placée dans le spectromètre diélectrique à large bande.

### 3.3.2 Mesure diélectrique par spectroscopie diélectrique à large bande

La nanomembrane (Fig. 3-14a) est placée entre les deux électrodes et incorporée dans le spectromètre diélectrique à large bande. Les mesures diélectriques sont réalisées à l'aide d'un spectromètre diélectrique à large bande Novocontrol Alpha analyseur (Fig. 3-15) sur une plage de fréquence de 1 Hz à  $10^6$  Hz et sur une plage de température de  $-80 \text{ }^\circ\text{C}$  à  $180 \text{ }^\circ\text{C}$ . Les rampes de températures sont réalisées avec une vitesse de 2 K/min et la tension est de 0,2 V. Pour contrôler la température, un flux d'azote gazeux pur est utilisé dans un cryostat fermé, fournissant des conditions expérimentales sans oxygènes et sans eau.



Figure 3-15. Photos de l'appareil de Spectroscopie Diélectrique à Large Bande. A gauche : Photo générale de l'appareil de mesure diélectrique. A droite : Photo du support, où la nanomembrane est déposée entre deux électrodes.

Les propriétés diélectriques sont déterminées via la mesure de la capacité diélectrique complexe définie comme suit (Eq. 3-24 à Eq. 3-28) :

$$\frac{U}{I} = \frac{1}{\omega C} \quad (\text{Eq. 3-24})$$

$$C = C' - iC'' \quad (\text{Eq. 3-25})$$

$$C = \epsilon \frac{A}{d} = \epsilon_0 \epsilon_r \frac{A}{d} \quad (\text{Eq. 3-26})$$

$$C' = \epsilon_0 \epsilon' \frac{A}{d} \quad (\text{Eq. 3-27})$$

$$C'' = \epsilon_0 \epsilon'' \frac{A}{d} \quad (\text{Eq. 3-28})$$

où C correspond à la capacité, U au voltage, I à l'intensité, C' la partie réelle et C'' la partie imaginaire de la fonction complexe, A au diamètre de l'électrode de mesure et d à la longueur des nanopores présents dans la nanomembrane étudiée (10  $\mu\text{m}$ , 5  $\mu\text{m}$  ou 2  $\mu\text{m}$ ).

### 3.3.3 Echantillon protéique

#### 3.3.3.1 Conservation de la protéine

Les protéines étudiées sont la sous-unité B pentamérique de la choléra toxine (C9903), la sous-unité B pentamérique de l'entérotoxine thermolabile (E8656) et le lysozyme de poule (62970) lyophilisées et achetées chez Sigma Aldrich.

Les protéines (CtxB<sub>5</sub>, LTB<sub>5</sub> et lysozyme) sont diluées dans une solution saline de phosphate (PBS : phosphate buffered saline composé de 10 mM de sodium de phosphate et 150 mM chlorure de sodium à un pH 7,4). Cette préparation de protéine est aliquotée et conservée à -20 °C. Chaque aliquot peut-être décongelé et recongelé trois fois au maximum, ainsi trois mesures diélectriques peuvent-être effectuées pour chaque aliquot.

Une solution dix fois concentrée de PBS est achetée chez Biosolve et diluée dix fois dans de l'eau distillée pour préparer la solution utilisée pour conserver la protéine à -20 °C. La solution de PBS 10X et d'eau distillée sont toutes les deux filtrées via des filtres de 22 µm pour limiter la contamination par des impuretés.

### 3.3.3.2 Vérification de la conformation protéique

Avant toute expérience diélectrique à partir d'un aliquot, les conformations protéiques des toxines et du lysozyme sont vérifiées par fluorescence. Deux échantillons protéiques sont testés, l'un est dilué dans de l'eau déionisée filtrée et l'autre est dilué dans une solution de PBS pour avoir une concentration finale d'environ 0,1 mg/ml.

La spectrofluorimétrie est une technique utilisée pour l'étude structurale des protéines en solution. L'intensité de fluorescence et la longueur d'onde maximum d'émission, fournissent des informations sur la conformation de la protéine. L'absorption correspond à l'émission d'une radiation lumineuse dans toutes les directions de l'espace. Cette radiation va exciter les fluorophores de la protéine et entraîner le passage de ces fluorophores d'un état fondamental singulet stable à un état singulet excité instable de haute énergie. La restitution de l'énergie absorbée par le fluorophore (passage de l'état singulet excité de plus basse énergie à l'état fondamental singulet stable) correspond à la fluorescence. La fluorescence correspond à une réémission rapide de photons dans toutes les directions de l'espace, avec émission d'une radiation lumineuse (92,93). La fluorescence est une méthode très sensible à l'environnement des fluorophores.

Les protéines possèdent une fluorescence intrinsèque grâce aux groupements fluorophores constitués par les chaînes latérales aromatiques des acides aminés (phénylalanine, tyrosine et tryptophane). Ces acides aminés aromatiques sont différents par leurs longueurs d'ondes d'excitations et d'émissions.

CtxB<sub>5</sub> possède un seul tryptophane (Trp) en position 88, situé à l'interface de deux monomères. D'après des travaux précédents (71,72), l'intensité de Trp mesure principalement l'état d'assemblage de la toxine permettant ainsi de suivre la dissociation ou la formation de CtxB<sub>5</sub> de par sa position à l'interface entre chaque monomère de la toxine. L'intensité de fluorescence de Trp est donc mesurée pour déterminer la conformation protéique de l'échantillon et vérifier qu'il s'agisse bien d'un pentamère (94). L'intensité de fluorescence est aussi utilisée pour regarder la conformation pentamérique de LTB et pour suivre sa dissociation (73). De même, l'intensité de fluorescence surveille les variations conformationnelles du lysozyme (95). Toutefois contrairement aux toxines CtxB<sub>5</sub> et LTB<sub>5</sub> qui possèdent un seul tryptophane en position 88 de chaque monomère, le lysozyme possède six tryptophanes avec trois tryptophanes dans le site de liaison du substrat (71,73,95).

L'appareil de fluorescence est un spectromètre de fluorescence Hitachi F-4500. La longueur d'onde d'émission pour une excitation à 295 nm (Trp) est mesurée de 320 à 380 nm. Les fentes d'excitation et d'émission sont de 5,0 nm. La vitesse de balayage est de 60 nm/min.

### 3.3.3.3 Préparation de la protéine pour les mesures diélectriques

Si les valeurs de fluorescence (longueur d'onde d'émission et intensité) des toxines correspondent aux valeurs des toxines sous forme pentamérique alors l'aliquot est dilué dans de l'eau déionisée préalablement filtrée. De même si les valeurs de fluorescence correspondent à la conformation monomérique pour le lysozyme. Les échantillons protéiques utilisés pour les mesures diélectriques sont composés de protéine, PBS et d'eau (distillée et déionisée).

La solution protéique est préparée en fonction du pourcentage de remplissage volumique souhaité dans les nanopores. Le pourcentage de remplissage volumique dans les nanopores ne doit pas être trop élevé pour limiter les effets d'encombrements macromoléculaires mais suffisamment important pour avoir un processus de relaxation détectable. Grâce à la densité protéique, les quantités de protéines et d'eau déionisée sont déterminées pour avoir le pourcentage de remplissage volumique désiré. Sauf précision contraire le taux de remplissage sera de 14,3 %.

Pour rappel, la sous-unité B pentamérique de la choléra toxine et la sous-unité B pentamérique de l'entérotoxine thermolabile utilisées sont composées de cinq monomères de 11 700 Dalton. Par conséquent, l'équation de Fischer et al, en 2004 (Eq. 3-29) détermine une densité protéique de 1,47 g/cm<sup>3</sup> par monomère (96). Le lysozyme est constitué d'un seul monomère de 14 300 Dalton (79), par conséquent la densité protéique est de 1,46 g/cm<sup>3</sup>.

$$\rho = [1,410 + 0,145 \cdot \exp(-\frac{M(kDa)}{13})] \quad (\text{Eq. 3-29})$$

Une goutte de 200 µL de solution est déposée sur la nanomembrane, puis chauffé à 50 °C pendant quinze minutes pour permettre aux protéines de rentrer dans les nanopores et évaporer ensuite l'eau libre. L'échantillon est ensuite refroidi à 30 °C pendant cinq minutes. La vitesse des changements de température est de 2 K/min. Cette procédure est nommée technique de la goutte. Une fois que les protéines sont rentrées dans les nanopores, la nanomembrane est placée entre deux électrodes pour réaliser la mesure diélectrique.

La même procédure est réalisée pour le contrôle du solvant, où le PBS (PBS 10X + eau distillée pour obtenir du PBS 1X) est dilué dans l'eau déionisée pour avoir le même ratio et une quantité équivalente à l'échantillon protéique. Ce contrôle est réalisé pour étudier la réponse diélectrique du PBS (PBS + eau distillée + eau déionisée) dans les mêmes conditions que la protéine.

#### *Exemple de calcul des quantités protéiques nécessaire pour un confinement donné :*

Protéine : sous-unité B pentamérique de la choléra toxine à une concentration de 2,6 mg/ml. Toutes les nanomembranes utilisées possèdent des dimensions de 10 mm \* 10 mm, soit une surface d'un cm<sup>2</sup>.

Dans cet exemple le confinement est de 40 nm de diamètre et 10 µm de long.

La porosité de la nanomembrane est de 12 % (0,12).

La densité de la protéine est de 1,47 g/cm<sup>3</sup>.

Volume d'un nanopore = surface nanomembrane (cm<sup>2</sup>) \* porosité nanomembrane \* longueur nanopores (cm) = 1 \* 0,12 \* 0,0010 = 0,00012 cm<sup>3</sup>

Remplissage des nanopores à 14,3 % = (14,3 \* 0,00012) / 100 = 0,0000172 cm<sup>3</sup>

Masse = densité \* volume nanopore = 1,47 \* 0,0000172 = 0,0000253 g = 0,0253 mg

Volume protéique = Concentration protéine / masse = 2,6 / 0,0253 = 103

Ainsi 1 µL de protéine à 2,6 mg/ml plus 102 µL d'eau déionisée sont nécessaires pour avoir un remplissage des nanopores à 14,3 %. Or comme 200 µL de solution protéique sont déposés sur la nanomembrane, la solution sera de 2 µL de protéine plus 204 µL d'eau déionisée. Les mêmes calculs sont réalisés pour les deux autres protéines étudiées. La densité des trois protéines est calculée à partir de la masse d'un monomère. Toutefois, le lysozyme est un monomère alors que les deux autres protéines sont des pentamères, par conséquent, le lysozyme ne possède pas d'interface et devrait donc être moins connecté que les deux autres protéines. Pour cette raison, le volume de protéine du lysozyme est multiplié par cinq. Dans l'exemple ci-dessus, la solution protéique serait composée de 10 µL de lysozyme à 2,6 mg/ml et de 204 µL d'eau déionisée.

#### 3.3.4 Essais diélectriques

La dynamique moléculaire est détectée en étudiant la permittivité complexe. La dynamique moléculaire est caractérisée par un pic de relaxation dans les données de pertes diélectriques ( $\epsilon''$ ) de la partie imaginaire de la permittivité complexe (Fig. 3-11). Pour étudier la dynamique moléculaire les données de pertes diélectriques seront donc analysées. La perte diélectrique ( $\epsilon''$ ) est mesurée en fonction de la température à différentes fréquences (partie 3.2.2). La variation de température s'effectue avec une vitesse de 2 K/min.

Le premier contrôle réalisé est celui de la nanomembrane vide pour vérifier l'absence de perte diélectrique. Pour cela, une fois que la nanomembrane est placée entre deux électrodes, elle est soumise à une température de 60 °C pendant dix minutes, puis la température diminue à -80 °C où elle est stabilisée pendant dix minutes et ensuite la température remonte à 30 °C. Les données de pertes diélectriques sont prélevées pendant chaque cycle de variation en température (refroidissement ou chauffage).

Pour l'échantillon protéique, une fois le processus de remplissage fini, la nanomembrane est chauffée à 60 °C pendant trois heures. Après ce traitement thermique, la température refroidit à -80 °C où elle est stabilisée pendant trente minutes avant de réchauffer à 60 °C. Un cycle de température est donc réalisé et celui-ci est composé d'une courbe de refroidissement (R) et d'une courbe de chauffage (C). L'échantillon protéique est de nouveau chauffé avec un traitement thermique de trois heures à 60 °C et un deuxième cycle de refroidissement/ chauffage est répété pour mesurer la reproductibilité du signal. Ensuite différents traitements thermiques à 80 °C, 100 °C, 140 °C et 180 °C sont réalisés pour étudier le comportement de la protéine dans un environnement confiné et en fonction de la

température (Fig. 3-16). Généralement, trois tests protéiques différents dans un même confinement sont réalisés pour assurer la reproductivité des données.

Pour le contrôle du solvant, la même rampe de température que pour l'échantillon protéique est réalisée. De même, deux tests de solvant au minimum sont réalisés dans chaque confinement pour vérifier la reproductibilité des résultats.

Les données de pertes diélectriques des courbes de refroidissement sont prélevées à la fin du traitement thermique de trois heures jusqu'à la fin des trente minutes à -80 °C (Fig. 3-16). Les valeurs de pertes diélectriques de la courbe de refroidissement après le premier traitement thermique à 60 °C sont nommées R1 60 °C, celles du deuxième traitement thermique à 60 °C R2 60 °C, celles après le traitement à 80 °C ; R 80 °C et ainsi de suite jusqu'à la courbe de refroidissement de 180 °C à -80 °C, nommée R 180 °C. Chaque expérience possède six courbes de refroidissement : R1 60 °C, R2 60 °C, R 80 °C, R 100 °C, R 140 °C et R 180 °C.

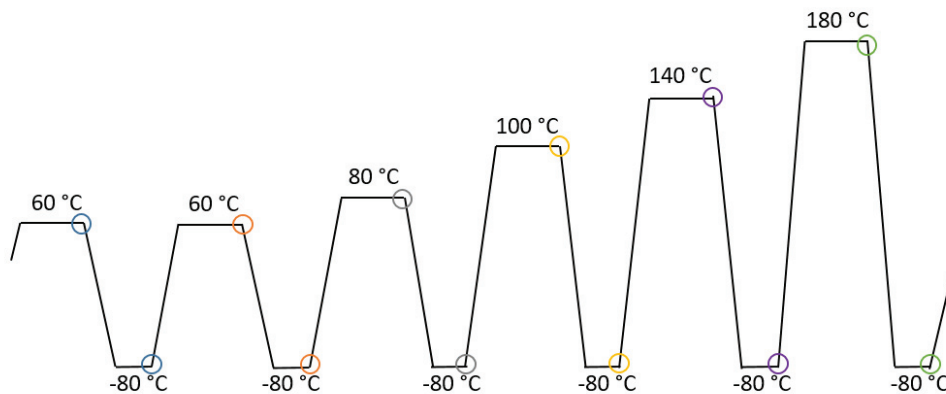


Figure 3-16. Rampe de température utilisée dans les expériences. Les ronds se réfèrent aux températures où les valeurs de pertes diélectriques sont prélevées pour obtenir une courbe de refroidissement.

Les données de pertes diélectriques peuvent aussi être analysées via les courbes de chauffage, et dans ce cas les valeurs de pertes diélectriques sont prélevées à la fin des trente minutes à -80 °C jusqu'à la température du traitement thermique. Comme les courbes de refroidissement, les courbes de chauffage seront nommées C1 60 °C, C2 60 °C, C 80 °C, C 100 °C... Dans tous les cas (courbe de refroidissement ou courbe de chauffage) les données de pertes diélectriques seront analysées en fonction de la température à différentes fréquences.

L'équation de la dérivée (Eq. 3-30) est utilisée pour tracer la dérivée en fonction de la température et ainsi déterminer la position du pic de relaxation en température [température où le signal d'intensité du pic ou des pics de relaxation est maximale ( $\theta_{max}$ )]:

$$\frac{\varepsilon''_{n+1} - \varepsilon''_{n-1}}{T_{n+1} - T_{n-1}} \quad (\text{Eq. 3-30})$$

Avec  $\varepsilon''$  correspondant aux valeurs de pertes diélectriques et T la température.



La dépendance en températures des pics de relaxation en fonction des temps de relaxation pour les protéines étudiées est modélisée par une équation d'Arrhenius (Eq. 3-22) (57,59).

Dans ce mémoire la spectroscopie diélectrique à large bande est utilisée sur les fréquences de 1 Hz à  $10^6$  Hz, soit des temps de relaxation de la seconde à la microseconde caractéristique de la dynamique des conformations protéiques. En effet, comme représenté dans la figure 2-7, les mouvements de la microseconde à la milliseconde sont associés aux mouvements des éléments de structures secondaires. Alors que les mouvements de la milliseconde à la seconde, sont associés aux mouvements de repliement, dépliement et plus précisément, aux mouvements des domaines, des interfaces, des chaînes mais aussi à l'assemblage, la dissociation des chaînes (3,53,54).

### 3.3.5 Analyse du spectre diélectrique

Un pic de relaxation est caractérisé par l'intensité, la position en température et/ou en fréquence, et la largeur du pic. Des exemples tirés d'étude de la dynamique de polymères en BDS sont présentés pour illustrer la dépendance du signal vis à vis de ces paramètres.

L'article de Serghei et al en 2010 (97) est utilisé pour illustrer le lien entre l'intensité du signal et la densité des dipôles. Les auteurs étudient l'écoulement du copolymère PS-b-PAVP (polystyrène-block-poly-4-vinylpyridine) dans les nanopores. L'écoulement du copolymère dans les nanopores entraîne une augmentation de l'intensité du pic de relaxation. L'intensité du pic de relaxation est proportionnelle à la densité de dipôle (Fig. 3-17).

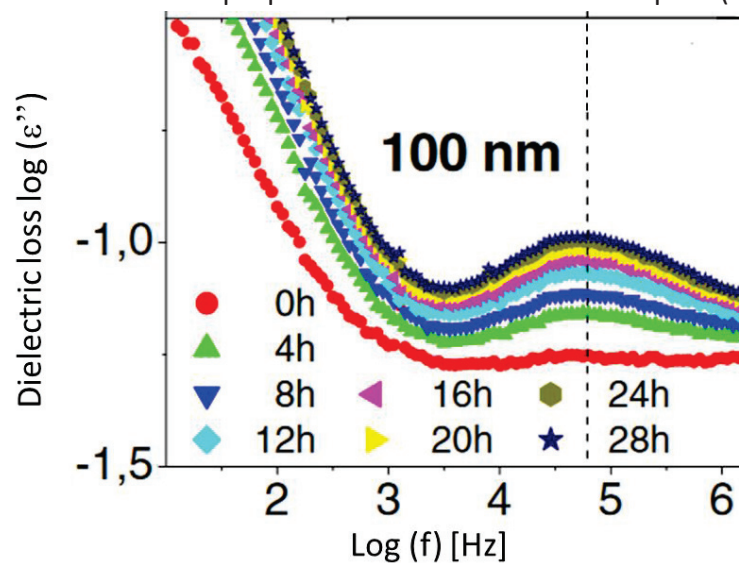


Figure 3-17. Variation de l'intensité du pic de relaxation en augmentant la quantité de copolymères dans les nanopores. Figure de l'article de Serghei et al, en 2010 (97).

L'article de Smiszerk-Lindert et al en 2018 (98) est utilisé pour illustrer comment la dépendance du signal en fréquence traduit un changement de dynamique lié à un changement de conformation. Dans la figure 3-18, le décalage de la position du pic de relaxation vers des fréquences plus basses dans le temps est associé à l'isomérisation de l'aberchrome 670 (conformère Z : trans et E : cis).

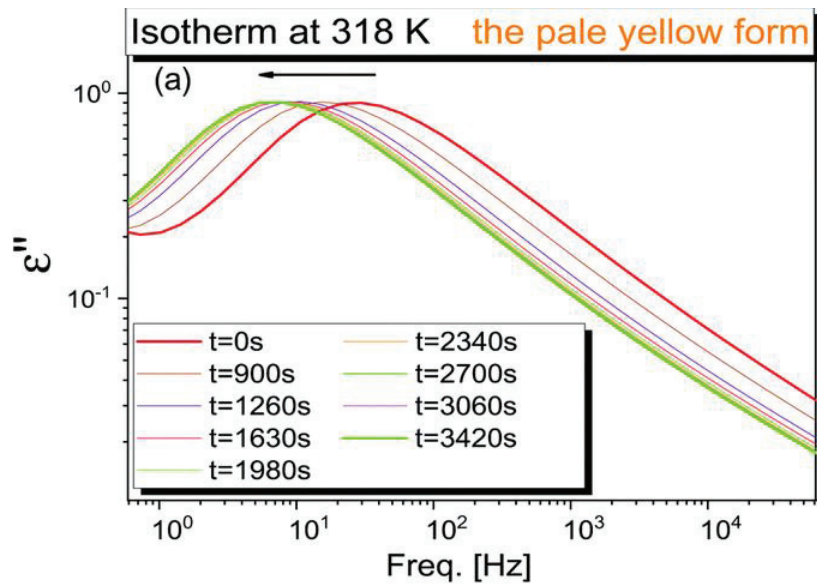


Figure 3-18. Variation de la position du pic de relaxation suite à un changement de conformation. Figure de l'article de Smiszerk-Lindert et al, en 2018 (98).

La largeur du pic de relaxation peut aussi augmenter indiquant une augmentation de l'hétérogénéité des fluctuations dipolaires, synonyme d'une dynamique résultante d'un mélange de dipôles.



## Chapitre 4

# Combinaison du nanoconfinement et de la spectroscopie diélectrique à la large bande pour l'étude de la dynamique des structures protéiques

Les travaux de ce chapitre 4, ont fait l'objet de deux publications présentées en annexe. La première publication intitulée : *Experimental Protein Molecular Dynamics : Broadband Dielectric Spectroscopy coupled with nanoconfinement* a été publiée dans *Scientific Reports* (Annexe A) et la deuxième en collaboration avec Lorenza Pacini intitulée *Analysis of Nanoconfined Protein Dielectric Signals Using Charged Amino Acid Network Models* a été publiée dans *Australian Journal of Chemistry* (Annexe B). Les résultats de ces deux publications sont réécrits pour en faire un chapitre.

Les protéines produisent leurs activités biologiques dans les organismes vivants via leur conformation protéique (structure 3D) et la dynamique moléculaire associée à cette conformation. La structure 3D repose sur les interactions entre les atomes des acides aminés qui composent la protéine.

Pour se former (repliement) et fonctionner la protéine couvre plusieurs échelles spatio-temporelles et propage les mouvements de façon dynamique d'une échelle locale (chaîne latérale) à une échelle plus large (structures secondaires, tertiaires, domaines, chaînes). Les mouvements temporels sont de l'ordre de la femtoseconde à la seconde, et les fluctuations moléculaires couvrent une gamme spatiale de l'Angströms aux nanomètres. Ces fluctuations moléculaires correspondent à la dynamique moléculaire des protéines (3,54).

La dynamique moléculaire de la femtoseconde à la nanoseconde caractérise des mouvements atomiques locaux (mouvements chaînes latérales), définis comme des mouvements rapides. Il existe aussi des mouvements moléculaires lents caractéristiques de fluctuations collectives de nombreux atomes dans des zones de plus grandes tailles, comme les éléments de structures secondaires (microsecondes à millisecondes), les structures tertiaires, les domaines, les chaînes (milliseconde à seconde) (Fig. 2-7) (3,53,54).

Pour suivre la dynamique des protéines, plusieurs approches expérimentales et théoriques existent. La cristallographie par rayon X et la cristallographie par laser à électrons

libres dans le domaine des rayons X (XFEL : X-ray free electron laser) sont déjà utilisées pour surveiller la résolution temporelle des mouvements conformationnels (99–102). L'ultra rapide RMN, la spectroscopie de Raman femtoseconde simulé et la spectroscopie infra-rouge ultra transitoire rapide mesurent des mouvements atomiques de la femtoseconde à la nanoseconde, soit des mouvements locaux (103–108). Des mouvements encore plus lents sont observés par des techniques de plus faible résolution spatiale comme la microscopie à force atomique (AFM) et la spectroscopie de fluorescence (FRET) (109–111).

Les approches théoriques telles que les simulations de dynamique moléculaire (MD) ont une résolution atomique et couvrent des mouvements jusqu'à la milliseconde grâce au super-ordinateur ANTON (54). Les simulations à haute résolution, sont limitées en terme de taille de protéine, de plus, il est difficile d'évaluer des mouvements au-dessus de la microseconde (112,113).

La plupart des approches pour couvrir les mouvements à l'échelle de temps de la dynamique de la structure protéique sont des approches intégratives, qui combinent des techniques expérimentales et théoriques (114–116). Ces techniques intégratives sont particulièrement efficaces pour surveiller les mouvements lents impliqués dans l'assemblage des protéines et la formation des pores (117–123).

La mesure de la dynamique spatio-temporelle des mouvements collectifs lents (microseconde à seconde) dans les protéines est cruciale car il s'agit de la plage temporelle où se produit le repliement des protéines et l'activité biologique. Malgré le fait que les mouvements lents des protéines puissent être étudiés via des approches intégratives, les mouvements lents sont globalement peu compris. En effet, les mouvements collectifs lents sont caractérisés par une hétérogénéité conformationnelle et/ou par des conformations peu peuplées et de courtes durées de vies. De plus, les mouvements lents des protéines impliquent des mouvements collectifs de nombreux atomes. Par conséquent, pour étudier les mouvements lents, les techniques utilisées doivent avoir une sensibilité élevée et pouvoir détecter les fluctuations reliées aux interactions moléculaires (61).

La spectroscopie diélectrique à large bande (BDS) est utilisée pour surveiller les fluctuations dipolaires dans la gamme de fréquence 1 Hz à  $10^6$  Hz (de la seconde à la microseconde) et le nanoconfinement est utilisé pour enrichir et/ou sélectionner des populations de conformations. En conclusion, le BDS devrait permettre de distinguer les conformations protéiques enrichies, sélectionnées par le nanoconfinement via leur dynamique moléculaire.

A notre connaissance, l'étude de la dynamique moléculaire des protéines par nanoconfinement et BDS n'a jamais été réalisée auparavant. Par conséquent, le but de ce chapitre est de déterminer si cette combinaison permet de distinguer des populations de conformation protéique (124). Pour cela, la perte diélectrique ( $\epsilon''$ ) est mesurée en fonction de la température à différentes fréquences, pour caractériser la conductivité et les processus de relaxation moléculaire (pics de relaxation) (Fig. 3-11). Le ou les processus de relaxation mesurés en fonction de la température à une fréquence donnée correspondent à la dynamique moléculaire.

## 4.1 Un signal diélectrique des protéines est-il détecté ?

### 4.1.1 Test nanomembrane vide

Avant de tester l'échantillon protéique pour savoir si la combinaison du nanoconfinement et de la spectroscopie diélectrique à large bande permet de détecter un signal diélectrique caractéristique de la conformation protéique, il est nécessaire de vérifier que les membranes avec les nanopores (nanomembranes) ne possèdent aucun signal similaire à la dynamique moléculaire (Fig. 4-1).

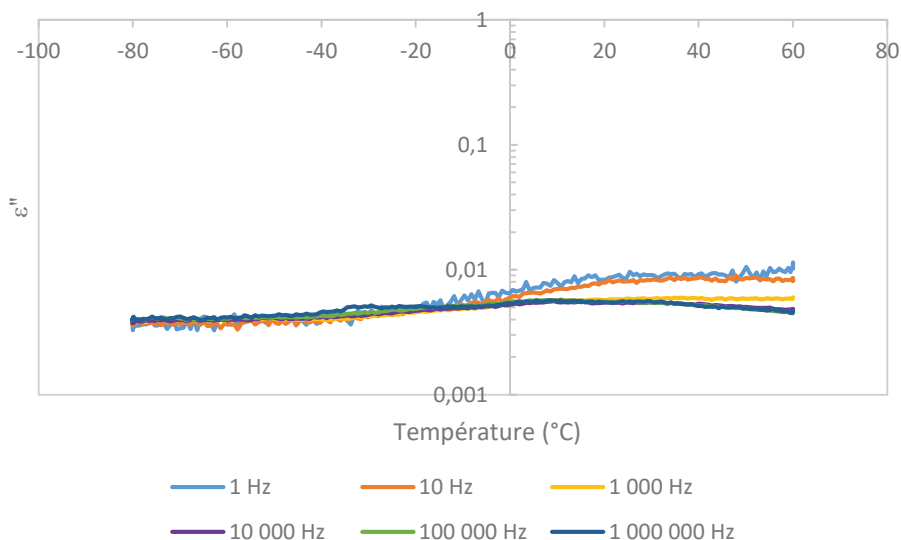


Figure 4-1. Nanomembrane vide avec un confinement à 40 nm de diamètre et 10  $\mu\text{m}$  de long. Représentation des courbes de refroidissement après un traitement thermique à 60  $^{\circ}\text{C}$ .

L'absence de signal de perte diélectrique pour la nanomembrane vide dans la gamme de fréquence étudiée, confirme que la membrane est utilisable pour étudier les protéines.

### 4.1.2 Test nanomembrane avec le solvant

Le solvant utilisé est constitué d'eau et de PBS (phosphate buffered saline) soit de NaCl, KCl,  $\text{KH}_2\text{PO}_4$ ,  $\text{Na}_2\text{HPO}_4$  et  $\text{H}_2\text{O}$ . Le solvant possède donc différents dipôles, suggérant qu'un pic de relaxation pourrait être détecté en regardant les données de pertes diélectriques à différentes fréquences et températures. Après avoir vérifié que les nanomembranes sont utilisables, il est donc nécessaire de caractériser le signal diélectrique du solvant. Pour réaliser ce contrôle, le solvant est testé dans les mêmes conditions que la protéine en termes de quantité d'eau et de PBS, pour la courbe de refroidissement de 60  $^{\circ}\text{C}$  à -80  $^{\circ}\text{C}$  après le premier traitement thermique à 60  $^{\circ}\text{C}$  (R1 60  $^{\circ}\text{C}$ ) (Fig. 4-2).

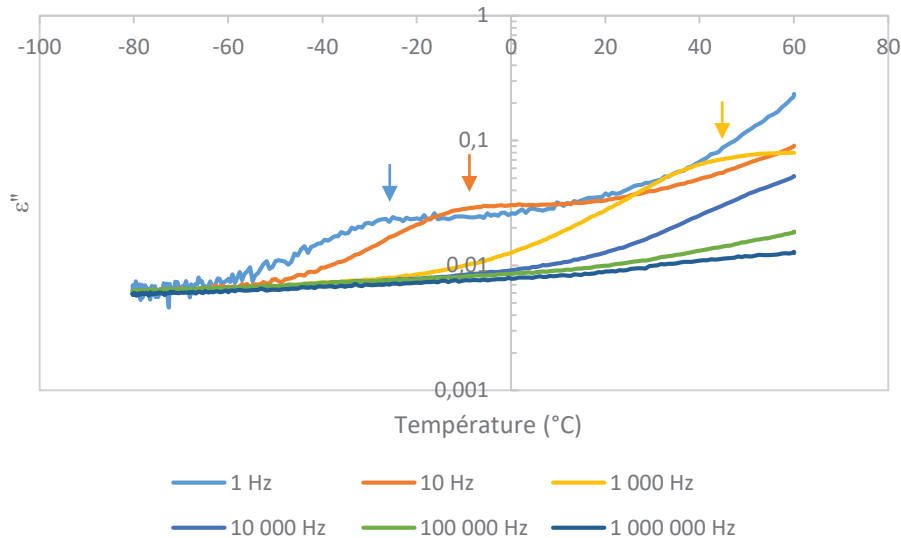
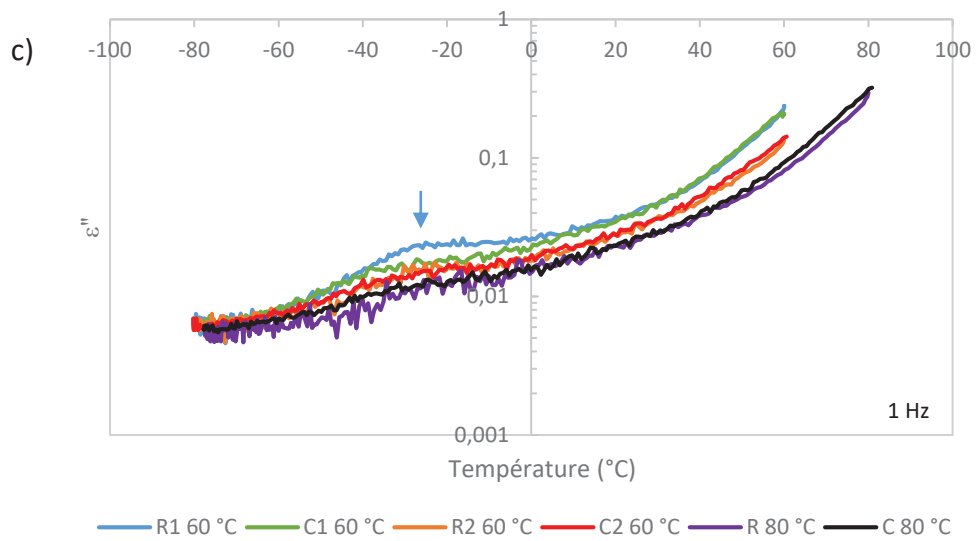
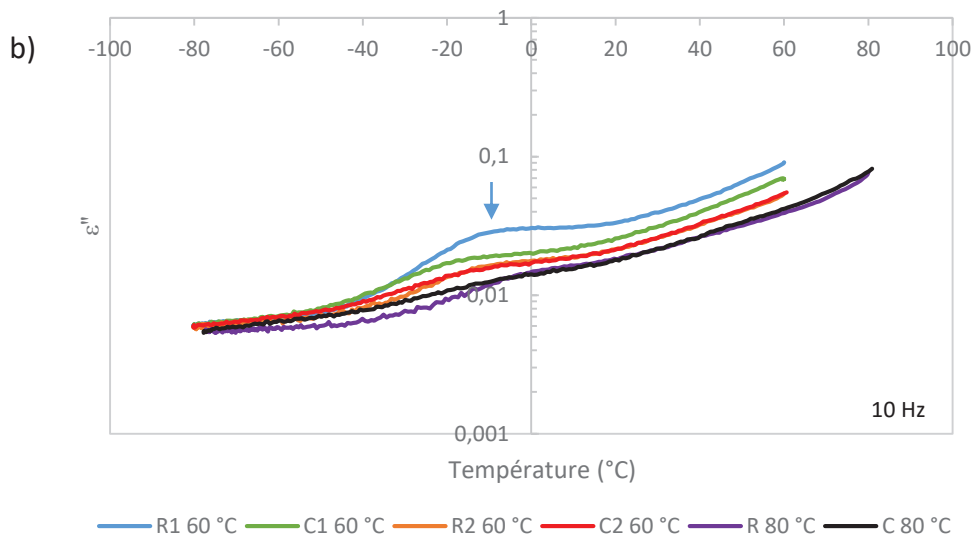
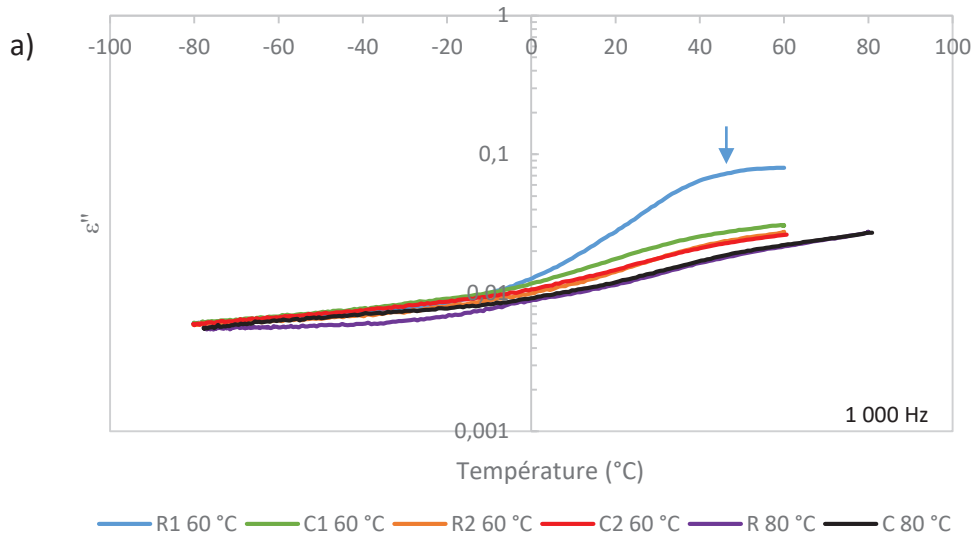


Figure 4-2. Superposition des courbes de refroidissement du solvant à différentes fréquences. La solution de solvant déposée sur la nanomembrane est composée de 2  $\mu\text{L}$  de PBS (PBS 1X) et de 204  $\mu\text{L}$  d'eau déionisée.

Comme le montre la figure 4-2, le solvant possède un processus de relaxation caractérisé par un pic de relaxation (symbole flèche) dans le spectre de perte diélectrique. À la fréquence de 1 Hz le pic apparaît à  $-27\text{ }^\circ\text{C}$ , à  $-8\text{ }^\circ\text{C}$  pour 10 Hz et à  $47\text{ }^\circ\text{C}$  pour  $10^3$  Hz. Toutefois, aucun processus de relaxation n'est détecté à partir de la fréquence de  $10^4$  Hz, car le pic est en dehors de la gamme de température utilisée. Une dynamique moléculaire associée au solvant est donc détectée après le premier traitement thermique à  $60\text{ }^\circ\text{C}$ .

Il est nécessaire d'analyser le comportement du solvant après différents traitements thermiques pour pouvoir ensuite étudier et analyser le comportement des protéines (Fig. 4-3). Dans un premier temps, le solvant est soumis à deux traitements thermiques à  $60\text{ }^\circ\text{C}$  pour vérifier que le signal diélectrique ne provient pas de l'eau, l'évaporation de l'eau conduit à des changements irréversibles du signal diélectrique qui sont facilement détectables par BDS. Puis le solvant est soumis à différents traitements thermiques à  $80\text{ }^\circ\text{C}$ ,  $100\text{ }^\circ\text{C}$ ,  $140\text{ }^\circ\text{C}$  et  $180\text{ }^\circ\text{C}$ . Suite au premier traitement thermique à  $60\text{ }^\circ\text{C}$ , une courbe de refroidissement est mesurée de  $60\text{ }^\circ\text{C}$  à  $-80\text{ }^\circ\text{C}$  (R1  $60\text{ }^\circ\text{C}$ ) puis une courbe de chauffage de  $-80\text{ }^\circ\text{C}$  à  $60\text{ }^\circ\text{C}$  (C1  $60\text{ }^\circ\text{C}$ ), et ainsi de suite pour les différents traitements thermiques réalisés.



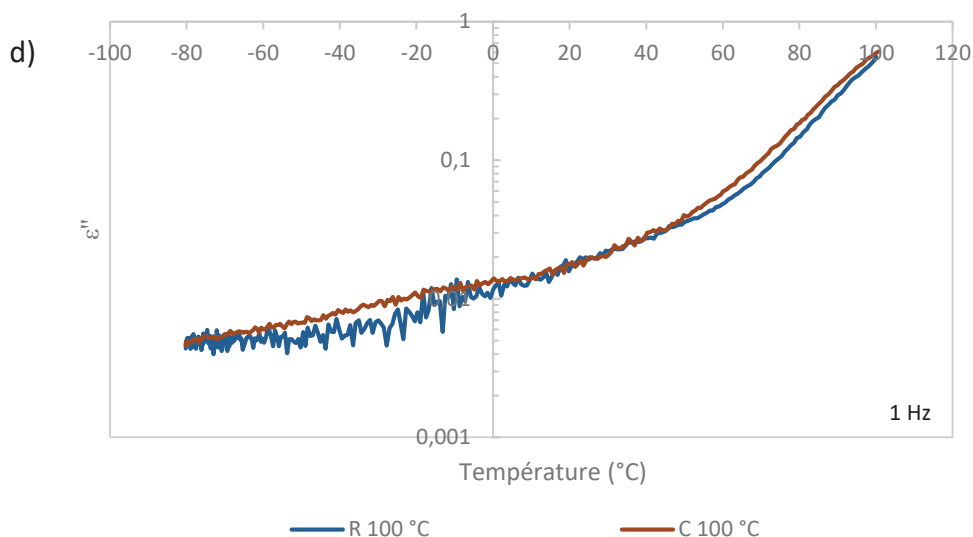


Figure 4-3. Superpositions des courbes de refroidissement et de chauffage mesurées à différentes fréquences après différents traitements thermiques pour le contrôle du solvant. a) Superposition des courbes à la fréquence de  $10^3$  Hz. b) Superposition des courbes à la fréquence de 10 Hz c) Superposition des courbes à fréquences de 1 Hz. d) Superposition des courbes à la fréquence de 1 Hz après le traitement thermique à 100  $^{\circ}\text{C}$ . Toutes les courbes de refroidissement et de chauffages sont obtenues après deux traitements thermiques à 60  $^{\circ}\text{C}$  et un traitement thermique à 80  $^{\circ}\text{C}$  sauf pour la figure d. 200  $\mu\text{L}$  de solution contenant 2  $\mu\text{L}$  de PBS 1X et 204  $\mu\text{L}$  d'eau déionisée est déposée sur la nanomembrane.

Pour les trois fréquences étudiées, la courbe de refroidissement après le premier traitement thermique à 60  $^{\circ}\text{C}$  possède un processus de relaxation caractéristique du solvant (MDb). Après le traitement thermique, l'intensité de ce pic de relaxation pour la courbe de refroidissement est supérieure à celle de la courbe de chauffage. Le solvant est composé de PBS et d'eau déionisée, par conséquent, la différence entre la courbe de refroidissement et celle de chauffage indique que l'évaporation de l'eau libre se produit toujours dans l'échantillon du solvant après trois heures à 60  $^{\circ}\text{C}$ . De plus, le pic de relaxation de la courbe de refroidissement est mieux défini que celui de la courbe de chauffage, suggérant que l'eau joue un rôle important dans le processus de relaxation et donc la dynamique moléculaire détectée pour le solvant (Fig. 4-3a à Fig. 4-3c).

Un deuxième traitement thermique à 60  $^{\circ}\text{C}$  est réalisé pour savoir si le solvant possède toujours un processus de relaxation. A la fréquence de 1 Hz, une MDb semble toujours détecté même si le signal diélectrique commence à être bruité. De même pour la fréquence à 10 Hz, une MDb semble toujours détectée vers -8  $^{\circ}\text{C}$ . Ainsi, la position de ce processus de relaxation en température ne change pas entre les deux traitements thermiques à 60  $^{\circ}\text{C}$ . A la fréquence de  $10^3$  Hz, aucun processus de relaxation et donc aucune dynamique moléculaire n'est détectée dans la fenêtre de notre expérience (Fig. 4-3a à Fig. 4-3c).

Pour les trois fréquences étudiées, l'intensité des pertes diélectriques mesurées pendant les courbes de refroidissement et de chauffage diminue par rapport à l'intensité des pertes diélectriques après le premier traitement thermique à 60  $^{\circ}\text{C}$ , ce qui tend à suggérer une évaporation plus importante de l'eau libre par rapport au premier traitement thermique. De

plus, après le deuxième traitement thermique à 60 °C, les intensités des courbes de refroidissement et des courbes de chauffage sont identiques. Cela indique, que l'eau libre s'est totalement évaporée pendant les trois heures à 60 °C du deuxième traitement thermique. La présence d'un pic de relaxation après l'évaporation de l'eau libre indique que le processus de relaxation du solvant provient de l'interaction entre l'eau liée et le PBS ou uniquement du PBS ou de l'eau liée.

Si un traitement thermique à 80 °C est réalisé à la suite des deux premiers traitements thermiques à 60 °C, aucun processus de relaxation du solvant n'est détecté pour les fréquences entre 10 Hz et 10<sup>3</sup> Hz, seulement la contribution diélectrique de la conductivité est détectée. A la fréquence de 1 Hz, le pic de relaxation du solvant ne semble plus être présent, même s'il est difficile de l'affirmer en raison du signal diélectrique bruité (Fig. 4-3a à Fig. 4-3c). Toutefois, si un traitement thermique à 100 °C est réalisé aucun MDb n'est détecté à la fréquence de 1 Hz (Fig. 4-3d).

Ces résultats tendent à suggérer que le processus de relaxation détecté pour le solvant dans la courbe de refroidissement après le premier traitement thermique à 60 °C, correspond principalement à l'eau libre. Toutefois, la reproductibilité de la 'légère' dynamique moléculaire lors du deuxième traitement thermique à 60 °C aux fréquences de 1 Hz et 10 Hz, suggère que l'eau liée, le PBS ou l'interaction eau liée-PBS participe aussi au processus de relaxation. Ces hypothèses sont émises en partant de l'observation que l'eau libre est totalement évaporée après 6h à 60 °C (3 h pour le premier traitement thermique à 60 °C et 3 h pour le deuxième), comme l'indique la stabilité de la réponse diélectrique mesurée après le deuxième traitement thermique à 60 °C. En conclusion, le processus de relaxation et donc la dynamique moléculaire associée au solvant n'est quasiment plus détectée après le deuxième traitement thermique à 60 °C. Maintenant, l'échantillon protéique est testé pour savoir si la combinaison du nanoconfinement et de la spectroscopie diélectrique à large bande est utilisable sur des protéines.

#### 4.1.3 Test nanomembrane avec CtxB<sub>5</sub>

##### 4.1.3.1 Etude de CtxB<sub>5</sub> à 1 000 Hz

Les courbes de refroidissement et de chauffage après le premier traitement thermique à 60 °C montrent un seul pic de relaxation dans le spectre de perte diélectrique caractéristique d'une dynamique moléculaire (symbole rond). Aucune différence significative n'est détectée dans les données de pertes diélectriques ( $\epsilon''$ ) pour les courbes de refroidissement et de chauffage, caractéristique d'une absence d'hystérésis (Fig. 4-4). La protéine reste donc dans un état stable pendant toute la mesure et subit un dépliement thermique principalement pendant les trois heures d'incubation. De plus, les intensités diélectriques des courbes de refroidissement et de chauffage sont similaires, l'intensité de la courbe de chauffage est même légèrement supérieure à celle de la courbe de refroidissement, ce qui laisse suggérer que, contrairement au solvant, l'eau libre est totalement évaporée pendant les trois heures à 60 °C. L'échantillon protéique est donc constitué de : protéine, eau liée et PBS.

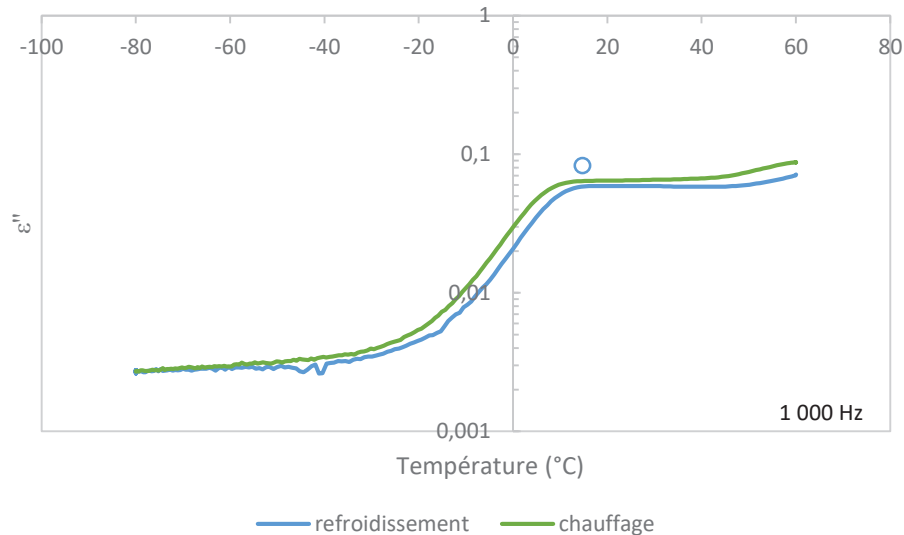


Figure 4-4. Superposition des courbes de refroidissement et de chauffage mesurées à la fréquence de  $10^3$  Hz après le premier traitement thermique à 60 °C. La solution utilisée est composée de 2  $\mu$ L de CtxB<sub>5</sub> à une concentration de 2,6 mg/ml et de 204  $\mu$ L d'eau déionisée, soit une concentration de 0,025 mg/ml.

Si un deuxième traitement thermique à 60 °C est réalisé, les données de pertes diélectriques sont reproductibles à la fois pour les courbes de refroidissement et de chauffage pour l'échantillon protéique (Fig. 4-5). Ceci démontre que les protéines étudiées sont stabilisées dans un état où il n'y a plus d'eau libre et où la quantité d'eau liée est également stabilisée, indiquant que trois heures d'incubation à 60 °C suffisent pour que la protéine atteigne un état conformationnel stable. Par conséquent, ce temps d'incubation peut être utilisé pour les traitements thermiques à des températures plus élevées.



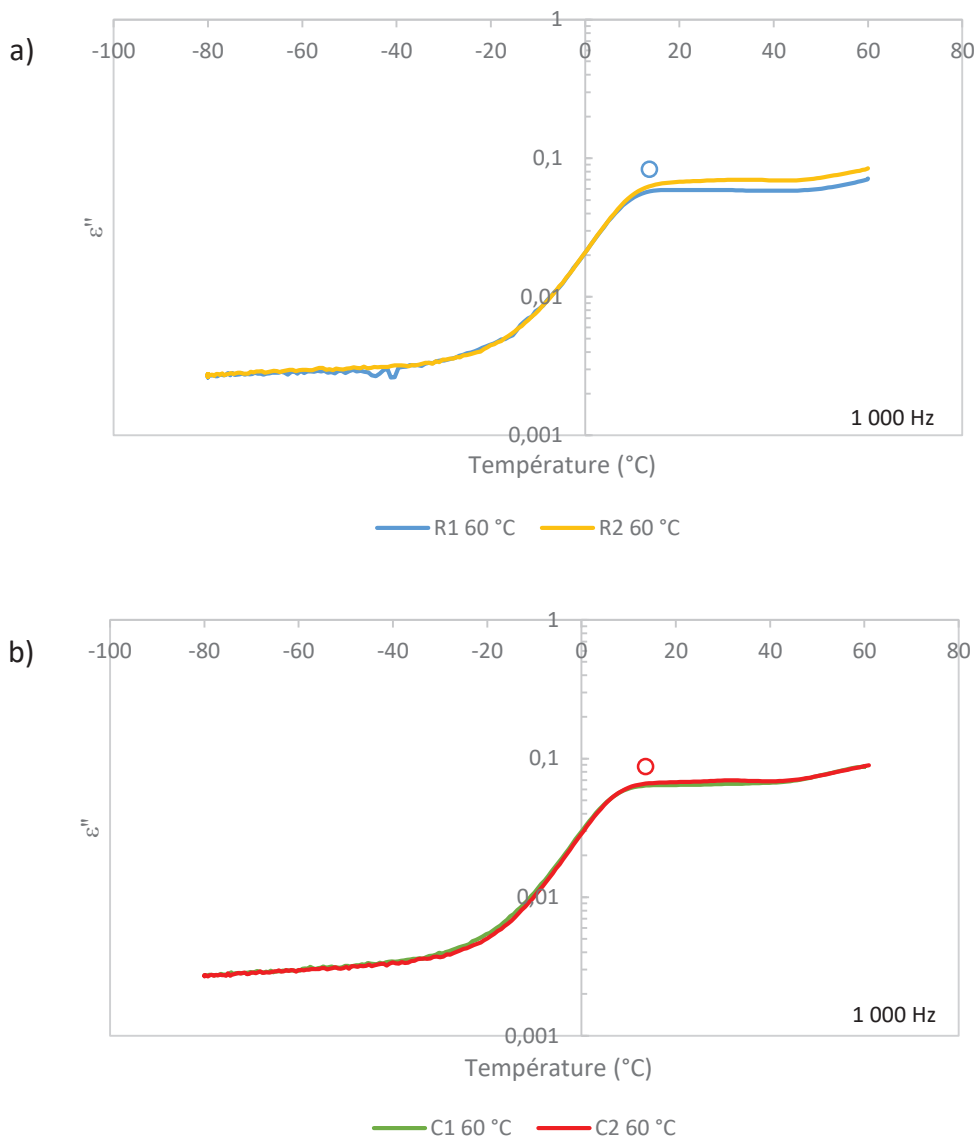


Figure 4-5. Superposition des courbes de refroidissement et de chauffage mesurées à la fréquence de  $10^3$  Hz après des traitements thermiques à 60 °C. a) Superposition des courbes de refroidissement du premier et du deuxième traitement thermique à 60 °C. b) Superposition des courbes de chauffage du premier et du deuxième traitement thermique à 60 °C.

Si une deuxième expérience est réalisée dans les mêmes conditions pour l'échantillon protéique, la position du processus de relaxation en température (température où l'intensité du pic de relaxation est maximale) est identique (Fig. 4-6). Ainsi, l'expérience est reproductible pour un même échantillon et un même confinement.

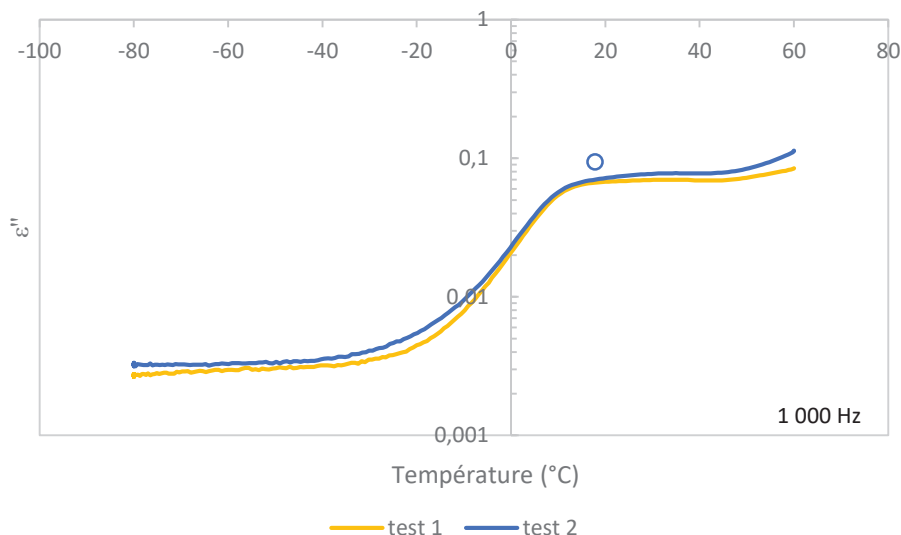


Figure 4-6. Reproduction de l'expérience avec un échantillon protéique à la fréquence de  $10^3$  Hz. Superposition des courbes de refroidissement après le deuxième traitement thermique à 60 °C.

La position en température du processus de relaxation du solvant et de la protéine ne se situe pas à la même température (47 °C pour le solvant et 13 °C pour la protéine) pour les courbes de refroidissement après le premier traitement thermique à 60 °C à la fréquence de  $10^3$  Hz. Toutefois, les intensités de pertes diélectriques sont du même ordre de grandeur (Fig. 4-7a). Ce qui n'est pas le cas pour le deuxième traitement thermique à 60 °C, où le processus de relaxation du solvant est inexistant (Fig. 4-7b). Pour ces raisons, additionnés avec le fait que les données de pertes diélectriques de l'échantillon protéique sont superposables après les deux traitements thermiques à 60 °C (Fig. 4-5), les données seront principalement analysées après le deuxième traitement thermique à 60 °C pour la suite du manuscrit.

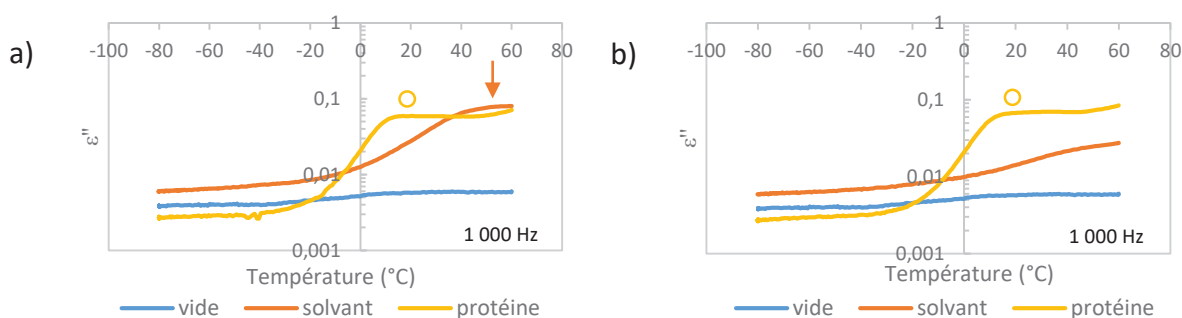


Figure 4-7. CtxB<sub>5</sub> dans le confinement à 40 nm de diamètre et 10 μm de long avec superposition des courbes de refroidissement à la fréquence de  $10^3$  Hz de la protéine, du solvant et de la nanomembrane vide. a) Superposition des courbes de refroidissement après le premier traitement thermique à 60 °C b) Superposition des courbes de refroidissement après le deuxième traitement thermique à 60 °C.

En conclusion :

- A la fréquence de  $10^3$  Hz et après le premier traitement thermique à 60 °C, le solvant et la protéine sont tous les deux caractérisés par un processus de relaxation.
- Le processus de relaxation est identique pour la courbe de refroidissement et de chauffage après le traitement thermique à 60 °C et reproductible si un deuxième

traitement thermique à 60 °C est réalisé. Par conséquent, la protéine est dans une conformation protéique stable.

- Le processus de relaxation n'est pas identique pour la courbe de refroidissement et de chauffage après le premier traitement thermique à 60 °C. De plus, le processus de relaxation n'est pas reproductible si un deuxième traitement thermique à 60 °C est réalisé pour le solvant à la fréquence de  $10^3$  Hz.

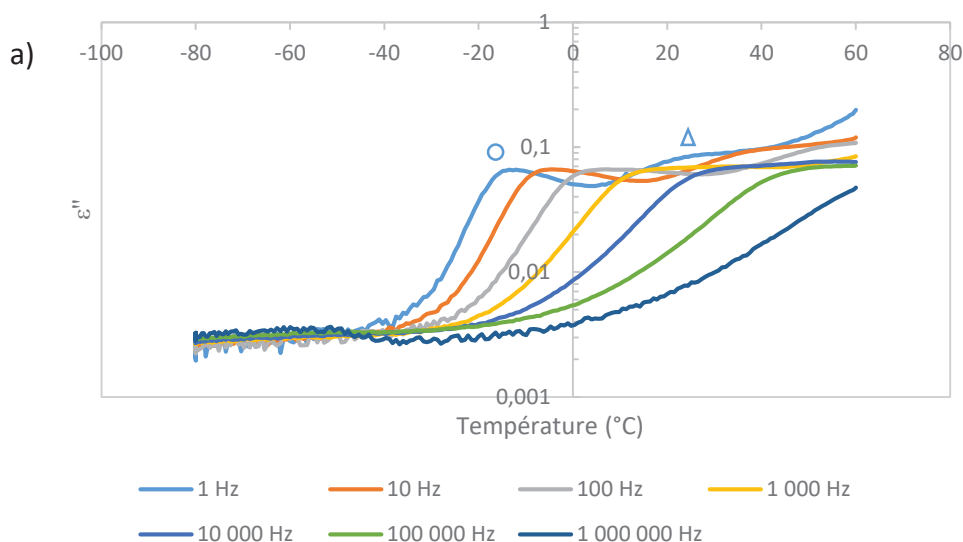
Par conséquent, la combinaison du nanoconfinement et de la spectroscopie diélectrique à large bande permet la détection d'un signal diélectrique caractéristique de la dynamique moléculaire de protéine.

Pour caractériser la dynamique moléculaire de la protéine, le signal de la perte diélectrique est mesuré en fonction de la température et de la fréquence. Après avoir validé la mesure diélectrique de la protéine à une fréquence donnée suite aux traitements thermiques à 60 °C, différentes fréquences seront systématiquement étudiées.

## 4.2 Analyse de CtxB<sub>5</sub> à différentes fréquences

Deux processus de relaxation sont visibles pour CtxB<sub>5</sub> aux fréquences de 1 Hz et 10 Hz. Le deuxième pic de relaxation (symbole triangle) est aussi présent à la fréquence de  $10^2$  Hz, même si celui-ci n'est pas assez bien défini pour pouvoir déterminer la position de ce pic en température. Un seul processus de relaxation est détecté aux fréquences de  $10^3$  Hz,  $10^4$  Hz et  $10^5$  Hz, pour aucun processus de relaxation à la fréquence de  $10^6$  Hz (Fig. 4-8). Au vu de ces résultats les fréquences à 1 Hz,  $10^3$  Hz et  $10^6$  Hz, seront analysées plus en détail.

Le premier processus de relaxation détecté dans les basses et hautes fréquences (1 Hz à  $10^5$  Hz) se nommera MD<sub>1</sub> (symbole rond). Le deuxième processus de relaxation plus lent (symbole triangle) possédant un pic plus large, moins distinct, se nommera MD<sub>2</sub>.



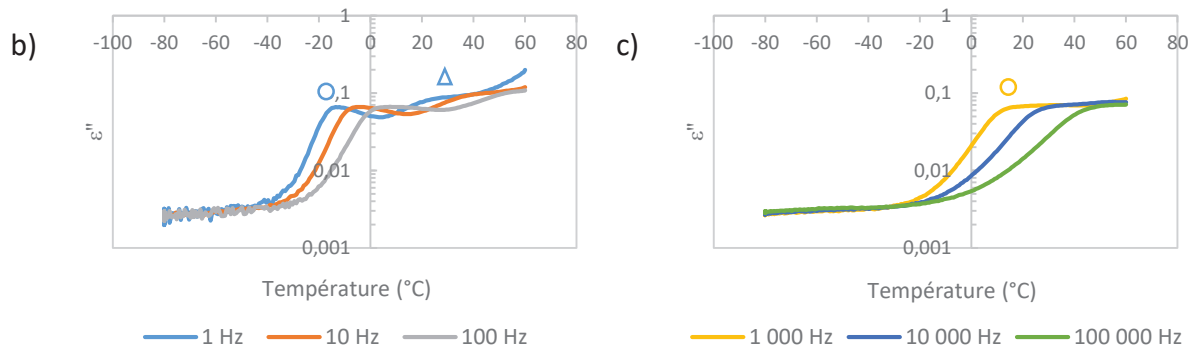


Figure 4-8. Analyse de CtxB<sub>5</sub> à différentes fréquences après le traitement thermique à 60 °C. a) Superpositions des courbes de refroidissement après le deuxième traitement thermique à 60 °C. c) Zoom sur les processus de relaxation de la fréquence 1 Hz à 10<sup>2</sup> Hz. b) Zoom sur les processus de relaxation de la fréquence 10<sup>3</sup> Hz à 10<sup>5</sup> Hz. Le premier processus de relaxation est représenté par le symbole rond, le deuxième processus de relaxation est représenté par le symbole triangle.

La température du pic de relaxation MD<sub>1</sub> (symbole rond) augmente avec l'augmentation de la fréquence (Fig. 4-8). A la fréquence de 1 Hz, la position du pic de relaxation en température se situe à -14 °C et à 13 °C à la fréquence de 10<sup>3</sup> Hz.

Le premier processus de relaxation (MD<sub>1</sub>) possède un décalage de 27 °C entre les fréquences de 1 Hz à 10<sup>3</sup> Hz (-14 °C à 13 °C), alors que le deuxième processus de relaxation (MD<sub>2</sub>) ne possède pas le même décalage pour la même gamme de fréquence. En effet, MD<sub>2</sub> n'est pas détecté à la fréquence de 10<sup>3</sup> Hz dans notre gamme de mesure. Par conséquent, les deux processus de relaxation détectés dès le traitement thermique à 60 °C jusqu'à 10<sup>2</sup> Hz ne possèdent pas les mêmes environnements dipolaires. Deux hypothèses pourraient expliquer la présence de deux pics de relaxation :

- Le confinement, ne permet pas d'enrichir et/ou sélectionner une seule population conformationnelle et les différentes populations conformationnelles enrichies, sélectionnées sont caractérisées par différents pics de relaxation.
- Une population conformationnelle est enrichie, sélectionnée (conformation moyenne sur la statistique de population de tous les nanopores) mais différentes zones de dipôles sont détectées, probablement à cause de la dénaturation thermique de la protéine dans une autre population conformationnelle.

Les deux pics de relaxation ne sont pas détectés dans la même gamme de fréquence. Or l'échantillon protéique utilisé est un pentamère, par conséquent, même si à cause de l'hétérogénéité conformationnelle les dipôles présents dans les conformations protéiques de l'échantillon ne sont pas les mêmes, la même conformation protéique 'globale' est présente. Par conséquent, toutes les conformations de pentamères devraient être détectées dans une gamme de fréquence similaire. Comme ce n'est pas le cas, ces résultats suggèrent qu'une population de conformation est enrichie, sélectionnée (conformation moyenne) et que les différents processus de relaxation détectés proviennent de la détection de différentes zones.

Dans ce cas, la largeur du pic de MD<sub>2</sub> pourrait suggérer un processus de relaxation associé à une population de conformations hétérogènes de la toxine, tandis que le pic plus

étroit de MD<sub>1</sub> suggérait des conformations de toxines moins hétérogènes (Fig. 4-8b). La présence d'un pic principal 'étroit' (MD<sub>1</sub>) observé après le traitement thermique à 60 °C pour les fréquences de 1 Hz à 10<sup>5</sup> Hz suggère que les nanopores fonctionnent comme un entonnoir conformationnel (Fig. 2-10) et réduisent l'hétérogénéité structurale de la solution en contraignant l'espace 3D dans lequel la toxine se trouve. D'autres diamètres et longueurs de nanopores seront testés et présentés dans les chapitres 5 et 6 pour valider cette hypothèse.

La dynamique moléculaire de la protéine est maintenant caractérisée en fonction de la température, pour cela la toxine est soumise à une dénaturation thermique à différents traitements thermiques.

### 4.3 Analyse du comportement de CtxB<sub>5</sub> en fonction de la température

La dénaturation thermique de la toxine est étudiée à l'aide de rampes de températures avec une série de traitements thermiques de trois heures à 60 °C, 80 °C, 100 °C, 140 °C et 180 °C suivis de cycle de refroidissement/ chauffage après chaque traitement pour mesurer la réponse diélectrique (Fig. 3-16).

Les protéines globulaires vont de structures compacts avec de nombreuses interactions atomiques et moléculaires qui maintiennent l'état natif à basses températures, à des structures dépliées moins compactes à des températures élevées, où les conformations perdent des interactions atomiques et moléculaires. La stabilité thermique varie considérablement pour les protéines des organismes mésophiles aux organismes thermophiles (125,126). Pour CtxB<sub>5</sub>, la stabilité thermique a été étudiée par calorimétrie différentielle à balayage et présente une transition à 74 °C (127,128). Les traitements thermiques perturbent les interactions atomiques et moléculaires natives de la toxine et devraient donc modifier les dipôles et/ou les environnements dipolaires, et par conséquent modifier les processus de relaxation et leurs caractéristiques.

La protéine reste à chaque fois trois heures à 80 °C, 100 °C, 140 °C, afin de mesurer les caractéristiques diélectriques de la protéine dans un état stable comme démontré par le traitement thermique à 60 °C (Fig. 4-4). L'état stable est révélé par l'absence d'hystérésis après les différents traitements thermiques, caractéristique d'un dépliement thermique pendant les trois heures d'incubations (Fig. 4-9). La stabilité de la protéine après le traitement thermique à 180 °C n'est pas étudiée. En effet, après la courbe de refroidissement de 180 °C à -80 °C, la température ne retourne pas à 180 °C (Fig. 3-16).

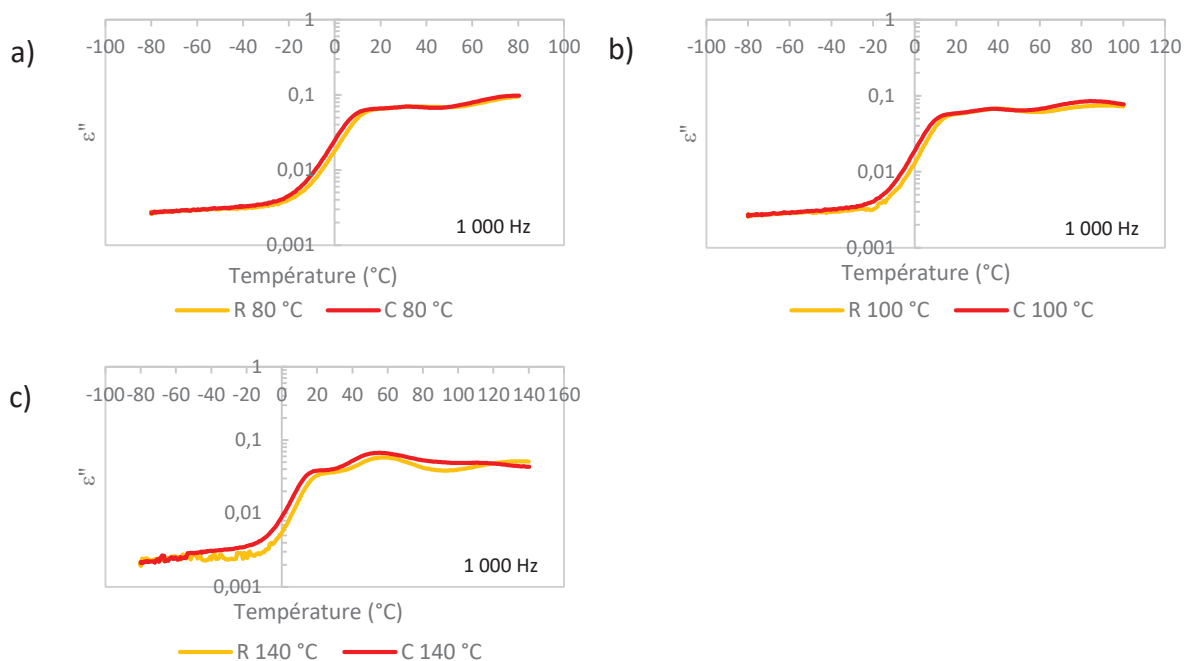
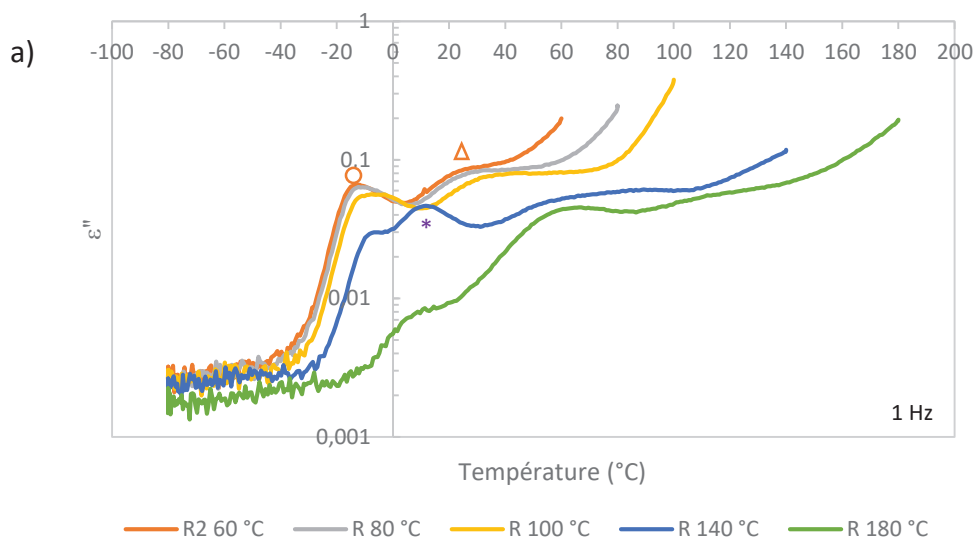


Figure 4-9. Étude de CtxB<sub>5</sub> à la fréquence de 10<sup>3</sup> Hz après différents traitements thermiques. a) Données des pertes diélectriques après le traitement thermique à 80 °C. b) Après le traitement thermique à 100 °C. c) Après celui à 140 °C.

Pour rappel, à la fréquence de 1 Hz, un seul processus de relaxation du solvant est observé jusqu'au traitement thermique à 80 °C (Fig. 4-3c). A la fréquence de 10<sup>3</sup> Hz, le processus de relaxation du solvant n'est plus observé à partir du deuxième traitement thermique à 60 °C (Fig. 4-3a). Alors qu'à la fréquence de 10<sup>6</sup> Hz, aucun processus de relaxation du solvant n'est observé dès le premier traitement thermique à 60 °C (Fig. 4-2). Ce qui n'est pas le cas pour la toxine, où la protéine maintient de multiples processus de relaxation à travers les différents traitements thermiques jusqu'à 180 °C et sur les fréquences de 1 Hz à 10<sup>6</sup> Hz (Fig. 4-10). Cela confirme que les processus de relaxation détectés pour l'échantillon protéique sont spécifiques des conformations protéiques.



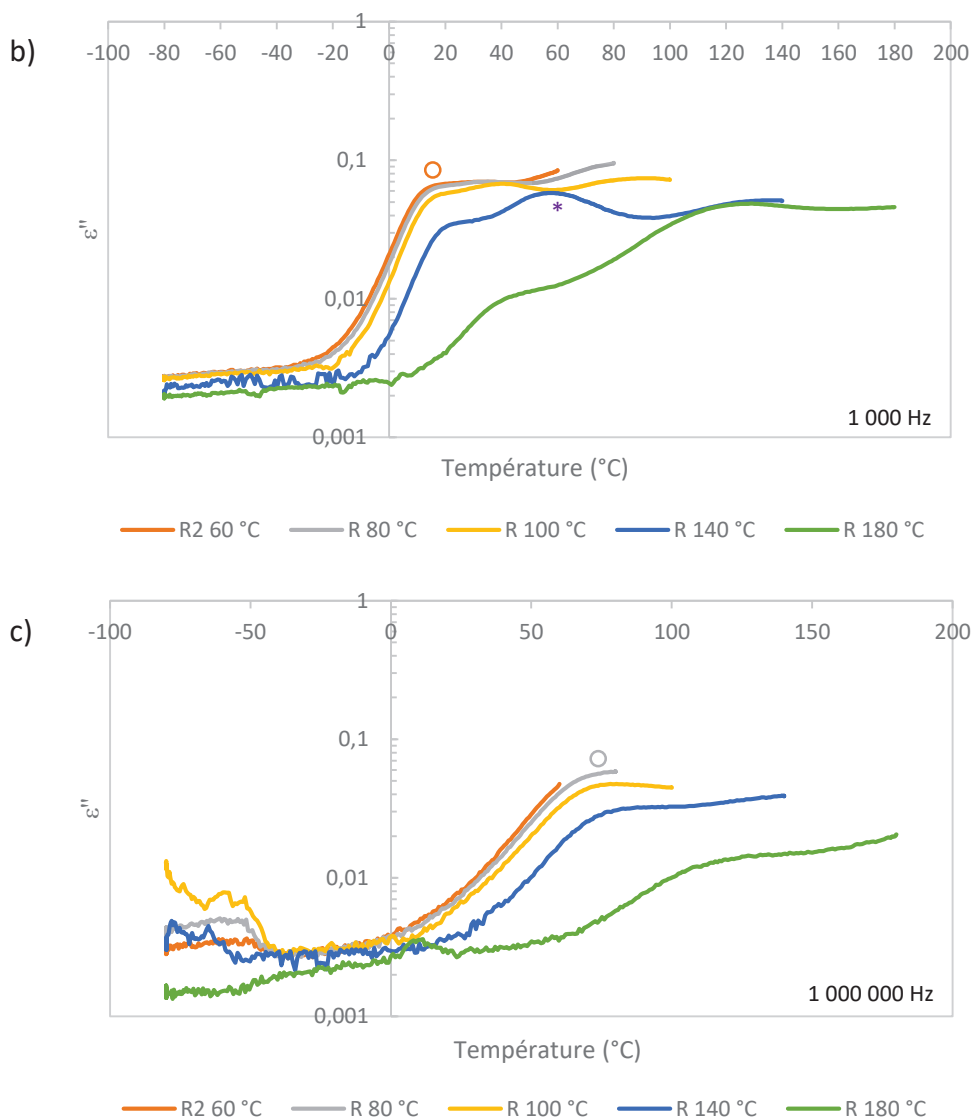


Figure 4-10. Etude du comportement de CtxB<sub>5</sub> après différents traitements thermiques. a) Analyse de CtxB<sub>5</sub> à la fréquence de 1 Hz. b) Analyse de CtxB<sub>5</sub> à la fréquence de 10<sup>3</sup> Hz c) Analyse de CtxB<sub>5</sub> à la fréquence de 10<sup>6</sup> Hz. MD<sub>1</sub> est représenté par le symbole rond, MD<sub>2</sub> par le symbole triangle et MD<sub>3</sub> par le symbole étoile.

A la fréquence de 1 Hz (Fig. 4-10a), les processus de relaxation sont similaires après le traitement thermique à 60 °C jusqu'au traitement thermique à 100 °C, suggérant que la dynamique moléculaire et les conformations protéiques détectées sont stables jusqu'à 100 °C. Les positions de MD<sub>1</sub> (symbole rond) en température changent de -14 °C (après les traitements thermiques à 60 °C et 80 °C) à -7 °C (après le traitement thermique à 100 °C) et de 30 °C (après les traitements thermiques à 60 °C et 80 °C) à 40 °C (après celui à 100 °C) pour MD<sub>2</sub> (symbole triangle). La variation de la position des pics de relaxation MD<sub>1</sub> et MD<sub>2</sub> entre les traitements thermiques à 80 °C et 100 °C traduit un changement d'environnement local pour ces deux processus de relaxation, probablement à cause d'une dénaturation thermique. Après incubation à 140 °C et 180 °C, l'intensité du pic de relaxation MD<sub>1</sub> diminue, avec en plus un élargissement du pic, et un décalage des positions en température de -7 °C à 10 °C après le traitement thermique à 180 °C. La baisse de l'intensité de perte diélectrique MD<sub>1</sub> suggère une

disparition progressive des conformations associées à ce processus de relaxation lorsque la température dépasse 100 °C. Les décalages de la position du pic de relaxation en température vers des températures plus élevées après les traitements thermiques à 100 °C et 180 °C suggèrent qu'à ces températures la protéine se déplie et donc change de conformation. Le dépliement thermique se traduit par la déstabilisation de plus de dipôles d'où une position en température plus élevée pour MD<sub>1</sub> indiquant un mouvement collectif plus grand. Le pic de MD<sub>2</sub> s'élargit et la position de ce pic en température passe de 40 °C à 80 °C après le traitement thermique à 140 °C. Comme pour MD<sub>1</sub>, l'intensité de la perte diélectrique MD<sub>2</sub> diminue lorsque la protéine est chauffée jusqu'à 140 °C, toutefois MD<sub>2</sub> est indétectable après le traitement à 180 °C. Cela suggère que, comme pour MD<sub>1</sub>, les conformations de la toxine associées à MD<sub>2</sub> disparaissent lorsque la température augmente au-dessus de 140 °C. Les décalages des pics de relaxation MD<sub>2</sub> après les traitements thermiques à 100 °C et 140 °C indiquent que plus de dipôles sont impliqués dans le processus de relaxation conduisant à une déstabilisation plus importante de la toxine dans cette conformation après ces deux traitements thermiques.

Après le traitement thermique à 140 °C, simultanément à la chute du signal MD<sub>1</sub>, un nouveau processus de relaxation (MD<sub>3</sub> : symbole étoile) apparaît avec un pic étroit et une position en température à 10 °C. Le pic MD<sub>3</sub> s'élargit après le traitement thermique à 180 °C et la position du pic en température passe à 70 °C, caractéristique d'une augmentation du nombre de dipôles détectés.

A la fréquence de 10<sup>3</sup> Hz (Fig. 4-10b), MD<sub>2</sub> est à peine détecté à partir du traitement thermique à 100 °C tandis que les caractéristiques de MD<sub>1</sub> dépendant de la température changent de façon similaire à la fréquence de 1 Hz. De plus, MD<sub>3</sub> commence à être détectée dès le traitement thermique à 80 °C.

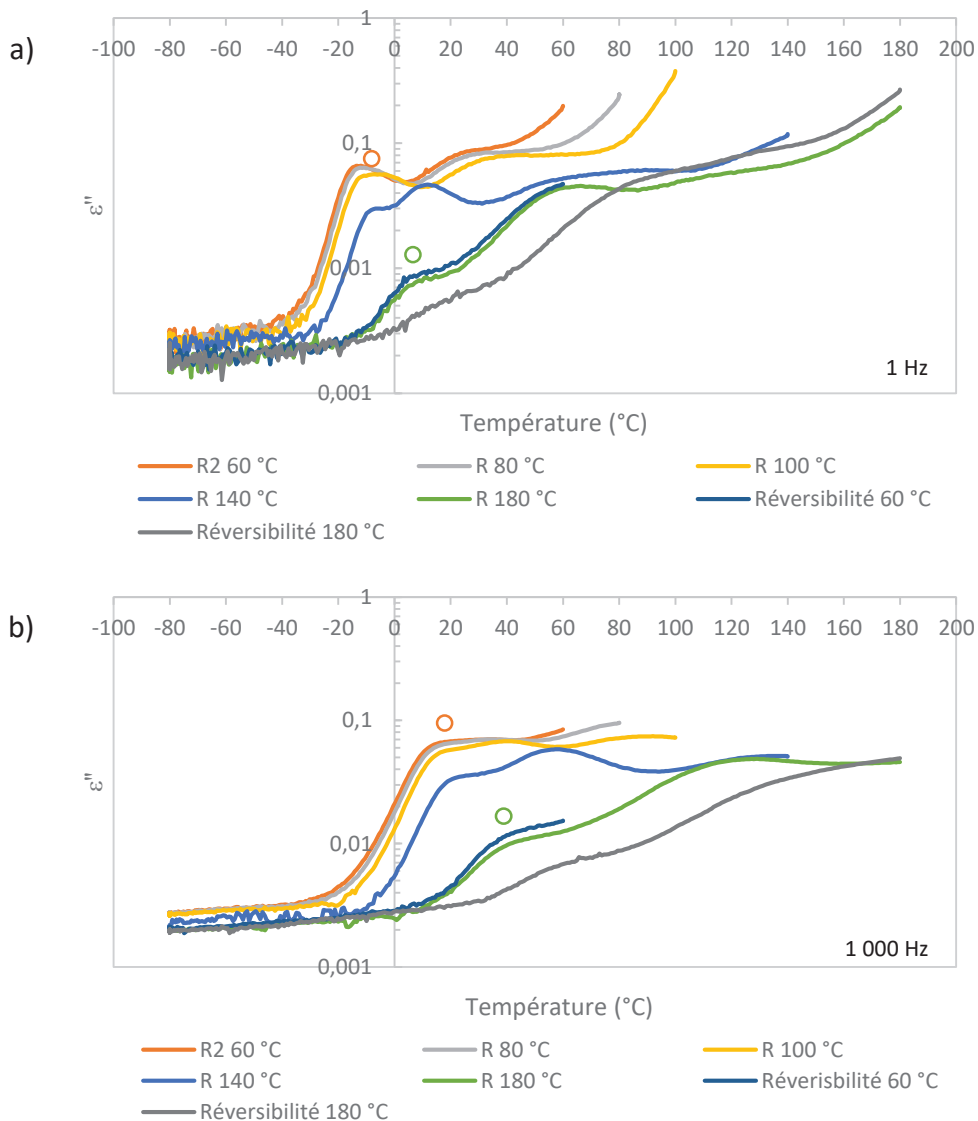
A la fréquence de 10<sup>6</sup> Hz (Fig. 4-10c), MD<sub>1</sub> commence à être détectable à partir du traitement thermique à 80 °C. Un changement dans la position du pic de relaxation en température puis dans l'intensité est observé comme pour les fréquences plus basses à partir du traitement thermique à 100 °C.

Le signal diélectrique est plus bruyant à des températures inférieures à -20 °C pour les différentes fréquences étudiées, mais à 10<sup>6</sup> Hz, le signal devient plus bruyant sur une plus grande plage de température.

En conclusion, le nanoconfinement et la spectroscopie diélectrique à large bande détectent trois processus de relaxation (MD<sub>1</sub>, MD<sub>2</sub> et MD<sub>3</sub>) caractéristiques de la conformation protéique de la toxine. Le processus de relaxation MD<sub>2</sub> ne couvre pas les hautes fréquences (supérieure à 10<sup>3</sup> Hz) contrairement aux deux autres processus de relaxation (MD<sub>1</sub> et MD<sub>3</sub>). De plus, MD<sub>2</sub> possède une position en température plus élevée que celle de MD<sub>1</sub> pour les différents traitements thermiques. Ces résultats suggèrent que la population conformationnelle associée à MD<sub>2</sub> possède un état de dépliement plus avancé que les deux autres processus de relaxation détectés.



Pour évaluer si la conformation protéique initiale détectée associée à MD<sub>1</sub> (symbole rond) peut être récupérée après le traitement thermique à 180 °C, l'échantillon est chauffé à 60 °C et maintenu pendant trois heures (Fig. 4-11).



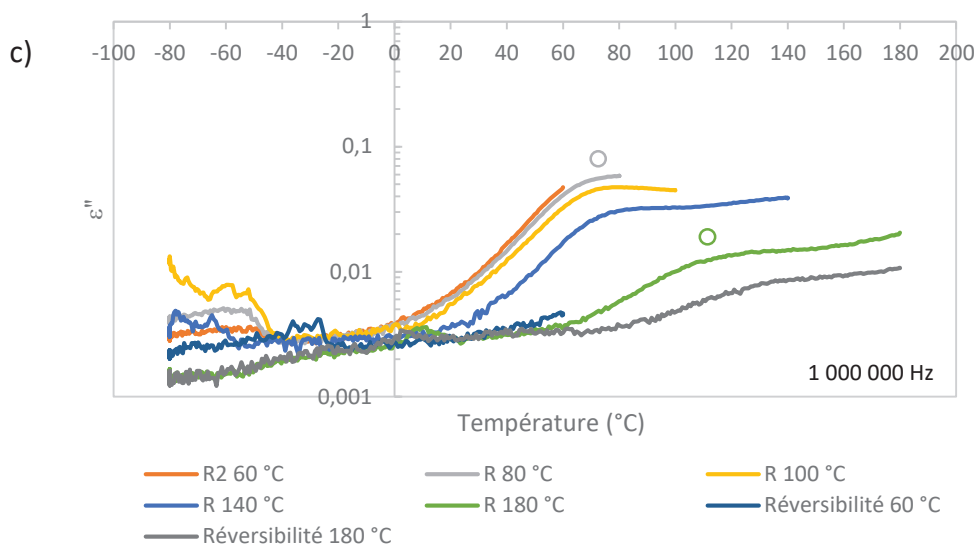


Figure 4-11. Test de réversibilité pour CtxB<sub>5</sub>. a) Etude de la réversibilité pour les courbes de refroidissement à la fréquence de 1 Hz b) Etude de la réversibilité pour les courbes de refroidissement à la fréquence de 10<sup>3</sup> Hz c) Et à 10<sup>6</sup> Hz

La position et les intensités du pic de relaxation MD<sub>1</sub> après le traitement thermique à 180 °C et après un 'nouveau' traitement thermique à 60 °C sont similaires (courbes R 180 °C et réversibilité 60 °C de la figure 4-11), indiquant que les dommages thermiques sont irréversibles et que la protéine ne se replie pas dans la conformation initiale malgré les trois heures à 60 °C. Si l'échantillon est réchauffé à 180 °C et conservé pendant trois heures supplémentaires, le signal diélectrique apparaît plus bas qu'après le premier traitement thermique à 180 °C et MD<sub>1</sub> devient presque indétectable aux fréquences étudiées indiquant que les conformations associées à ce processus de relaxation ont presque disparu.

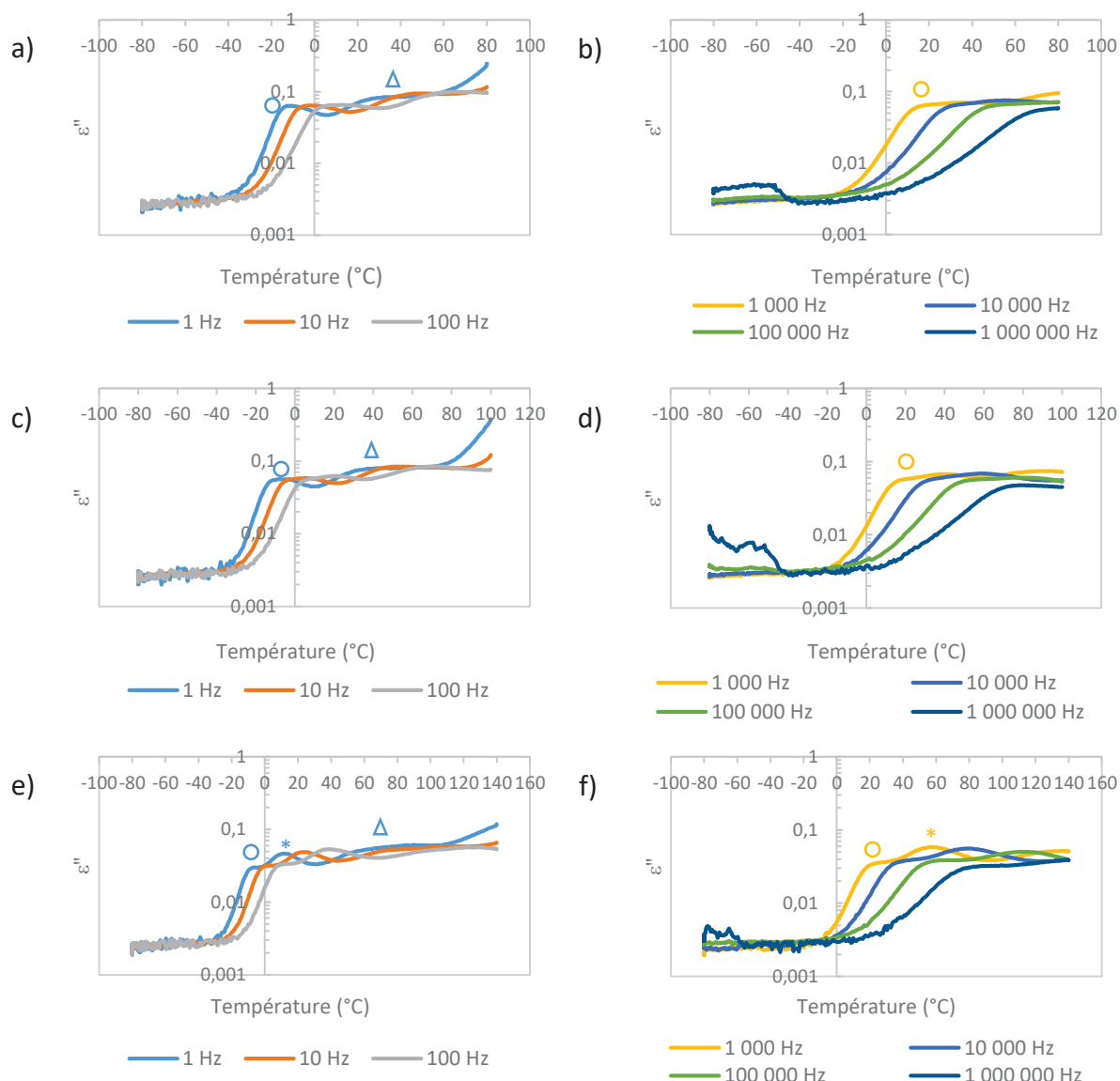
#### 4.4 Analyse du comportement de CtxB<sub>5</sub> en fonction de la fréquence

La dépendance en fréquence des trois processus de relaxation est étudiée en fonction des différents traitements thermiques dans la figure 4-12 et après le traitement thermique à 60 °C dans la figure 4-8. MD<sub>1</sub> est détecté dans une large gamme de fréquence allant de 1 Hz à 10<sup>5</sup> Hz après le traitement thermique à 60 °C et de 1 Hz à 10<sup>6</sup> Hz pour les traitements thermiques de 80 °C à 180 °C. MD<sub>2</sub> n'est détecté qu'aux basses fréquences (de 1 Hz à 10<sup>2</sup> Hz) pour les traitements thermiques de 60 °C à 140 °C. A 10<sup>3</sup> Hz, MD<sub>2</sub> est à peine détecté à partir du traitement thermique à 100 °C et jusqu'au traitement thermique à 140 °C. MD<sub>3</sub> couvre les fréquences de 10<sup>2</sup> Hz à 10<sup>4</sup> Hz après le traitement thermique à 80 °C, de 10<sup>2</sup> Hz à 10<sup>5</sup> Hz après le traitement à 100 °C, les basses et hautes fréquences (1 Hz à 10<sup>5</sup> Hz) après le traitement thermique à 140 °C et majoritairement les basses fréquences après le traitement thermique à 180 °C (1 Hz à 10<sup>4</sup> Hz).

La figure 2-7 a résumé que la protéine possédait des mouvements de chaînes latérales à l'échelle de la nanoseconde, des mouvements plus lents de plus grandes échelles (structures secondaires) et des mouvements encore plus lents sur des plus grandes surfaces (repliement,

déplie ment des domaines, interfaces et des structures tertiaires/ assemblage et dissociation de chaînes polypeptidiques). Le premier processus de relaxation (symbole rond) implique plusieurs échelles de mouvements collectifs associées à des mouvements de structures secondaires, domaines, interfaces et structures tertiaires (chaînes polypeptidiques). Le deuxième processus de relaxation détecté (symbole triangle) dans les basses fréquences est lié à des mouvements de domaines, interfaces, structures tertiaires, alors que le troisième processus de relaxation (symbole étoile) correspond à des mouvements de structures secondaires, domaines et structures tertiaires.

Les différents processus de relaxation suggèrent la détection de populations conformationnelles avec différents états de déplie ment en combinant le nanoconfinement et la spectroscopie diélectrique à large bande.



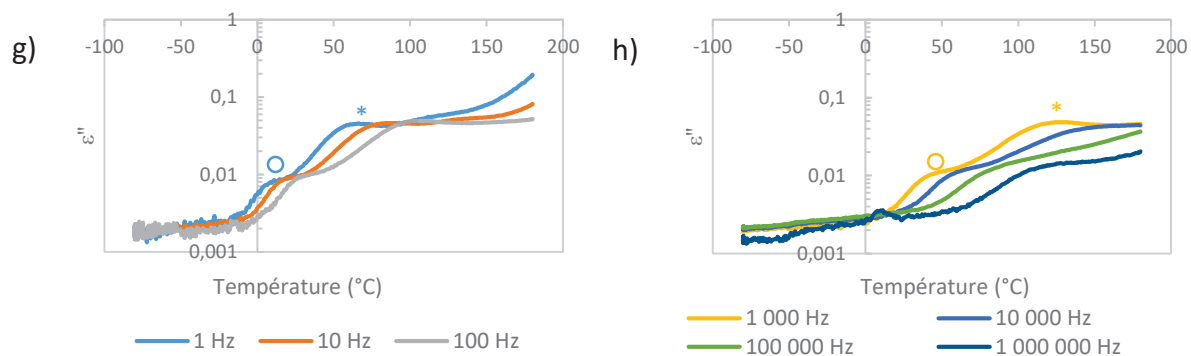
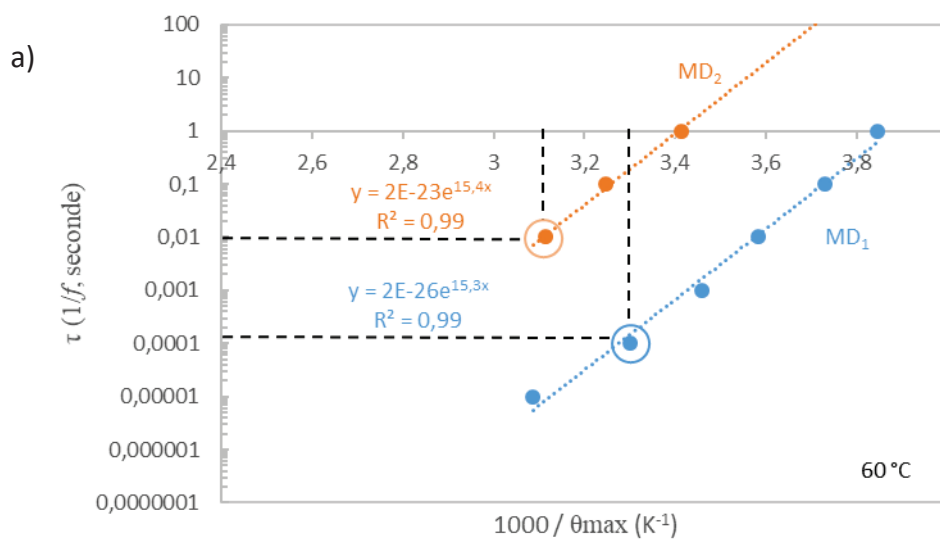
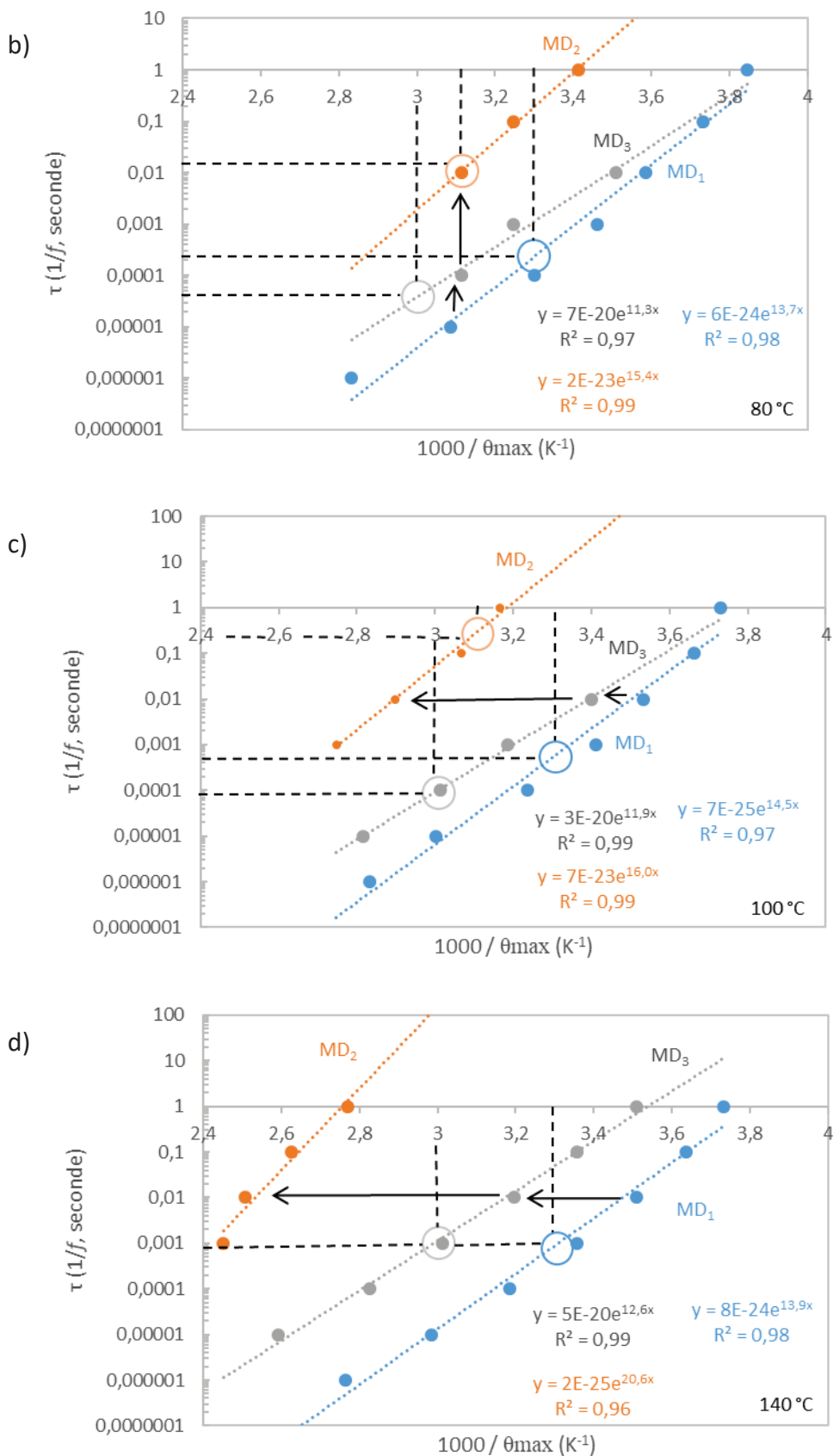


Figure 4-12. Analyse de la dépendance en fréquence des processus de relaxation après différents traitements thermiques. a) Superposition des courbes de refroidissement de la fréquence 1 Hz à 10<sup>2</sup> Hz après le traitement thermique à 80 °C (R 80 °C). b) Superposition des courbes R 80 °C de 10<sup>3</sup> Hz à 10<sup>6</sup> Hz. c) Superposition des courbes R 100 °C de 1 Hz à 10<sup>2</sup> Hz. d) Superposition des courbes R 100 °C de 10<sup>3</sup> Hz à 10<sup>6</sup> Hz. e) Superposition des courbes R 140 °C de 1 Hz à 10<sup>2</sup> Hz. f) Superposition des courbes R 140 °C de 10<sup>3</sup> Hz à 10<sup>6</sup> Hz. g) Superposition des courbes R 180 °C de 1 Hz à 10<sup>2</sup> Hz. h) Superposition des courbes R 180 °C de 10<sup>3</sup> Hz à 10<sup>6</sup> Hz.

Les dépendances en température des temps de relaxation tau  $\tau$  ( $\tau = 1/f$ ) des trois processus de relaxation sont représentées dans la figure 4-13. Les temps de relaxation dépendent linéairement de la température pour les trois processus de relaxation.





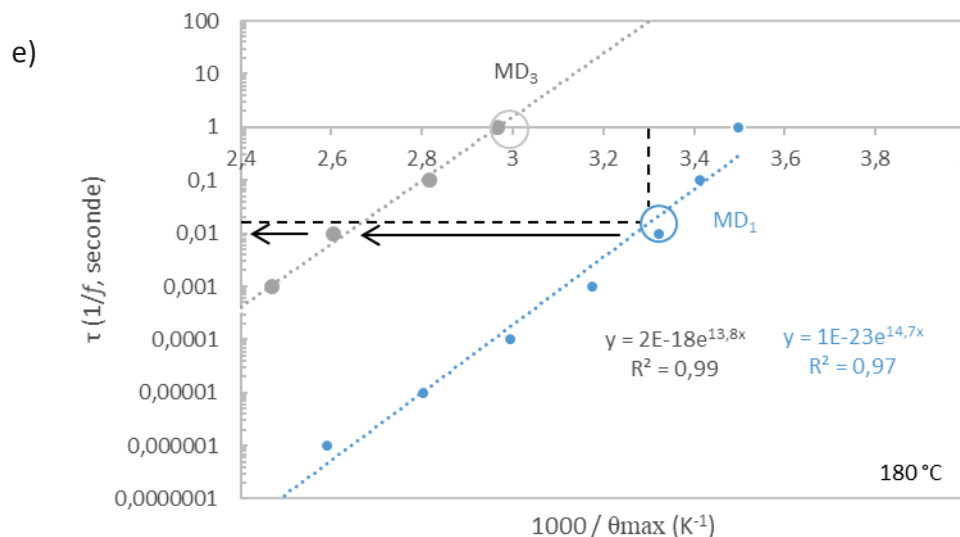


Figure 4-13. Dépendance en température des temps de relaxation après différents traitements thermiques. a) Analyse des dépendances des courbes de refroidissement après le deuxième traitement thermique à 60 °C. b) Après le traitement thermique à 80 °C. c) A 100 °C. d) A 140 °C. e) A 180 °C. Les flèches verticales regardent la variation du temps de relaxation des différents processus de relaxation pour une position du pic de relaxation en température identique. Les flèches horizontales montrent la variation de la position des différents processus de relaxation en température avec un temps de relaxation identique. Les ronds visualisent les temps de relaxation pour un pic de relaxation en température identique après chaque traitement thermique pour les trois processus de relaxation.

Trois caractéristiques du signal changent en fonction du traitement thermique et des processus de relaxation.

Premièrement, les temps de relaxation des dipôles avec des positions en température (température où l'intensité du pic est maximale) similaires deviennent beaucoup plus lents pour MD<sub>2</sub> et MD<sub>3</sub> à mesure que les températures du traitement thermique augmentent (cercle) après le traitement thermique à 80 °C. Pour MD<sub>1</sub>, les temps de relaxation des dipôles avec une même position en températures du pic de relaxation restent similaires jusqu'au traitement thermique à 140 °C et ne ralentisse qu'après le traitement thermique à 180 °C (cercle).

Les temps de relaxation des dipôles avec des positions des pics de relaxation en température similaire deviennent plus lents à travers les processus de relaxation de MD<sub>1</sub> à MD<sub>3</sub> à MD<sub>2</sub> (flèches verticales). Ce qui signifie que le dépliement thermique conduit à des mouvements de plus grande échelle spatiale sans impacter l'environnement local du dipôle (car pas de changement dans les positions des pics de relaxation en température). Ces mouvements pourraient correspondre à des mouvements d'interfaces.

Les temps de relaxation dans MD<sub>3</sub> sont dix fois plus lents que ceux de MD<sub>1</sub> à 100 °C, cent fois plus lents à 140 °C et dix mille fois plus lents à 180 °C quelles que soient les positions des pics de relaxation en température. De même, les temps de relaxation des dipôles dans MD<sub>2</sub> sont milles fois plus lents que dans MD<sub>1</sub>. Ces résultats indiquent que les dipôles fluctuants à différentes fréquences sont indépendants les uns des autres, même pendant le dépliement thermique de la toxine.

Deuxièmement, les positions des pics de relaxation en température avec des temps de relaxation similaires augmentent (diminution de  $1000/K$ ) avec l'augmentation de la température des traitements thermiques (flèches horizontales). Les positions des pics de relaxation en température avec des temps de relaxation similaires augmentent de MD<sub>1</sub> à MD<sub>3</sub> à MD<sub>2</sub>. Ceci signifie que le dépliement thermique conduit à des changements au voisinage des dipôles (changements dans les positions des pics de relaxation en température) sans impact sur le temps de relaxation des dipôles (pas de changement de  $\tau$ ). Ces résultats pourraient correspondre à un dipôle allant d'une position enfouie à une position en surface mais en gardant la même fluctuation.

Troisièmement, les pics de relaxation changent en termes de temps de relaxation et de position en température. Par exemple, il n'y a pas de pic de relaxation avec un temps de relaxation lent pour MD<sub>3</sub> après le traitement thermique à 100 °C mais il y en a après celui à 140 °C. Cela signifie que les dipôles associés à MD<sub>3</sub> avec des temps de relaxation rapides après le traitement thermique à 100 °C ont ralenti et que leur position en température a changé à 140 °C. Par conséquent, dans ce cas, le dépliement thermique entraîne des changements locaux qui se propagent à plus grande échelle entraînant la détection des dipôles à des fréquences plus basses, synonymes de mouvements collectifs plus lents impliquant plus de dipôles.

La dépendance en température des temps de relaxation des trois processus de relaxation correspond à une équation d'Arrhenius (Eq. 3-22) pour les cinq traitements thermiques (60 °C, 80 °C, 100 °C, 140 °C et 180 °C). Les énergies d'activations sont d'environ 27-30 kcal/mol pour les trois processus de relaxation et pour les différentes températures de traitement thermique, à l'exception de MD<sub>2</sub> après le traitement thermique à 140 °C où l'énergie d'activation est d'environ 41 kcal/mol. De telles énergies d'activations similaires à travers les processus de relaxation et les températures des traitements thermiques sont compatibles avec la surveillance de tous les processus de relaxation mais dans des environnements moléculaires différents lorsque que la toxine se déplie localement (autour du dipôle) ou à plus grande échelle.

En conclusion, outre la détection de MD<sub>1</sub>, le nanoconfinement couplé avec le BDS distingue également deux autres processus de relaxation (MD<sub>2</sub> et MD<sub>3</sub>) correspondant à deux autres populations de conformations protéiques induites par les traitements thermiques. Les informations sur l'assemblage et le désassemblage de CtxB<sub>5</sub> en condition macroscopique seront utilisées pour interpréter la nature des populations conformationnelles induites par les traitements thermiques (73,77,127–130). Petite précision, la toxine dans les conditions nanoconfinée et faiblement hydratée est plus résistante aux traitements thermiques que la toxine en conditions macroscopique et hydratée, comme le prouve la détection d'un processus de relaxation jusqu'à 180 °C dans nos expériences versus 74 °C en macroscopique (128).

Un tel impact sur la stabilité des toxines est compatible avec la détection d'espèces désassemblées et/ou dépliées, rarement observées autrement, et qui seraient rendues plus

stables dans les conditions de nanoconfinement. Le nanoconfinement, la déshydratation ou une combinaison des deux pourraient être responsable de la haute résistance aux températures. Les conditions de déshydratation sont connues pour avoir un impact sur la dynamique des protéines (131–133). Les liaisons hydrogènes impliquant l'eau pourraient être affaiblies en cas de déshydratation tandis que la présence du solvant (PBS) pourrait favoriser les ponts salins et conduire à une toxine plus stable. Ceci serait compatible avec le rôle stabilisant des ponts salins vis-à-vis de la résistance à la température observée dans les protéines thermophiles (134–136).

Des expériences de BDS sur la myoglobine et d'autres protéines dans des conditions hydratées macroscopiques rapportent un processus de relaxation rapide correspondant à la dynamique de l'eau libre qui disparaît lors de la déshydratation et des processus de relaxation plus lents associés à la dynamique protéine-eau et aux protéines, toujours détectées lors de la déshydratation (56–58). Shiro et al en 2009 ont également montré en utilisant le BDS que la myoglobine confinée dans des gels de silice ne présente que des processus de relaxation lents comme observés dans nos expériences, et peu de changement de dynamique suivant le niveau de déshydratation de l'échantillon. La seule différence observée dans la myoglobine confinée correspond à un léger ralentissement de la dynamique moléculaire mais sans changement dans l'intensité des pertes diélectriques en diminuant le niveau d'hydratation (59). Ainsi le niveau d'hydratation n'impacterait pas l'enrichissement et/ou la sélection conformationnelle mais plutôt la fréquence à laquelle la dynamique moléculaire serait détectée. Le nanoconfinement pourrait augmenter la compacité de la structure protéique (137) et ainsi stabiliser la protéine comme observé pour les protéines thermophiles (125). L'étude de différents nanoconfinements et de différentes conditions de déshydratations serait une façon de tester le rôle de ces deux paramètres dans la résistance à la température de la toxine. Les chapitres 5 et 6, analyseront l'effet du confinement en faisant varier dans un premier temps la longueur puis le diamètre du confinement.

Les structures dynamiques des protéines détectées dans des conditions expérimentales inhabituelles (nanoconfinée et faiblement hydratée) montrent la robustesse de la structure des protéines et de la dynamique des protéines à une large gamme de conditions, en accord avec la durabilité des protéines dans le temps et dans les différentes conditions expérimentales (138–140). La présence d'un ensemble de structures et de dynamiques plutôt que de solution unique pour une fonction protéique est l'un des paramètres fondamentaux expliquant cette durabilité (5).

Le signal diélectrique spécifique à la conformation des protéines évolue avec la température des traitements thermiques, et est sensible à la perturbation des interactions atomiques et moléculaires locales. Ainsi, le nanoconfinement couplé avec le BDS permet de caractériser une protéine en termes de conformations, de stabilité (température) et de dynamique grâce à de multiples mesures, telles que les fréquences et les positions des pics de relaxation en température. Cette approche pourrait donc être utilisée pour détecter des mutations en comparant les signaux diélectriques d'une protéine et de ses séquences variantes, et ainsi diagnostiquer l'impact des mutations.



## 4.5 Interprétation du signal diélectrique

Actuellement avec les données de BDS, seul le signal diélectrique global peut être utilisé pour distinguer les processus de relaxation et les conformations protéiques. En effet, il est impossible de déterminer les dipôles locaux responsables du signal ou l'amplitude des fluctuations collectives sondées. En d'autres termes, il est impossible de déterminer si la protéine entière est sondée ou si uniquement des sous-domaines sont regardés. De plus, l'échantillon protéique est un mélange de trois composés : la protéine, le PBS et l'eau liée aux protéines. Comme chaque composé possède de nombreux dipôles, la participation individuelle de chaque composé ou comment les signaux individuels se combinent pour donner le signal global est indissociable. En revanche, il est possible de regarder nos données par rapport aux données expérimentales et théoriques connues sur la dynamique des protéines pour tirer des hypothèses sur l'échelle des fluctuations collectives détectées par le BDS. Des hypothèses sur le dépliement thermique des toxines sont émises en accord avec les données établies sur le dépliement et repliement des toxines (72,94,128,141).

Les mouvements des chaînes latérales des acides aminés déterminés par d'autres approches expérimentales et simulations de dynamique moléculaire sont plus rapides de plusieurs ordres de grandeur ( $\leq$  nanoseconde) que la plage de fréquences mesurée par le BDS. Il est donc raisonnable de présumer que de tels mouvements ne sont pas détectés dans le signal diélectrique. Selon les simulations de dynamique moléculaire et les approches intégratives, les mouvements de microsecondes à millisecondes sont attribués aux éléments structuraux 2D, tandis que les mouvements de milliseconde à secondes sont attribués aux mouvements de repliement, dépliement des domaines, interfaces et des chaînes (3,54). Des fluctuations de dipôles de la microseconde à la seconde ont déjà été observées sur d'autres protéines via le BDS et ces mouvements ont été attribués au squelette peptidique (backbone), aux domaines à large échelle ou aux chaînes latérales enfouies (56,59).

Comme MD<sub>1</sub> est le pic principal détecté après les traitements thermiques de 60 °C à 140 °C, nous supposons qu'il correspond à l'état le plus stable de la toxine et à sa conformation la plus repliée, soit un pentamère de type natif. MD<sub>3</sub> est similaire en terme de plage de fréquence mais avec des temps de relaxation plus lents que MD<sub>1</sub>, nous supposons qu'il s'agit d'un pentamère partiellement replié où les dipôles ont un environnement moins compact permettant des mouvements à plus grande échelle. MD<sub>2</sub> est le processus de relaxation le plus lent avec un large pic, nous supposons qu'il s'agit de l'espèce CtxB la moins repliée détectée et donc un mélange d'intermédiaires d'assemblage de CtxB (tétramères, trimères et dimères). Alternativement, MD<sub>1</sub> pourrait être un monomère CtxB replié, MD<sub>3</sub> un monomère légèrement moins replié et MD<sub>2</sub> plusieurs états dépliés de monomère CtxB. Toutefois, les processus de relaxation disparaissent à des températures élevées, par conséquent il est difficile d'imaginer quelles seraient les espèces non détectées si des monomères dépliés sont détectés. La première hypothèse est également cohérente avec le (dés)assemblage et le (dé, re) plissement des intermédiaires CtxB précédemment identifiés macroscopiquement (72,94).

Dans la première hypothèse, il est raisonnable de supposer que les interfaces des sous-unités de toxine sont sondées pendant l'expérience puisque les intermédiaires d'assemblage de CtxB seraient détectés (MD<sub>2</sub>). Il s'ensuit que les monomères CtxB seraient non détectés. Le pentamère de la toxine native possède trois zones d'interfaces distinctes : I<sub>1</sub> (interface N-terminale), I<sub>2</sub> (interface β principale) et I<sub>3</sub> (interface de l'hélice α), mais il est peu probable que les interfaces natives soient détectées car elles impliquent des mouvements moléculaires locaux avec des fluctuations à l'échelle de la nanoseconde (3).

Dans MD<sub>1</sub>, des temps de relaxation rapides et lents sont détectés et lorsque la toxine se déplie avec le traitement à haute température, seuls les temps de relaxation lents sont détectés pour MD<sub>2</sub> aux différentes températures de traitements thermiques et pour MD<sub>3</sub> après le traitement thermique à 180 °C (Fig. 4-12). Les interfaces de toxine I<sub>1</sub>, I<sub>2</sub> et I<sub>3</sub> existeraient comme natives (échelle nanoseconde : non détectée), déstabilisées (fluctuations de la microseconde à la milliseconde : mobilité des structures 2D) et altérées avec des mouvements plus grands et plus lents (milliseconde à seconde : mouvements de domaines, interfaces et structures tertiaires). Ces hypothèses sont cohérentes avec les mouvements d'échelle de temps de la microseconde à la milliseconde mesurés par les échanges d'hydrogènes impliqués dans le dépliement thermique des protéines (142).

Le modèle suppose que les intermédiaires d'assemblage de CtxB sont détectés via les mouvements lents des interfaces et plus précisément les interfaces déstabilisées et/ou altérées.

Une approche intégrative est développée pour mieux comprendre les interfaces potentiellement impliquées dans le signal diélectrique. Pour cela, Lorenza Pacini (doctorante au laboratoire Ampère) et moi-même avons travaillé ensemble pour modéliser un réseau composé des interactions intermoléculaires atomiques (distance seuil de 5 Å) entre les acides aminés à partir de la structure pentamérique native de la choléra toxine obtenue par diffraction de rayon X (PDB : 1EEI) (91).

Les acides aminés regardés dans ce réseau sont les acides aminés chargés et leurs voisins (acides aminés chargés et non chargés) à une distance seuil de 5 Å. Ces acides aminés sont choisis pour différentes raisons :

- Les structures protéiques et la dynamique protéique dépendent des interactions hydrophobes non-spécifiques (Vander Waals) et d'interactions électrostatiques spécifiques (pont salin, dipôle chargé). Les interactions électrostatiques devraient participer davantage au signal diélectrique car elles impliquent des forces coulombiennes. De plus, les interactions électrostatiques sont essentielles pour le repliement et la stabilité des protéines et joueraient même un rôle majeur dans la stabilité thermique des protéines (71).
- Les résidus chargés et leurs voisins proches sont les principaux constituant responsables des interactions électrostatiques (143).
- Les interactions électrostatiques dans la structure des protéines dépendent du pH et des conditions salines. Or dans les expériences de BDS des protéines, de l'eau et du sel (PBS) sont présents.

- Dans un milieu ionisé les dipôles électrostatiques deviennent prépondérants, par conséquent il est possible de négliger les autres dipôles.

Dans les conditions du confinement à 40 nm de diamètre et 10  $\mu\text{m}$  de long utilisées dans ce chapitre, le pH des solutions testées en BDS est de 6,9. D'après le pKa des acides aminés, uniquement les extrémités N et C-terminales, et les acides aminés : lysine, arginine, acide glutamique et acide aspartique sont considérés comme protonnés à ce pH.

Précédemment nous avons suggéré que les conformations associées à MD<sub>1</sub> et MD<sub>3</sub> seraient des pentamères non natifs, et que MD<sub>2</sub> correspondrait à des intermédiaires d'assemblage. Par conséquent, il est raisonnable de supposer que le signal de BDS sonde les interfaces (I<sub>1</sub>, I<sub>2</sub> et I<sub>3</sub>) de la toxine. Les interfaces doivent être diélectriquement détectables, de plus ces interfaces doivent avoir des sensibilités thermiques distinctes pour sonder les mouvements spatio-temporels de la protéine. Ces trois interfaces sont-elles-même sous-divisées en plusieurs interfaces représentées dans la figure 4-14 :

- L'interface I<sub>1</sub> correspondant au domaine N-terminal est divisée en deux interfaces (Fig. 4-14c) : I<sub>1a</sub> qui comprend les résidus 1 à 3 de la chaîne E en interactions avec les résidus 92 et 93 de la chaîne adjacente F / I<sub>1b</sub> qui correspond à l'interaction des résidus 1 à 12 de la chaîne E avec les résidus 28 à 39 de la chaîne F.
- L'interface I<sub>2</sub> correspondant à l'interface principale comprenant l'interaction entre le brin beta numéro 2 (résidus 26 à 32) d'un monomère M et le brin beta numéro 6 (résidus 96 à 103) du monomère adjacent M<sub>+1</sub> : I<sub>2a</sub> comprend les résidus 25 à 32 de la chaîne E en interactions avec les résidus 88, et de 96 à 103 de la chaîne adjacente D (Fig. 4-14d) / I<sub>2b</sub> comprend les résidus 28 à 39 de la chaîne E en interactions avec les résidus 58 à 68 de la chaîne D (les résidus 58 à 81 participe à l'hélice central) (Fig. 4-14d) / I<sub>2c</sub> comprend les résidus 101 à 103 de la chaîne E en interactions avec les résidus 73 à 77 de la chaîne adjacente F (Fig. 4-14e).
- L'interface I<sub>3</sub> de l'hélice  $\alpha$  centrale (résidu 60 à 78) est une interface tripartie composée des résidu 63 à 77 de la chaîne E en interactions avec les résidus 63 à 81 de la chaîne D, et les résidus 67 à 81 de la chaîne E en interactions avec les résidus 66 à 77 de la chaîne F. Il s'agit d'une interface tripartie car elle implique des interactions entre la chaîne E et D ainsi qu'entre la chaîne E et F ; mais elle constitue toutefois une interface unique (Fig. 4-14f).

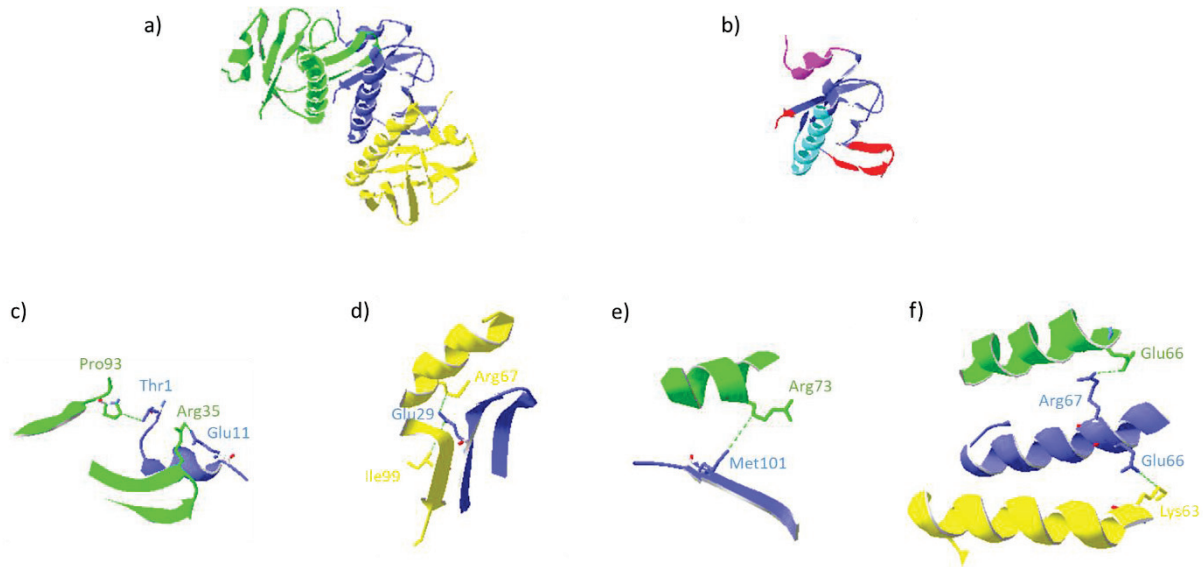


Figure 4-14. Représentation des interfaces dans CtxB. a) Les chaînes D (jaune), E (bleu) et F (vert) de la toxine sont représentées. b) Zoom sur la chaîne E, en violet résidus impliqués dans l'interface  $I_1$ , en rouge résidus impliqués dans l'interface  $I_2$  et en bleu ciel résidus impliqués dans l'interface  $I_3$ . c) Représentation de l'interface  $I_1$  avec la formation d'un dipôle induit (Thr1E-Pro93F) dans l'interface  $I_{1a}$  et la formation d'un dipôle ionique (Glu11E-Arg35F) dans l'interface  $I_{1b}$ . d) Représentation de l'interface  $I_{2a}$  avec un dipôle induit (Glu29E-Ile99D) et l'interface  $I_{2b}$  avec un dipôle ionique (Glu29E-Arg67D). e) Représentation de l'interface  $I_{2c}$  avec un dipôle induit (Met101E-Arg73F). f) Représentation de l'interface  $I_3$  avec deux dipôles ioniques (Arg67E-Glu66F et Glu66E-Lys63D). Les chaînes latérales des acides aminés formant un dipôle sont de la même couleur que leur chaîne.

Basé sur le nombre d'interactions atomiques et de paires d'acides aminés, les interfaces sont classées comme suit : les interfaces  $I_{1a}$  et  $I_{2c}$  sont les plus faibles avec respectivement 46 interactions atomiques, 5 paires d'acides aminés et 22 interactions atomiques et 5 paires d'acides aminés. Les interfaces  $I_{2a}$  et  $I_3$  sont les interfaces les plus fortes avec 198 interactions atomiques et 26 paires d'acides aminés et 214 interactions atomiques et 24 paires d'acides aminés. Par conséquent, les interfaces  $I_{1b}$  et  $I_{2b}$  sont dites modérées avec 137 interactions atomiques, 19 paires d'acides aminés et 136 interactions atomiques et 19 paires d'acides aminés.

Toutefois, pour être visible diélectriquement, il est préférable que les paires d'acides aminés et les interactions atomiques se produisent via des acides aminés chargés. L'interaction entre deux acides aminés chargés donnera un dipôle ionique, alors que l'interaction entre un acide aminé chargé et un acide aminé non chargé donnera un dipôle induit.

Les interfaces  $I_{1a}$  et  $I_{2c}$  possèdent le plus faible nombres d'interactions et pas de dipôles ioniques. Toutefois, ces deux interfaces possèdent chacune deux dipôles induits. Par conséquent ces deux interfaces sont considérées comme invisible ou faiblement visible diélectriquement. L'interface  $I_{2a}$  possède de nombreuses interactions, pas de dipôle ionique, mais quatre dipôles induits. Par conséquent l'interface  $I_{2a}$  est aussi considérée comme invisible ou faiblement visible diélectriquement. Les interfaces  $I_{1b}$  et  $I_{2b}$  possèdent respectivement deux et un dipôles ioniques et quelques dipôles induits, par conséquent ces interfaces devraient

être visible diélectriquement. L'interface  $I_3$  est l'interface diélectrique la plus puissante car elle possède 14 dipôles ioniques et 5 dipôles induits.

L'interface  $I_3$  possède la plus forte stabilité (nombre d'interactions atomiques et nombre de paires d'acides aminés) avec le plus grand nombre de dipôles ioniques, par conséquent, cette interface devrait être détectée dans l'état le plus déplié de la toxine. Ceci suggère que les dipôles de  $MD_2$  correspondent à cette interface. Les interfaces  $I_{1a}$  et  $I_{2c}$  sont les interfaces les plus faibles donc les interfaces déstabilisées en premiers, toutefois la détection de ces interfaces est peu visible diélectriquement. Le dépliement de l'interface  $I_{1a}$  (résidu 1 à 3 de la chaîne E) conduit à la perturbation de l'interface  $I_{1b}$  (résidu 1 à 12 de la même chaîne). Ces résultats suggèrent que l'interface  $I_{1b}$  et potentiellement les interfaces  $I_{1a}$  et  $I_{2c}$  participent au signal diélectrique  $MD_1$  et plus particulièrement aux dipôles avec des mouvements lents. La perturbation de l'interface  $I_{2c}$  devrait entraîner la perturbation de l'interface  $I_{2a}$  et donc la détection de cette interface dans les mouvements rapides de  $MD_1$  et lors des traitements thermiques à haute température. L'interface  $I_{2b}$  participerait par conséquent, au processus de relaxation  $MD_3$ . Un modèle possible des interfaces perturbées pendant la dénaturation thermique de CtxB<sub>5</sub>, cohérent avec le (dés)assemblage et le (dé, re) pliage des intermédiaires CtxB identifiés macroscopiquement (72,94) est représenté dans la figure 4-15.

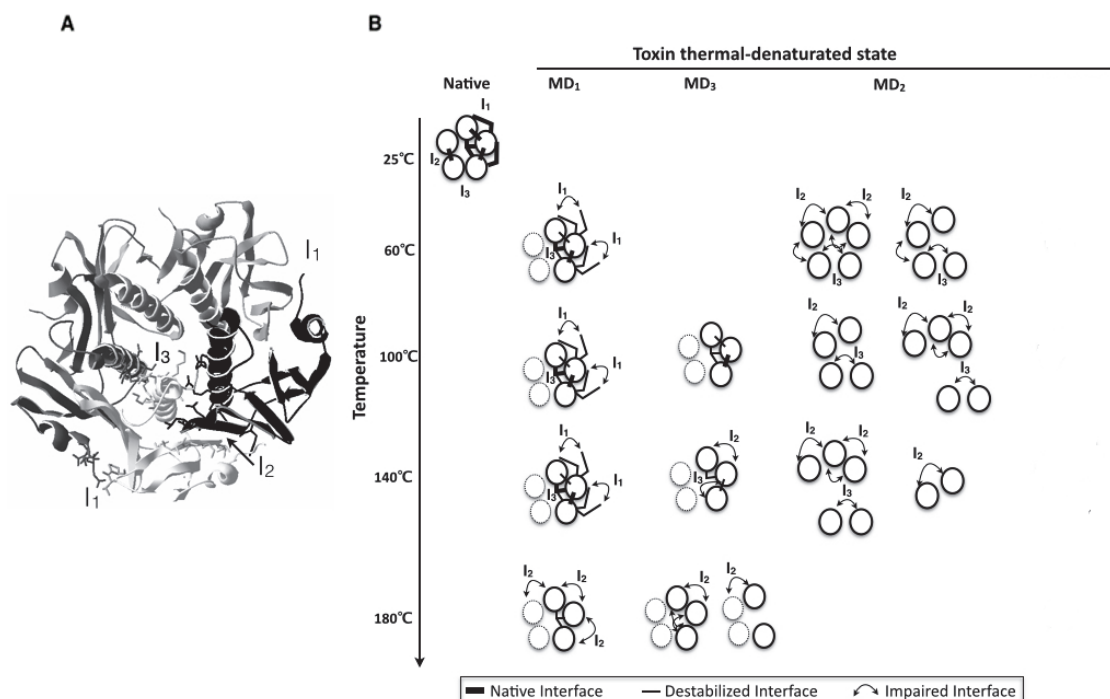


Figure 4-15. Modèle de la dénaturation thermique de CtxB<sub>5</sub>. a) Représentation de la structure atomique de CtxB<sub>5</sub> avec les trois interfaces. b) Le monomère de CtxB est représenté par un cercle et CtxB<sub>5</sub> est donc composé de cinq cercles. Uniquement les interfaces déstabilisées et altérées sont détectées par BDS.

Dans ce modèle, les monomères ne sont pas détectés, toutefois nous ne pouvons pas exclure l'hypothèse que les dipôles intramoléculaires de CtxB<sub>5</sub> ne participent pas aussi au processus de relaxation. Par conséquent, les monomères et plus particulièrement les monomères repliés pourraient potentiellement être détectés dans  $MD_2$ .

En conclusion, le nanoconfinement et la spectroscopie diélectrique à large bande sont utilisables pour étudier les processus de relaxation des conformations protéiques car celles-ci possèdent un signal diélectrique caractéristique et reproductible. Les signaux diélectriques des protéines évoluent avec la température des traitements thermiques, confirmant la sensibilité des changements d'interactions moléculaires par rapport au dépliement thermique des protéines.

La dynamique moléculaire de conformations protéiques distinctes, avec des conformations rarement détectées macroscopiquement, deviennent accessible expérimentalement en combinant le nanoconfinement et la spectroscopie diélectrique à large bande. Cette nouvelle approche capture les fluctuations dipolaires sous-jacentes à la dynamique moléculaire des conformations protéiques dont les mouvements lents s'associent au dépliement des protéines. L'avantage du BDS est sa capacité à caractériser les caractéristiques dipolaires à la fois à l'échelle locale avec la position du pic de relaxation en température mais aussi à une échelle plus large correspondant aux fluctuations des mouvements collectifs. Ainsi, le niveau de description fourni par le BDS offre un grand potentiel pour étudier la dynamique des protéines et est un équivalent expérimental à la simulation de la dynamique moléculaire sur la plage de fréquence de  $10^3$  Hz à  $10^6$  Hz.

Dans la partie 3.3.5, nous avons expliqué comment interpréter les résultats diélectriques sur d'autres molécules que les protéines en utilisant la spectroscopie diélectrique à large bande. Ce chapitre montre que les interprétations émises dans cette partie sont applicables pour les protéines nanoconfinées comme le montre la figure 4-16 issue de la figure 4-10a.



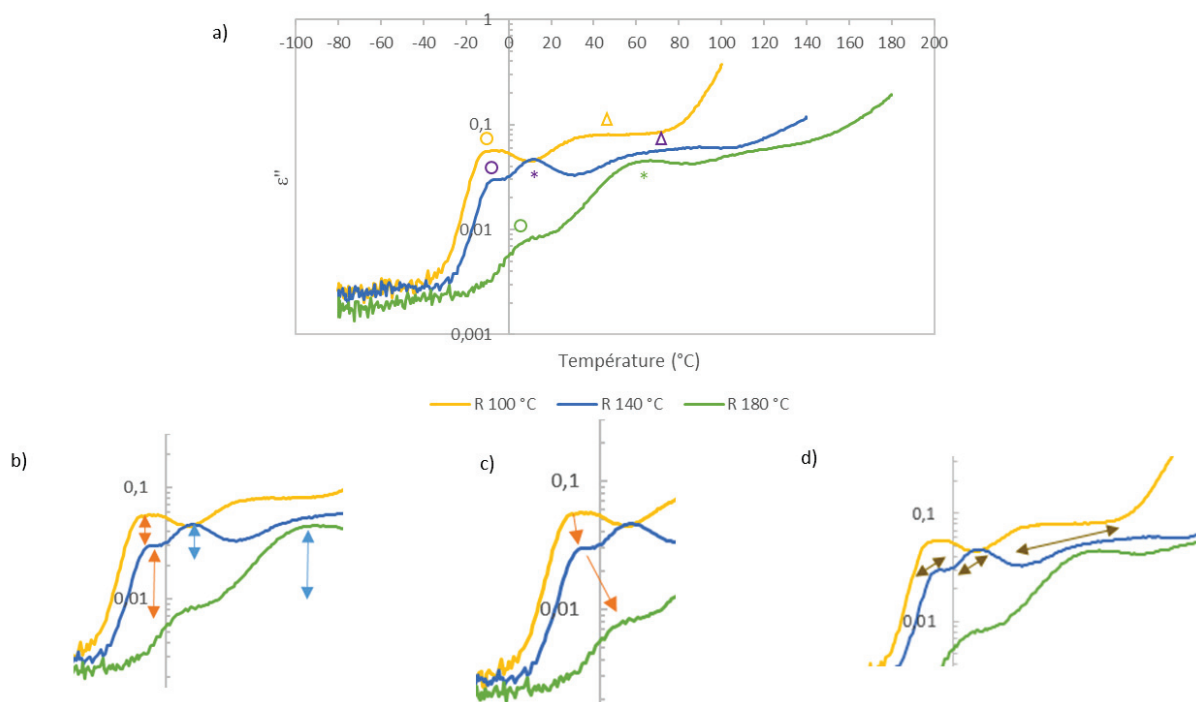


Figure 4-16. Résumé des paramètres pour analyser le signal diélectrique détecté. a) Représentation des courbes de refroidissement après différents traitements thermiques à la fréquence de 1Hz. b) Représentation du paramètre intensité du pic. c) Représentation du paramètre position du pic de relaxation. d) Représentation du paramètre largeur du pic.

Dans la figure 4-16a, après un même traitement thermique plusieurs pics de relaxation avec des positions en température différentes sont détectés. Ces différences de positions, traduisent la présence de différentes populations conformationnelles. Le symbole rond représente la population conformationnelle MD<sub>1</sub>, le symbole triangle la population conformationnelle MD<sub>2</sub> et le symbole étoile la population conformationnelle MD<sub>3</sub> pour les différents traitements thermiques. Dans la figure 4-16b, l'intensité du pic de relaxation associé à MD<sub>1</sub> diminue avec l'augmentation des traitements thermiques, traduisant une diminution du nombre de dipôles détectés. Inversement le processus de relaxation MD<sub>3</sub> non détecté après le traitement thermique à 100 °C apparaît après le traitement thermique à 140 °C, traduisant la détection d'une nouvelle population conformationnelle. De plus, l'intensité de ce pic de relaxation augmente après le traitement thermique à 180 °C, caractéristique d'un changement de densité de dipôles détectés. Dans la figure 4-16c, la position du pic de relaxation en température augmente avec l'augmentation des traitements thermiques dans une population conformationnelle traduisant un changement structural à l'intérieure de cette population. Dans la figure 4-16d, la largeur du pic du processus de relaxation MD<sub>2</sub> est plus importante que la largeur des pics de relaxation MD<sub>1</sub> et MD<sub>3</sub> caractéristique d'une hétérogénéité conformationnelle plus importante.

## Chapitre 5

# Impact du confinement : Variation de la longueur

L'état natif des protéines est défini comme un ensemble de conformations plutôt que par une conformation unique (144). Les différences structurales entre les conformations caractérisent la diversité conformationnelle de la protéine.

Le concept d'ensemble conformationnel est essentiel pour expliquer les propriétés des protéines (145,146), telles que :

- La promiscuité des enzymes et des anticorps, en effet un anticorps se lie à un antigène mais grâce à sa diversité conformationnelle un anticorps est aussi capable de se lier à plusieurs antigènes (147).
- La transduction du signal qui repose sur les propriétés dynamiques intrinsèques des protéines, plus précisément, la capacité des protéines à changer de conformation suite à la liaison d'un ligand, d'une autre protéine... (18).
- La reconnaissance protéine-protéine : les protéines possèdent plusieurs conformations protéiques grâce à la mobilité des acides aminés et adaptent leurs conformations suite à l'interaction avec d'autres protéines. Cette adaptation semble facilitée par la présence de certains acides aminés moins mobiles que d'autres qui sont pré-organisés dans une conformation semblable à l'état lié. Ces acides aminés participent généralement à la liaison avec d'autres protéines (148).

Au jour d'aujourd'hui, plusieurs bases de données sur la diversité conformationnelle des protéines existent :

- CoDNaS (Conformational Diversity of Native State) est une base de données issue de structures tridimensionnelles redondantes obtenues via différents protocoles expérimentaux pour une même protéine (8). Chaque protéine dans la base de données à en moyenne 15 conformères référencés.
- PCDB (protein conformational diversity database) est une base de données possédant une compilation redondante de toutes les structures cristallographiques obtenues via différentes conditions, et de structures RMN (9). La diversité conformationnelle est estimée à partir d'une collection redondante de structures protéiques de chaque domaine déposé dans la base de données.



Dans les deux bases de données, les différences de structures redondantes sont obtenues suite à la variation de pH, température et à la présence de mutation, ligand, etc... Par conséquent, ces structures possèdent à la fois des diversités structurales locales et globales.

L'ensemble conformationnel est responsable des propriétés de robustesse et d'adaptabilité des protéines. La description de l'ensemble conformationnel est nécessaire pour diagnostiquer l'impact des mutations fragilisant la protéine. Au jour d'aujourd'hui la caractérisation de l'ensemble des conformations protéiques adoptées par chaque protéine reste un défi majeur.

Dans le chapitre précédent (Chapitre 4), nous avons montré que la combinaison du nanoconfinement et de la spectroscopie diélectrique à large bande (BDS) dans des conditions faiblement hydratées permettait l'enrichissement et/ou la sélection d'une population de conformation protéique. De plus, le dépliement d'un pentamère non natif de CtxB<sub>5</sub> a pu être suivi via cette approche et plusieurs conformations non natives ont pu être détectées par des pics de relaxation distincts.

Dans ce chapitre, nous allons étudier si la variation de la longueur du confinement dans des conditions faiblement hydratées permet de cribler différentes conformations protéiques. Pour cela, nous allons analyser le confinement à 40 nm de diamètre et 5 µm de long puis le confinement à 40 nm de diamètre et 2 µm de long et ensuite les résultats de ces deux confinements seront comparés à celui du chapitre précédent (Chapitre 4 : confinement à 40 nm de diamètre et 10 µm de long). Ces comparaisons permettront de déterminer si la variation de longueur du confinement permet de cribler différentes conformations protéiques.

Pour rappel, les dimensions du confinement sont confirmées en utilisant la microscopie électronique à balayage. Puis la fluorescence est utilisée pour vérifier l'état pentamérique de la protéine CtxB<sub>5</sub> avant les mesures de BDS. Ensuite, l'échantillon protéique (composé de CtxB<sub>5</sub>, d'eau liée et de PBS) est déposé sur la nanomembrane avec le confinement choisi via la technique de la goutte pour un remplissage volumique des nanopores à 14,3 %. La quantité de protéine présente dans les nanopores augmente en augmentant la longueur du confinement, mais la proportion d'occupation des protéines est la même dans tous les confinements car le pourcentage de remplissage est identique (Chapitre 3, partie 3.3.3.3).

L'échantillon protéique est ensuite soumis à différents traitements thermiques de trois heures pour des températures allant de 60 °C à 180 °C. Après chaque traitement thermique l'échantillon est refroidi à -80 °C où la température est stabilisée pendant trente minutes avant de remonter à la température du prochain traitement thermique (Fig. 3-16). A la fin de chaque traitement thermique, les données de pertes diélectriques sont prélevées jusqu'à la fin des trente minutes à -80 °C et représentées dans les courbes de refroidissement (exemple : R 80 °C). Les courbes de refroidissement seront étudiées en fonction de la température à différentes fréquences. Le ou les pics de relaxation détectés correspondront à la dynamique moléculaire.

5.1 Confinement à 40 nm de diamètre et 5  $\mu\text{m}$  de long

## 5.1.1 Contrôle de l'échantillon protéique

L'échantillon protéique CtxB<sub>5</sub> à la fréquence de 10<sup>3</sup> Hz après le traitement thermique à 60 °C possède un pic de relaxation à -1,5 °C (symbole rond), toutefois, le contrôle du solvant possède aussi un pic de relaxation à -22 °C (symbole flèche) (Fig. 5-1a). De même, à la fréquence de 1 Hz, l'échantillon du solvant et l'échantillon protéique possède tous les deux un pic de relaxation à -50 °C et -29 °C (Fig. 5-1b).

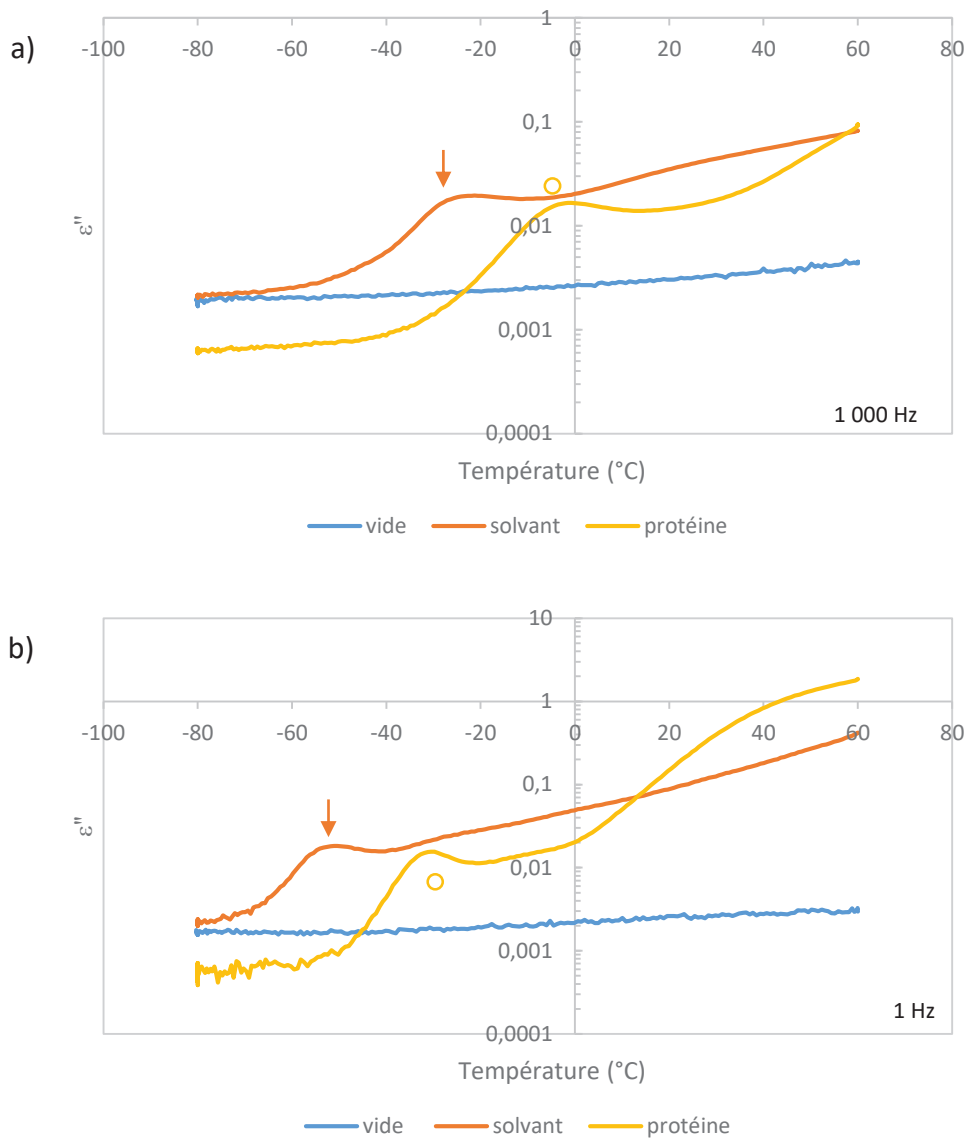


Figure 5-1. CtxB<sub>5</sub> dans le confinement de 40 nm de diamètre et 5  $\mu\text{m}$  de long. a) A la fréquence de 10<sup>3</sup> Hz. b) A la fréquence de 1 Hz. Courbe de refroidissement après le traitement thermique à 60 °C. La courbe en jaune représente l'échantillon protéique composé de 2  $\mu\text{L}$  de toxine à 2,6 mg/ml et 410  $\mu\text{L}$  d'eau déionisée. La concentration protéique finale est de 0,013 mg/ml. La courbe orange représente le contrôle du solvant (2  $\mu\text{L}$  de PBS 1X et 410  $\mu\text{L}$  d'eau déionisée). La courbe bleue correspond à la membrane vide.

Contrairement au confinement à 40 nm de diamètre et 10  $\mu\text{m}$  de long (Fig. 4-3), le pic de relaxation du solvant ne disparaît jamais dans le confinement de 40 nm de diamètre et 5

$\mu\text{m}$  de long suite aux différents traitements thermiques (Fig. 5-2). A la fréquence de 1 Hz (Fig. 5-2), le solvant possède un seul processus de relaxation jusqu'au traitement thermique à 80 °C. La position de ce pic de relaxation en température ne change pas par rapport au traitement thermique à 60 °C et reste à -50 °C. A cette fréquence, après le traitement thermique à 100 °C, un deuxième processus de relaxation devient apparent à -33 °C, de plus la position du premier processus de relaxation en température est décalée à -46 °C. Ces deux processus de relaxation associés au solvant sont détectables jusqu'au traitement thermique à 180 °C. Après le traitement thermique à 80 °C, les intensités des pics de relaxation du solvant diminuent mais sans jamais disparaître. Dans le chapitre 4, cette diminution d'intensité est associée à l'évaporation de l'eau libre, de plus, nous avons supposé que le processus de relaxation du solvant (MDb) provenait principalement de l'eau libre. La présence de MDb (symbole flèche) jusqu'au traitement thermique à 180 °C dans ce confinement suggère que le signal diélectrique du solvant est beaucoup plus complexe et ne provient pas uniquement de l'eau libre. En effet, il est raisonnable de supposer qu'après six heures à 60 °C, trois heures à 80 °C, trois heures à 100 °C et trois heures à 140 °C, l'eau libre devrait être totalement évaporée. Ces résultats suggèrent que le signal diélectrique de MDb provient du PBS, de l'eau liée ou de l'interaction PBS-eau liée.

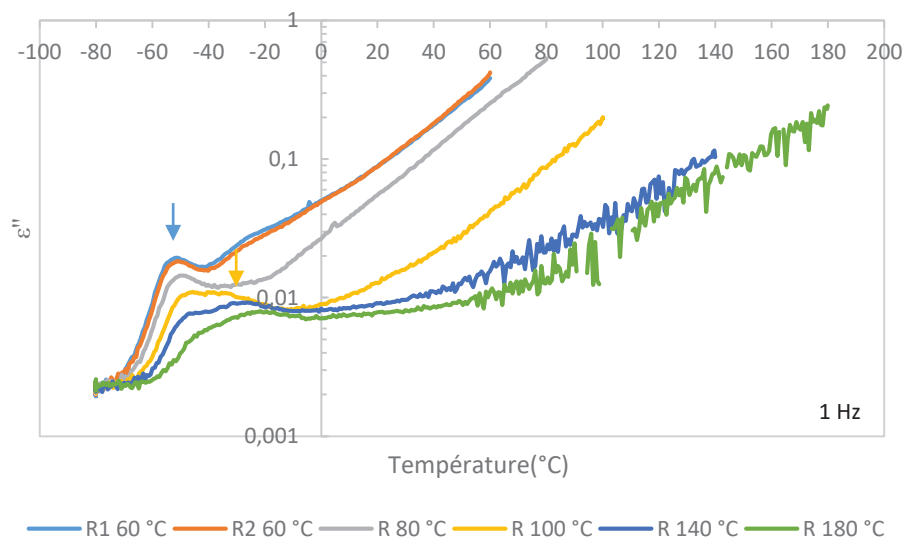


Figure 5-2. Analyse du comportement du solvant à la fréquence de 1 Hz après différents traitements thermiques.

La position des pics de relaxation en température associés au solvant et à la protéine ne sont pas les mêmes pour les différents traitements thermiques (Fig. 5-1b et Fig. 5-3). De même, pour le nombre de processus de relaxation détectés entre les deux échantillons. L'échantillon protéique possède plus de processus de relaxation que le solvant après les traitements thermiques à 80 °C et 100 °C (Fig. 5-3a et Fig. 5-3b). De plus, après le traitement thermique à 180 °C (Fig. 5-3d), l'échantillon protéique ne possède aucun processus de relaxation, ce qui n'est pas le cas pour le solvant. Ces résultats suggèrent que les processus de relaxation détectés dans l'échantillon protéique sont bien caractéristiques des conformations protéiques. Le signal du solvant sera analysé plus en détail ultérieurement (Chapitre 7).

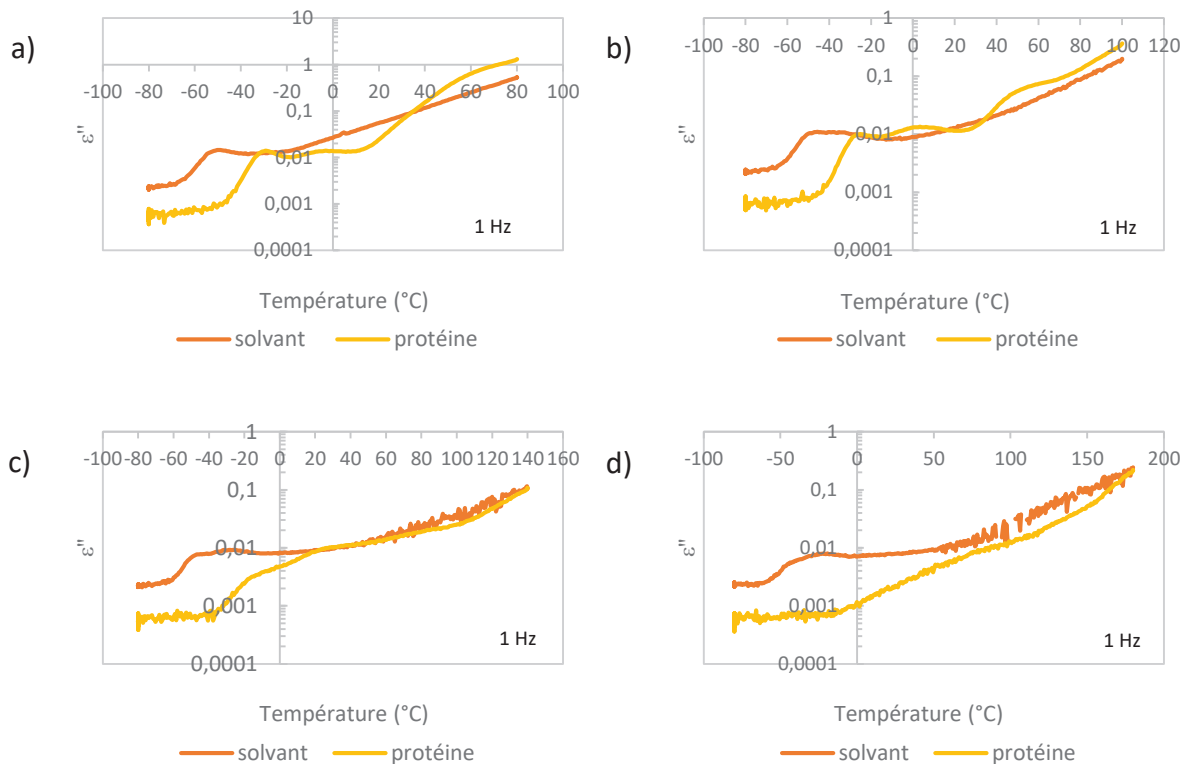


Figure 5-3. Comparaison du signal diélectrique pour les courbes de refroidissement du solvant et de l'échantillon protéique à la fréquence de 1 Hz après différents traitements thermiques. a) Superposition des courbes R 80 °C. b) Superposition des courbes R 100 °C. c) Superposition des courbes R 140 °C. d) Superposition des courbes R 180 °C.

Dans le chapitre 4, les processus de relaxation de l'échantillon protéique ont été principalement attribués aux interfaces électrostatiques des monomères. Cependant, les processus de relaxation détectés dans le confinement à 40 nm de diamètre et 5  $\mu\text{m}$  de long pour le solvant, suggèrent que les interactions électrostatiques PBS-protéine, PBS-eau liée-protéine, eau liée-protéine pourraient aussi participer au signal diélectrique de l'échantillon protéique. Néanmoins, comme son nom l'indique, l'eau liée est liée à la protéine, et le PBS joue un rôle dans la stabilité conformationnelle de la protéine, par conséquent, des informations sur la conformation protéique seront obtenues. Ces résultats sont en accord avec les différentes études protéiques via le BDS (55–58,60) où en raison du couplage étroit entre les molécules d'eau liée et les protéines, la distinction entre la contribution de chaque élément est souvent sujette à débat dans le signal diélectrique (55).

### 5.1.2 Analyse de CtxB<sub>5</sub> à différentes fréquences

La position du pic de relaxation en température varie en fonction de la fréquence (Fig. 3-12), par conséquent, les données de pertes diélectriques en fonction de la température sont étudiées à différentes fréquences. La température du pic de relaxation augmente avec l'augmentation de la fréquence (Fig. 5-4).

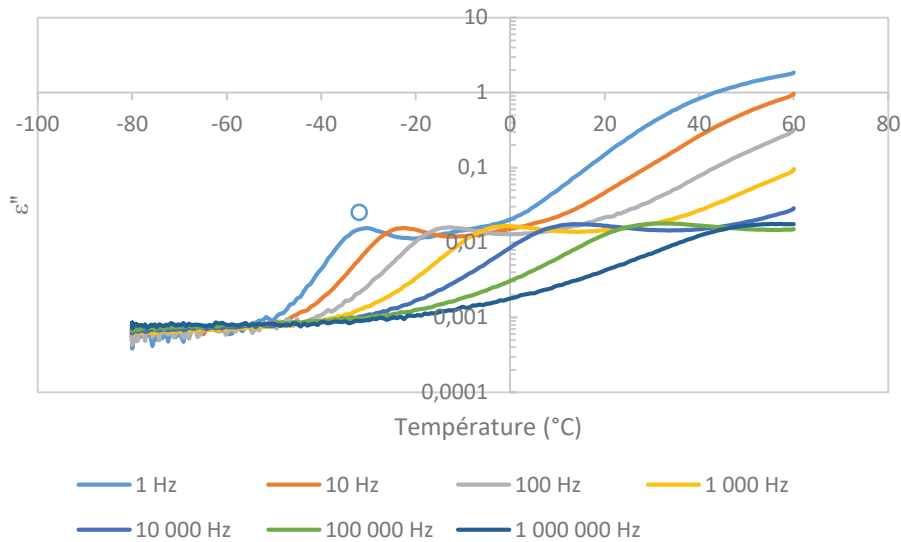
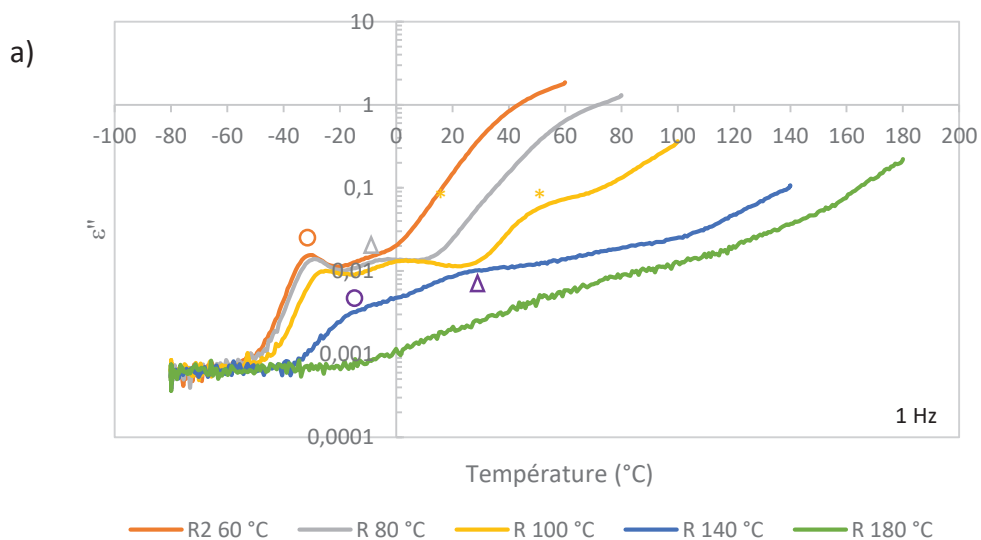


Figure 5-4. Analyse de CtxB<sub>5</sub> à différentes fréquences après le deuxième traitement thermique à 60 °C. Le symbole rond correspond au premier processus de relaxation détecté.

CtxB<sub>5</sub> possède un seul processus de relaxation (symbole rond) de 1 Hz à 10<sup>6</sup> Hz. La position en température du pic de relaxation se situe à -29 °C à la fréquence de 1 Hz, à -1,5 °C à la fréquence de 10<sup>3</sup> Hz et à 54 °C pour la fréquence de 10<sup>6</sup> Hz.

### 5.1.3 Analyse du comportement de CtxB<sub>5</sub> en fonction de la température

La dénaturation thermique de la protéine est étudiée en réalisant différents traitements thermiques. Un ou plusieurs processus de relaxation sont détectés jusqu'au traitement thermique à 140 °C dans la gamme de fréquence étudiée pour l'échantillon protéique de CtxB<sub>5</sub> (Fig. 5-5).



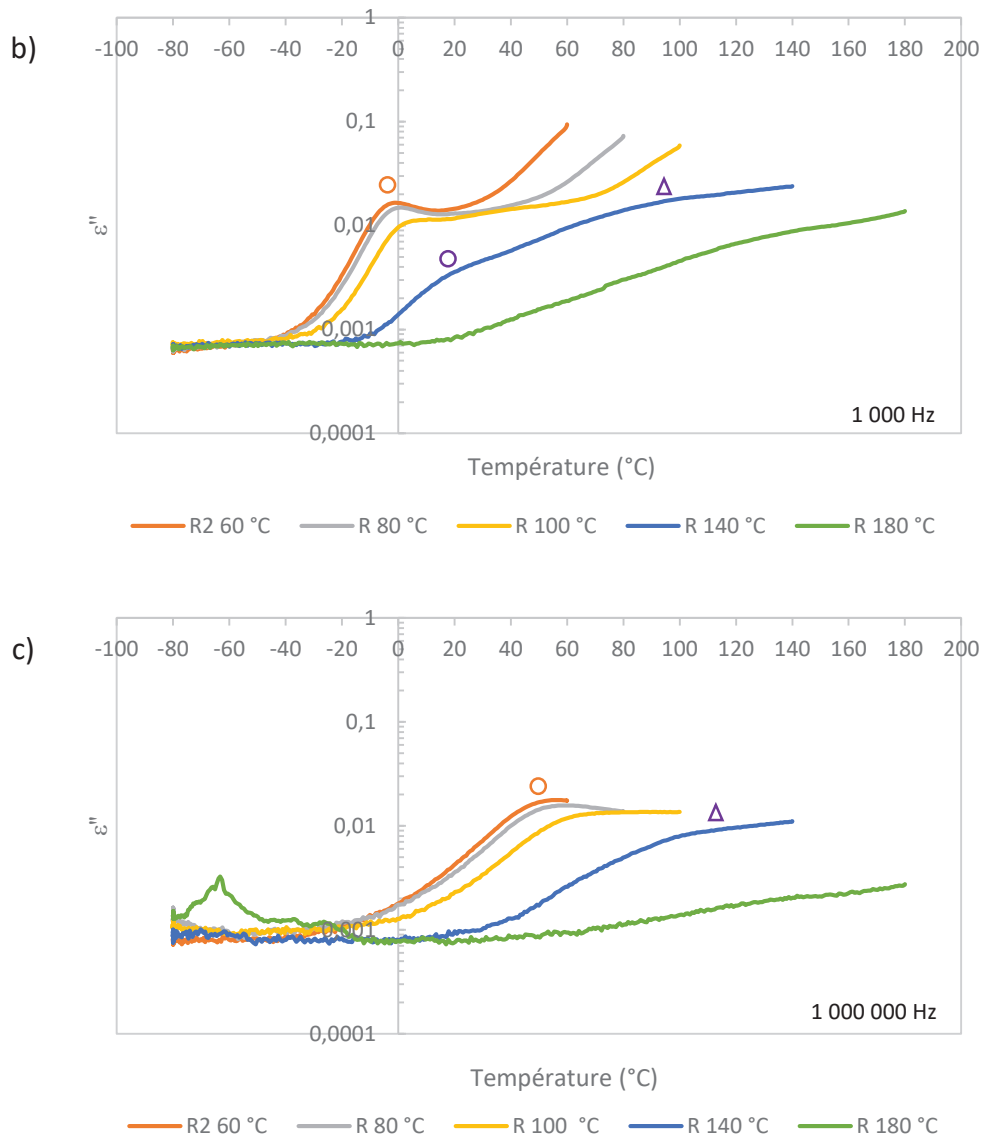


Figure 5-5. Étude du comportement de CtxB<sub>5</sub> après différents traitements thermiques. a) Superposition des courbes de refroidissement de CtxB<sub>5</sub> à la fréquence de 1 Hz. b) Analyse des courbes de refroidissement de CtxB<sub>5</sub> à 10<sup>3</sup> Hz. c) Analyse de CtxB<sub>5</sub> à 10<sup>6</sup> Hz.

A la fréquence de 1 Hz (Fig. 5-5a), après le traitement thermique à 80 °C, le premier pic de relaxation détecté (symbole rond) possède la même position en température qu'après le traitement thermique à 60 °C, soit -29 °C. Toutefois, un deuxième pic de relaxation (symbole triangle) commence à apparaître à -2 °C. Après le traitement thermique à 100 °C, un troisième pic de relaxation (symbole étoile) apparaît à 53 °C. De plus, la position en température des deux premiers pics de relaxation change par rapport au traitement thermique précédent à 80 °C. En effet, la position de ces deux processus de relaxation en température est respectivement décalée à -25 °C et à 4 °C, traduisant un ou plusieurs changements conformationnels des populations protéiques détectées. Après le traitement thermique à 140 °C, la forme du signal diélectrique est totalement différente des températures précédentes. Toutefois, le premier pic de relaxation est toujours détecté à -14 °C et le deuxième à 34 °C. Après 180 °C, aucun processus de relaxation n'est détecté.

A la fréquence de  $10^3$  Hz (Fig. 5-5b), un seul processus de relaxation correspondant au premier pic de relaxation à 1 Hz est détecté jusqu'au traitement thermique à 100 °C. La position du pic de relaxation en température ne change pas jusqu'au traitement thermique à 80 °C et reste à -1,5 °C. Toutefois, la position en température du pic de relaxation est décalée à 5 °C après le traitement thermique à 100 °C. Après le traitement thermique à 140 °C, la forme du signal diélectrique change totalement, comme observé à la fréquence de 1 Hz. Le premier pic de relaxation détecté (symbole rond) est toujours présent mais un second pic de relaxation (symbole triangle) apparaît, toutefois ces deux processus de relaxation sont difficiles à caractériser. A 180 °C, aucun processus de relaxation n'est détecté.

A la fréquence de  $10^6$  Hz (Fig. 5-5c), un seul processus de relaxation est présent (symbole rond). La position du pic de relaxation en température change entre 80 °C et 100 °C de 54 °C à 67 °C. Après le traitement thermique à 140 °C, le premier pic de relaxation n'est plus détectable toutefois un deuxième pic de relaxation (symbole triangle) apparaît à 117 °C. Après 180 °C, le signal diélectrique est linéaire.

Dans ce confinement, le comportement protéique change principalement entre les traitements thermiques à 100 °C et 140 °C pour les fréquences étudiées. Ainsi, comme observé dans le chapitre 4, le confinement protège la protéine de la dénaturation thermique par rapport à la toxine à l'état macroscopique (127,128).

#### 5.1.4 Analyse du comportement de CtxB<sub>5</sub> en fonction de la fréquence

La dépendance en fréquence des différents traitements thermiques est représentée dans la figure 5-6 et dans la figure 5-4 pour le traitement thermique à 60 °C. Le premier processus de relaxation détecté (symbole rond) couvre toute la gamme de fréquence entre 1 Hz et  $10^6$  Hz jusqu'au traitement thermique à 100 °C. Puis uniquement les basses fréquences (1 Hz à  $10^3$  Hz) après le traitement à 140 °C. Le deuxième processus de relaxation détecté (symbole triangle) est présent à 1 Hz après le traitement thermique à 80 °C, dans les basses fréquences (1 Hz à  $10^2$  Hz) après le traitement thermique à 100 °C et dans les basses et hautes fréquences après 140 °C (1 Hz à  $10^6$  Hz). Un troisième processus de relaxation est présent dans les basses fréquences après le traitement thermique à 100 °C (1 Hz à  $10^2$  Hz).

Comme le montre la figure 2-7, des mouvements de structures secondaires de l'échelle de la microseconde à la milliseconde et des mouvements de repliement, dépliement et d'assemblage, dissociation de chaînes de l'ordre de la milliseconde à la seconde sont présents dans une protéine. Par conséquent, le premier pic de relaxation couvrant les basses et hautes fréquences dès le traitement thermique à 60 °C est lié à des mouvements de structures secondaires (fréquence supérieure à  $10^3$  Hz) mais aussi à des mouvements de dépliement dans les structures tertiaires, les domaines, les interfaces ou des mouvements de dissociation de chaînes. A partir de 140 °C, uniquement des mouvements de dépliement sont détectés. Le deuxième pic de relaxation détecté dans les basses fréquences correspond à des mouvements de dépliement après les traitements thermiques à 80 °C et 100 °C, cependant des mouvements de structures secondaires sont aussi détectés après le traitement thermique à 140 °C. Le troisième pic de relaxation détecté uniquement dans les basses fréquences après

le traitement thermique à 100 °C est associé à des mouvements de dépliement (ou de dissociation de chaîne).

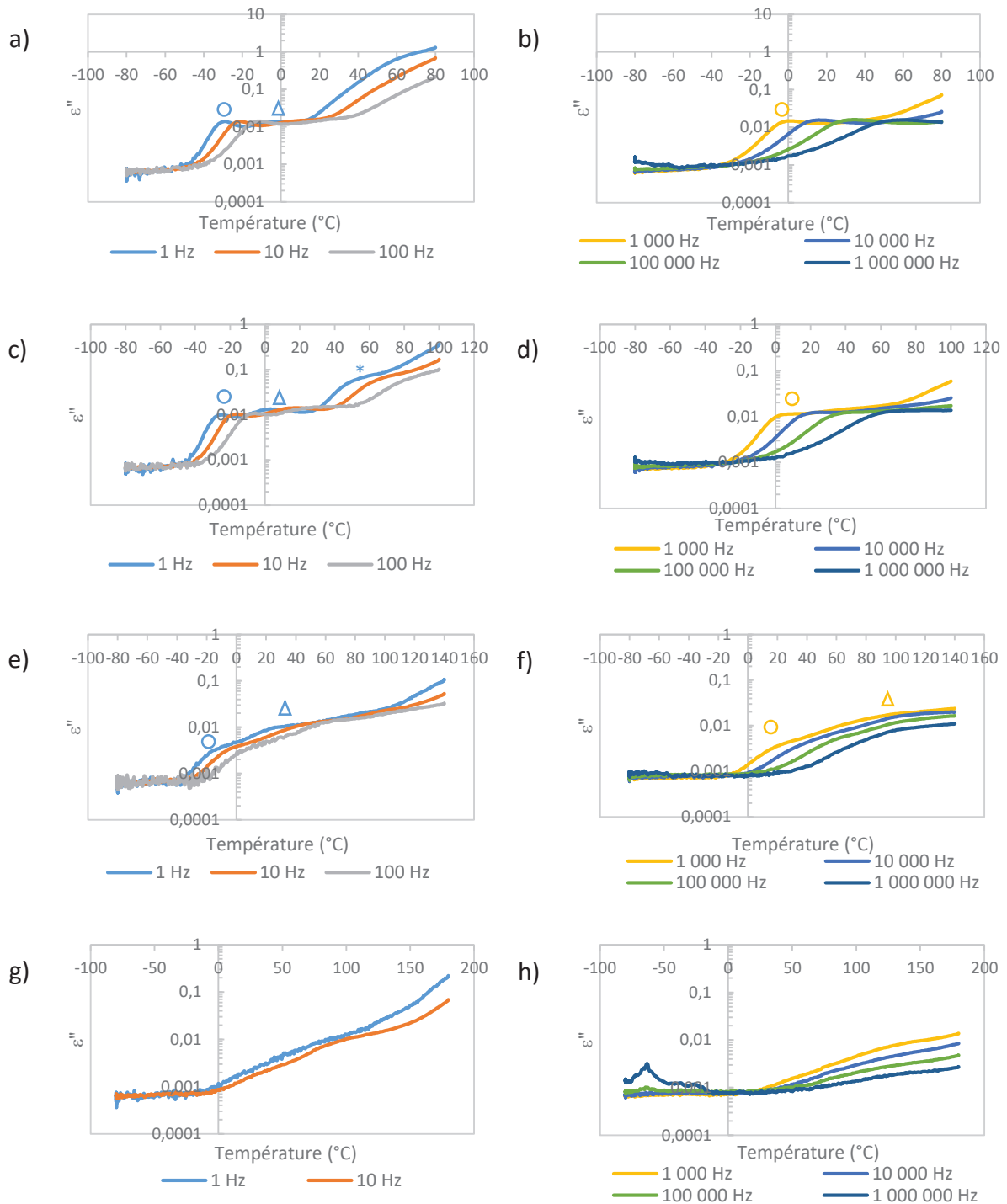
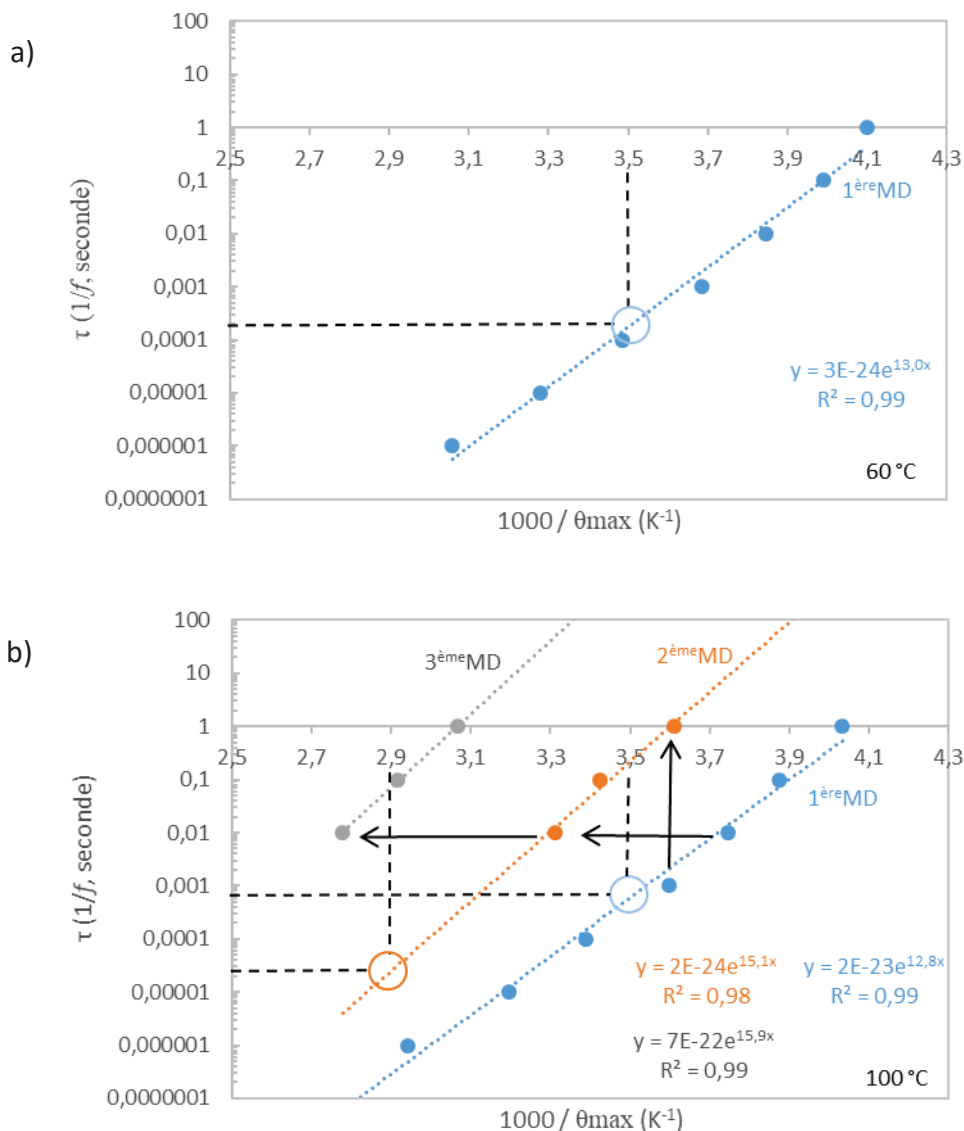


Figure 5-6. Analyse de la dépendance en fréquence des processus de relaxation pour le confinement à 40 nm de diamètre et 5 μm de long après différents traitements thermiques. a) Superposition des courbes de refroidissement de 1 Hz à 10<sup>2</sup> Hz après le traitement thermique à 80 °C (R 80 °C). b) Superposition des courbes R 80 °C de 10<sup>3</sup> Hz à 10<sup>6</sup> Hz. c) Superposition des courbes de refroidissement de 1 Hz à 10<sup>2</sup> Hz après le traitement thermique à 100 °C (R 100 °C). d) Superposition des courbes R 100 °C de 10<sup>3</sup> Hz à 10<sup>6</sup> Hz. e) Superposition des courbes de refroidissement de 1 Hz à 10<sup>2</sup> Hz après le traitement thermique à 140 °C (R 140 °C). f) Superposition des courbes R 140 °C de 10<sup>3</sup> Hz à 10<sup>6</sup> Hz. g) Superposition des courbes de refroidissement de 1 Hz à 10<sup>2</sup> Hz après le traitement thermique à 140 °C (R 140 °C). h) Superposition des courbes R 140 °C de 10<sup>3</sup> Hz à 10<sup>6</sup> Hz.



Superposition des courbes de refroidissement de 1 Hz à 10 Hz après le traitement thermique à 180 °C (R 180 °C). h) Superposition des courbes R 180 °C de 10<sup>3</sup> Hz à 10<sup>6</sup> Hz.

Les dépendances en température des temps de relaxation ( $\tau$ ) des trois processus de relaxation sont représentées dans la figure 5-7. La dépendance en température de la courbe de refroidissement après le traitement thermique à 80 °C n'est pas représentée, car les températures associées au premier pic de relaxation sont identiques à celles après le traitement thermique à 60 °C. De plus, après le traitement thermique à 80 °C, le deuxième processus de relaxation n'est détecté qu'à la fréquence de 1 Hz. Les temps de relaxation dépendent linéairement de la température pour les trois processus de relaxation.



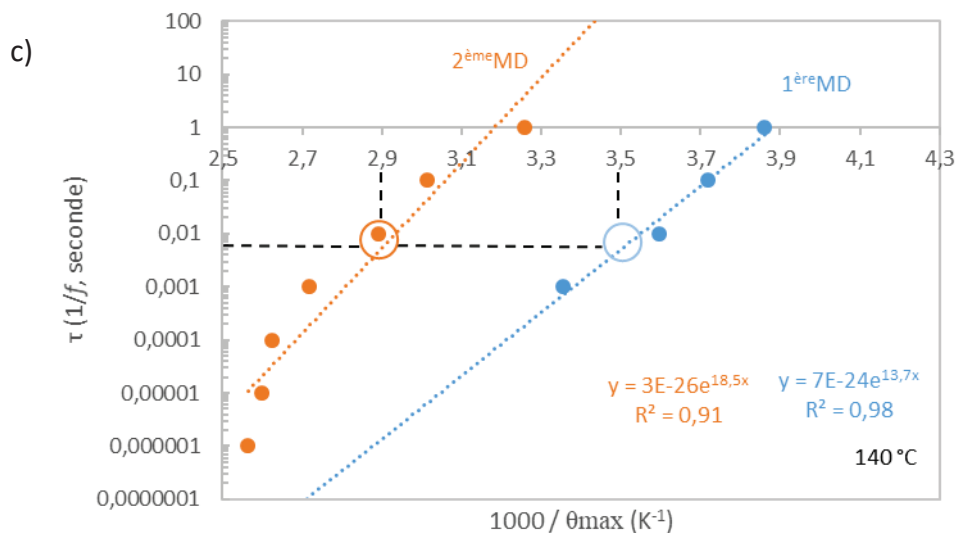


Figure 5-7. Dépendance en température des temps de relaxation dans le confinement à 40 nm de diamètre et 5  $\mu\text{m}$  de long après différents traitements thermiques. a) Dépendance en température du premier processus de relaxation après le traitement thermique à 60 °C. b) Dépendance en température des processus de relaxation après le traitement à 100 °C. c) Dépendance en température des processus de relaxation après le traitement thermique à 140 °C.

Les temps de relaxation en fonction de la position en température sont similaires (cercle) jusqu'au traitement thermique à 140 °C pour le premier processus de relaxation détecté. Ce qui n'est pas le cas pour le deuxième processus de relaxation détecté où les temps de relaxation deviennent plus lents entre 100 °C et 140 °C (Fig. 5-7).

Les temps de relaxation avec des positions des pics de relaxation en température similaires deviennent plus lents du premier processus de relaxation détecté au deuxième processus de relaxation détecté (flèche verticale). Les positions des pics de relaxation en température avec des temps de relaxation similaires augmentent du premier processus de relaxation détecté au deuxième processus de relaxation puis au troisième (flèches horizontales).

Le premier processus de relaxation détecté couvre une large gamme de fréquence (basses et hautes fréquences) jusqu'au traitement thermique à 100 °C. De plus, il s'agit du processus de relaxation avec les temps de relaxation les plus rapides, suggérant une population de conformation plus stable et plus repliée. Toutefois, au vu de l'échelle des temps de relaxation mesurée dans l'expérience (seconde à microseconde), le pentamère détecté ne correspond pas au pentamère natif, où uniquement des mouvements de chaînes latérales (nanoseconde) devraient être présents (3). Ce processus de relaxation est détecté dans les plus basses températures (dès le traitement thermique à 60 °C) et correspond donc à la conformation protéique la plus proche de l'état natif, pour ces raisons et en accord avec les données macroscopiques (127,128), nous supposons qu'il s'agit d'une population de pentamère de type natif.

Le troisième processus de relaxation détecté dans les basses fréquences possède les temps de relaxation les plus lents. Par conséquent, il s'agit des conformations protéiques les

moins repliées. Or dans le chapitre précédent, nous avons émis l'hypothèse que les monomères dépliés n'étaient pas détectés (124). Cette hypothèse est en accord avec l'absence de processus de relaxation après le traitement thermique à 180 °C dans ce confinement. En effet, si les monomères dépliés (état final de la protéine) étaient détectés, un processus de relaxation devrait toujours être présent quel que soit la température. Le troisième processus de relaxation détecté dans ce confinement correspond à une population d'intermédiaire d'assemblage et/ ou de monomères repliés.

Le deuxième processus de relaxation détecté possède des temps de relaxation intermédiaire par rapport aux deux autres processus. Ainsi, la population de conformation correspond à un état entre le pentamère de type natif et les intermédiaires d'assemblage. Il est donc raisonnable de supposer que ce processus de relaxation correspond à une population de pentamères non repliés ou partiellement repliés.

Si le même raisonnement que pour le chapitre précédent est suivi (Chapitre 4), alors le premier processus de relaxation détecté et associé à un pentamère de type natif correspond à MD<sub>1</sub>, le deuxième processus de relaxation associé à une population de pentamères partiellement ou non repliés est MD<sub>3</sub> et le troisième processus de relaxation correspondant aux intermédiaires d'assemblage et/ou aux monomères repliés est MD<sub>2</sub>.

Les positions des pics de relaxation en température dans ce confinement sont différentes de celles du chapitre précédent où la longueur du confinement est plus grande, suggérant un enrichissement et/ou une sélection conformationnelle différente si la longueur du confinement varie. Une analyse plus complète sera réalisée dans la troisième partie de ce chapitre. Maintenant un autre confinement avec une longueur de pore plus petite est étudié.

## 5.2 Confinement à 40 nm de diamètre et 2 µm de long

### 5.2.1 Contrôle de l'échantillon protéique

La membrane vide et le contrôle du solvant ne détecte pas de processus de relaxation alors que l'échantillon protéique CtxB<sub>5</sub> possède deux processus de relaxation dans le confinement à 40 nm de diamètre et 2 µm de long à la fréquence de 10<sup>3</sup> Hz après le traitement thermique à 60 °C. Par conséquent, les pics de relaxation avec des positions en température à -8 °C et 45 °C sont caractéristiques de la protéine (Fig. 5-8).

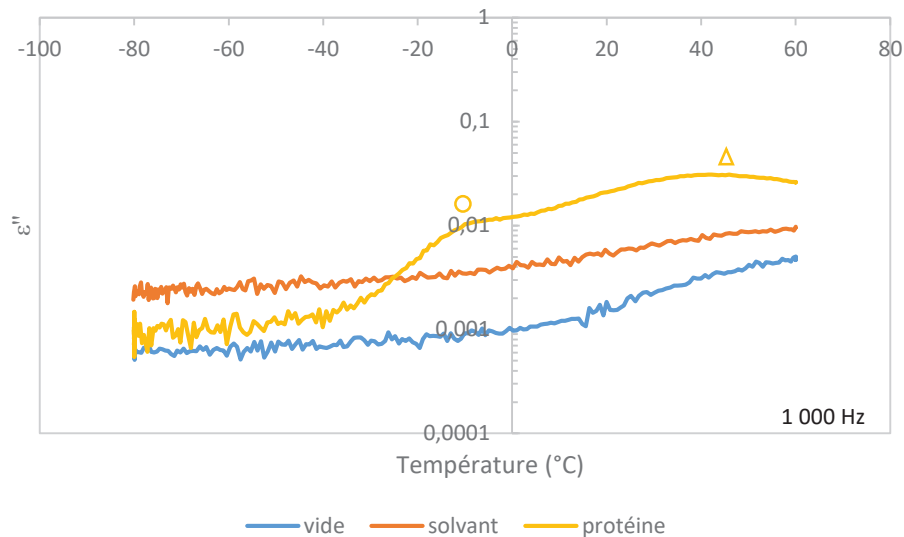


Figure 5-8. CtxB<sub>5</sub> dans le confinement à 40 nm de diamètre et 2 μm de long à la fréquence de 10<sup>3</sup> Hz. Superposition des courbes de refroidissement après le deuxième traitement thermique à 60 °C. La courbe jaune correspond à l'échantillon protéique où 200 μL de solution est déposée sur la nanomembrane. La solution protéique est composée de 2 μL de CtxB<sub>5</sub> à une concentration de 2,6 mg/ml et de 1 016 μL d'eau déionisée. La concentration protéique finale est de 0,0051 mg/ml. La courbe orange correspond à l'échantillon avec le solvant composé de 2 μL de solution de PBS 1X et de 1 016 μL d'eau déionisée. La courbe bleue correspond à la membrane vide.

### 5.2.2 Analyse de CtxB<sub>5</sub> à différentes fréquences

La position des processus de relaxation en température varie en fonction de la fréquence (Fig. 3-12). Par conséquent, les données de pertes diélectriques en fonction de la température sont étudiées à différentes fréquences. La température des pics de relaxation augmente avec l'augmentation de la fréquence (Fig. 5-9).

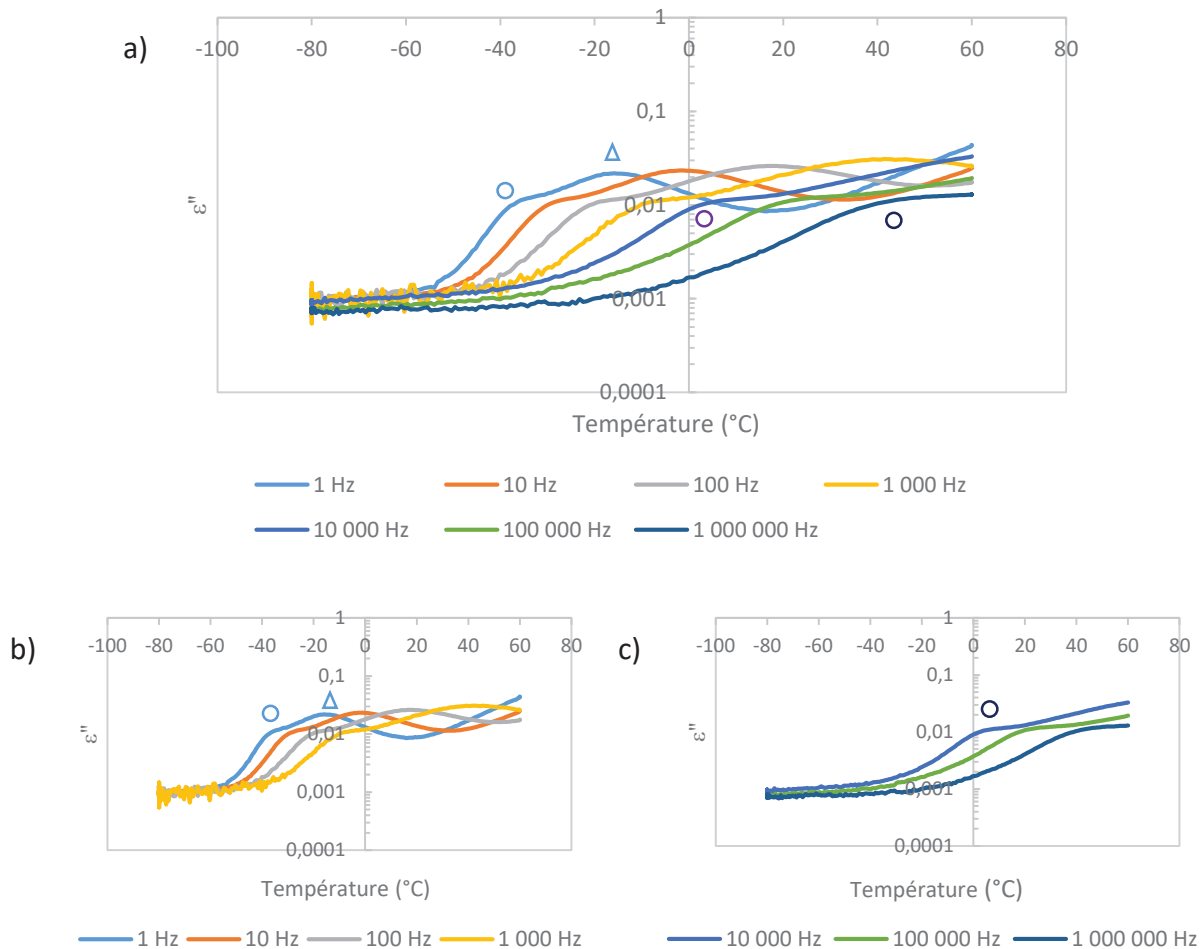


Figure 5-9. Analyse de CtxB<sub>5</sub> à différentes fréquences. a) Superposition des courbes de refroidissement après le deuxième traitement thermique à 60 °C (R2 60 °C). b) Zoom sur les processus de relaxation de 1 Hz à 10<sup>3</sup> Hz. c) Zoom sur le premier processus de relaxation de 10<sup>4</sup> Hz à 10<sup>6</sup> Hz. Le symbole rond représente le premier processus de relaxation détecté alors que le symbole triangle représente le deuxième.

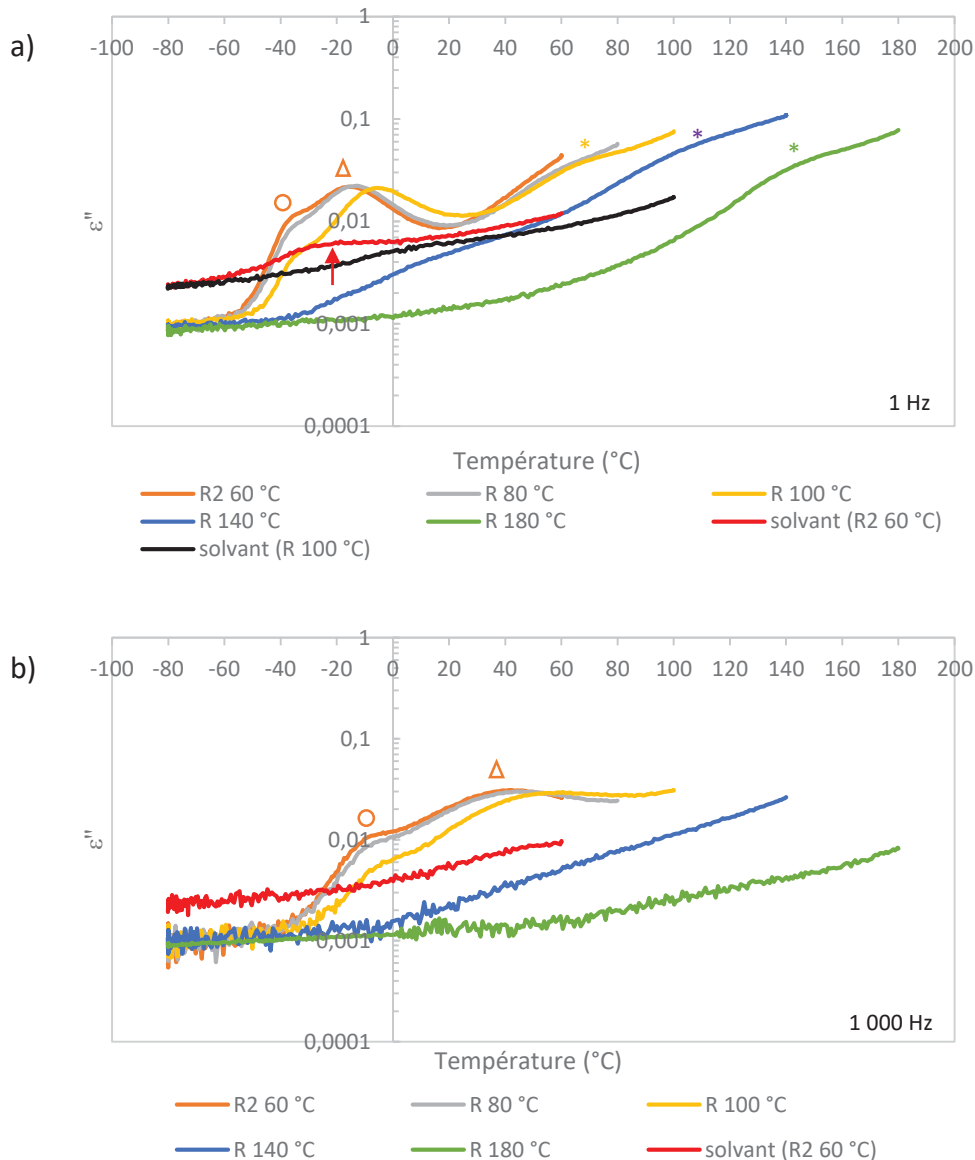
CtxB<sub>5</sub> possède deux processus de relaxation de 1 Hz à 10<sup>3</sup> Hz (Fig. 5-9a et Fig. 5-9b), alors qu'un seul processus de relaxation est détecté de 10<sup>4</sup> Hz à 10<sup>6</sup> Hz (Fig. 5-9a et Fig. 5-9c). Le premier processus de relaxation (symbole rond) possède un pic de relaxation à -33 °C à la fréquence de 1 Hz ; à -8 °C à la fréquence de 10<sup>3</sup> Hz ; 27 °C à 10<sup>5</sup> Hz et à 50 °C à 10<sup>6</sup> Hz. Le deuxième processus de relaxation détecté de 1 Hz à 10<sup>3</sup> Hz possède un pic de relaxation à -14 °C à la fréquence de 1 Hz et à 45 °C à la fréquence de 10<sup>3</sup> Hz.

Le premier processus de relaxation possède un décalage de 25 °C de 1 Hz à 10<sup>3</sup> Hz (-33 °C à -8 °C), alors que le deuxième processus de relaxation est décalé de 59 °C entre ces deux fréquences (-14 °C à 45 °C). Par conséquent, les deux processus de relaxation détectés dès le traitement thermique à 60 °C ne possèdent pas les mêmes environnements dipolaires.

Une analyse du comportement de la toxine en fonction de la température dans ce confinement est réalisée pour les fréquences 1 Hz, 10<sup>3</sup> Hz et 10<sup>5</sup> Hz (Fig. 5-10) pour étudier la dénaturation thermique de CtxB<sub>5</sub> dans ce confinement.

5.2.3 Analyse du comportement de CtxB<sub>5</sub> en fonction de la température

Pour le solvant, aucun processus de relaxation n'est détecté à la fréquence de 10<sup>5</sup> Hz dès le premier traitement thermique à 60 °C, le signal du solvant disparaît à partir du deuxième traitement thermique à 60 °C pour la fréquence de 10<sup>3</sup> Hz et à partir de 100 °C à la fréquence de 1 Hz. A la fréquence de 1 Hz, le signal diélectrique du solvant est différent du signal diélectrique de la protéine après le deuxième traitement thermique à 60 °C (Fig. 5-10a). De plus, pour la protéine un ou plusieurs processus de relaxation sont détectés jusqu'au traitement thermique à 100 °C quels que soient les fréquences étudiées confirmant que le signal diélectrique de l'échantillon protéique caractérise bien les conformations protéiques.



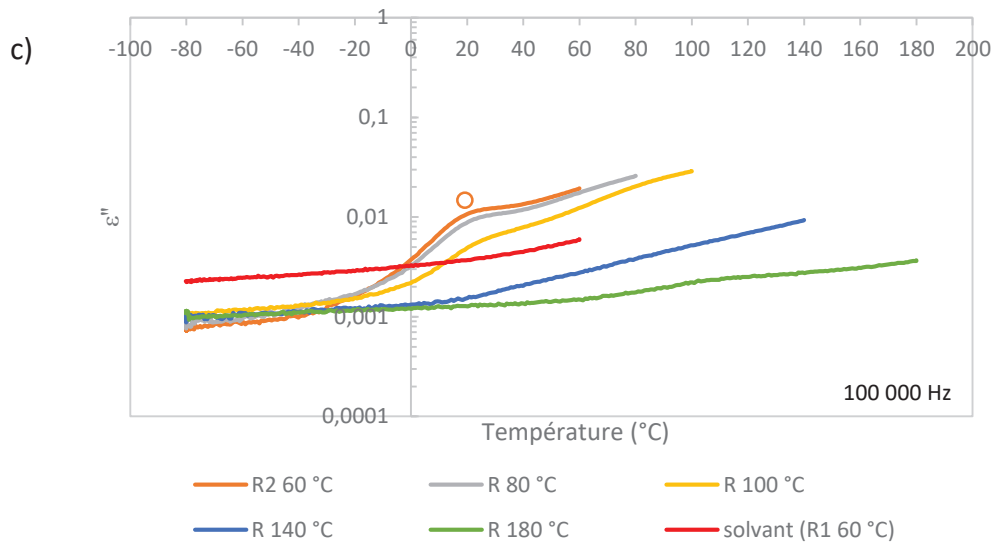


Figure 5-10. Etude du comportement de CtxB<sub>5</sub> après différents traitements thermiques. a) Analyse de CtxB<sub>5</sub> et du solvant à la fréquence de 1 Hz. b) Analyse de CtxB<sub>5</sub> et du solvant à 10<sup>3</sup> Hz. c) Analyse de CtxB<sub>5</sub> et du solvant à 10<sup>5</sup> Hz. La concentration protéique de l'échantillon est de 0,0051 mg/ml.

A la fréquence de 1 Hz (Fig. 5-10a), les deux processus de relaxation sont détectés jusqu'au traitement thermique à 80 °C sans changement de position du pic de relaxation en température. A partir du traitement thermique à 100 °C, la position du premier processus de relaxation en température détecté ne change pas (-33 °C). En revanche, l'intensité du signal diminue d'un facteur de 2 par rapport au traitement thermique à 60 °C. Après le traitement thermique à 140 °C ce processus de relaxation n'est plus détecté. Ces résultats tendent à suggérer que même si la conformation protéique ne change pas avec l'augmentation de la température, cette population conformationnelle se dépeuple quand la température des traitements thermiques augmente, jusqu'à disparaître ou à être en quantité insuffisante pour pouvoir être détectée. La position du deuxième processus de relaxation (symbole triangle) détecté passe de -14 °C à -5 °C, ce qui suggère un changement de conformation entre 80 °C et 100 °C, même si l'intensité des pertes diélectriques ne change pas. Par conséquent, même si un changement de conformation se produit entre 80 °C et 100 °C, suite à une déstabilisation plus importante des interfaces, le nombre de protéines associées à cette population de conformation ne change pas.

A partir du traitement thermique à 100 °C, un troisième processus de relaxation (symbole étoile) apparaît. Ce troisième processus de relaxation n'est pas assez défini pour pouvoir déterminer avec précision la position de ce pic de relaxation en température. Ce processus de relaxation semble avoir majoritairement disparu après le traitement thermique à 140 °C. Néanmoins, le signal est récupéré après le traitement thermique à 180 °C, même si comme à 100 °C, la position de ce processus de relaxation en température n'est pas assez bien définie pour pouvoir être déterminée avec précision.

A la fréquence de 10<sup>3</sup> Hz (Fig. 5-10b), deux processus de relaxation sont détectés jusqu'au traitement thermique à 100 °C. La position du pic de relaxation en température du premier processus de relaxation ne change pas (-8°C) avec l'augmentation de la température des traitements thermiques, toutefois, comme à la fréquence de 1 Hz, l'intensité des données

du signal diminue. Après le traitement thermique à 100 °C, le deuxième processus de relaxation détecté passe de 45 °C (après le traitement thermique à 80 °C) à 56 °C. Après le traitement à 140 °C aucun processus de relaxation n'est détecté.

A la fréquence de  $10^5$  Hz (Fig. 5-10c), un seul processus de relaxation est détecté jusqu'au traitement thermique à 100 °C. Comme précédemment, la position du pic de relaxation en température ne change pas et reste à 27 °C à travers les différents traitements thermiques même si l'intensité des données de pertes diélectriques diminue. De plus, comme à la fréquence de  $10^3$  Hz, aucun processus de relaxation n'est détecté à partir du traitement thermique à 140 °C.

Pour les différentes fréquences étudiées, la position du premier processus de relaxation en température est identique après le traitement thermique à 60 °C jusqu'au traitement thermique à 100 °C. Ces résultats suggèrent que dans ce confinement les conformations protéiques associées à ce processus de relaxation sont stables jusqu'à 100 °C. Le comportement de la toxine est toujours différent de l'état macroscopique, où la dénaturation de CtxB<sub>5</sub> se produit à 74 °C (127,128). Par conséquent, le confinement protège CtxB<sub>5</sub> de la dénaturation thermique.

Dans le confinement à 40 nm de diamètre et 2 µm de long, la position des pics de relaxation en température est différente de la position des processus de relaxation détectés dans les deux autres confinements, suggérant que la variation de la longueur du confinement permet bien d'enrichir et/ou sélectionner différentes conformations protéiques. Une analyse plus détaillée de l'effet de la longueur du confinement sera abordée dans la troisième partie de ce chapitre.

#### 5.2.4 Analyse du comportement de CtxB<sub>5</sub> en fonction de la fréquence

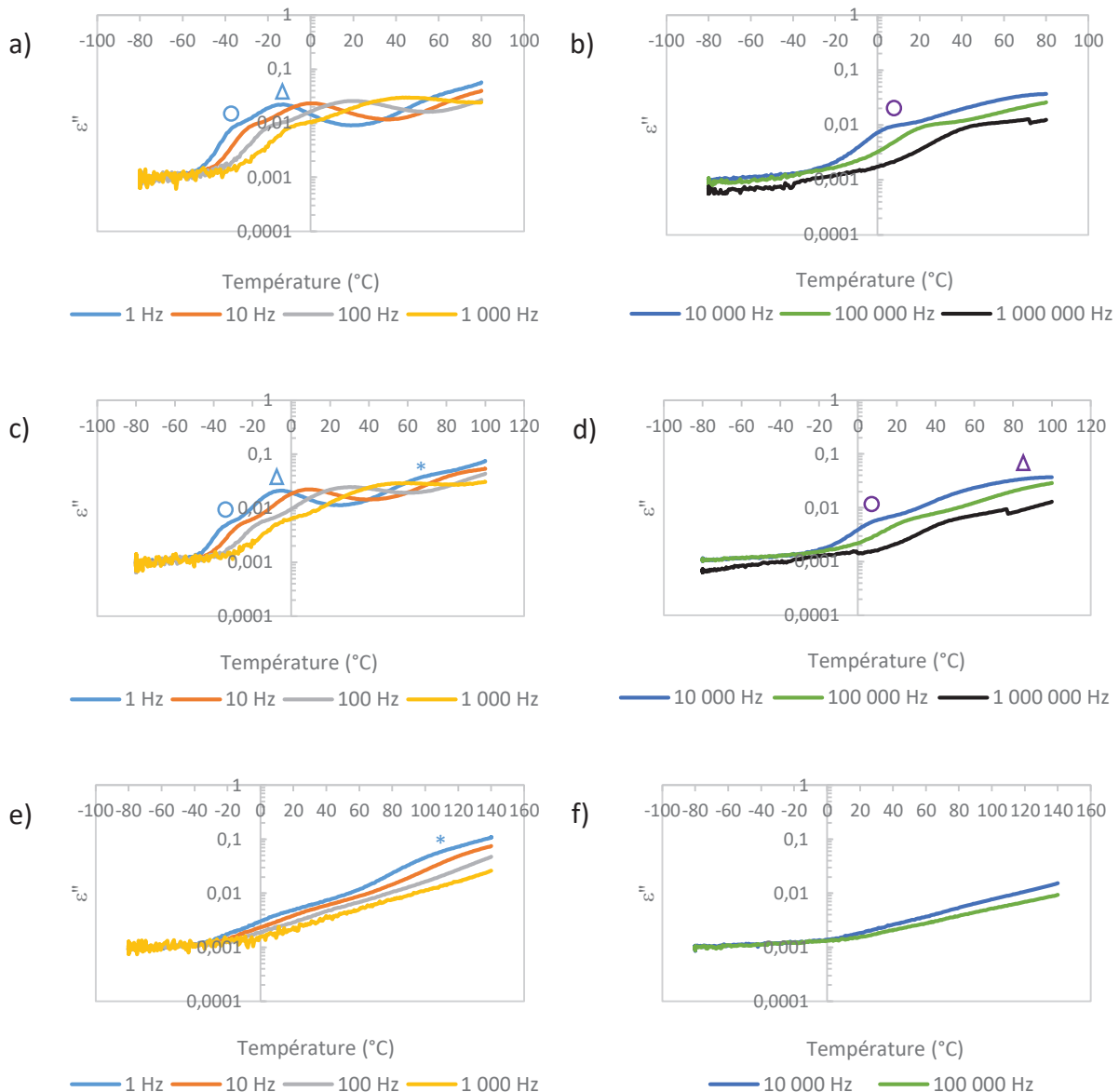
La dépendance en fréquence des processus de relaxation est étudiée en fonction des différents traitements thermiques dans la figure 5-11 et dans la figure 5-9 pour le deuxième traitement thermique à 60 °C. Le premier processus de relaxation détecté (symbole rond) couvre toute la gamme de fréquence de 1 Hz à  $10^6$  Hz pour les traitements thermiques jusqu'à 100 °C. Le deuxième processus de relaxation est détecté jusqu'à la fréquence de  $10^3$  Hz pour les traitements thermiques à 60 °C et 80 °C et jusqu'à  $10^4$  Hz pour le traitement thermique à 100 °C. Un troisième processus de relaxation commence à apparaître dans les basses fréquences (1 Hz et 10 Hz) après le traitement thermique à 100 °C et à 1 Hz après les traitements thermiques à 140 °C et 180 °C.

D'après la figure 2-7, la protéine possède des mouvements de chaînes latérales à l'échelle de la nanoseconde, des mouvements plus lents de plus grandes échelles (structures secondaires) et des mouvements encore plus lents sur des plus grandes surfaces (dépliement, repliement de domaines, interfaces, structures tertiaires et assemblage, dissociation de ces structures tertiaires). Le premier processus de relaxation détecté dans les basses et hautes fréquences dès le traitement thermique à 60 °C est lié à des mouvements de structures secondaires (fréquence supérieure à  $10^3$  Hz) mais aussi à des mouvements de dépliement de



chaînes, domaines, interfaces et/ou des mouvements de dissociation de chaînes. Le deuxième processus de relaxation détecté principalement dans les basses fréquences correspond à des mouvements de dépliement, comme le troisième processus de relaxation détecté après le traitement thermique à 100 °C.

A partir du traitement thermique à 140 °C, la fréquence à 10<sup>6</sup> Hz n'est pas représentée et analysée car les courbes de refroidissement deviennent trop bruitées. Toutefois, la position des pics de relaxation en température augmente avec l'augmentation de la fréquence. Par conséquent, il est raisonnable de supposer que la non détection de ces processus de relaxation à la fréquence de 10<sup>5</sup> Hz pour les traitements thermiques à 140 °C et 180 °C, sous-entend que ces processus de relaxation ne seraient pas détectés à la fréquence de 10<sup>6</sup> Hz pour les mêmes traitements thermiques.



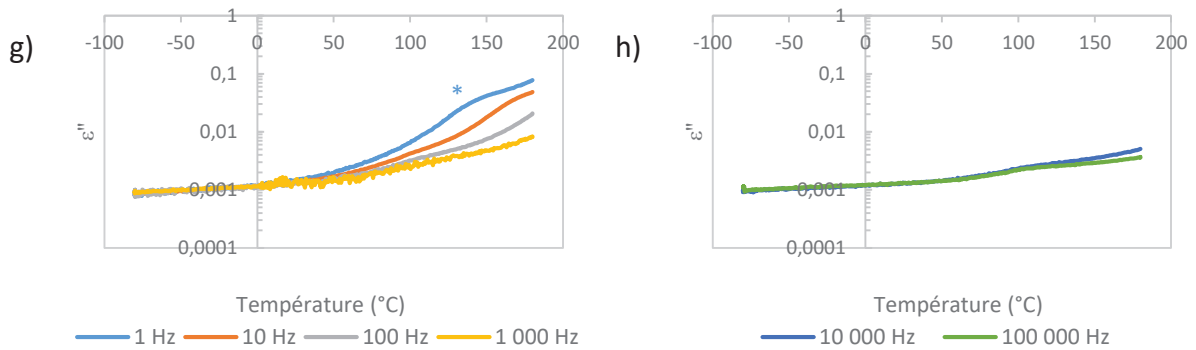


Figure 5-11. Analyse de la dépendance en fréquence des processus de relaxation dans le confinement à 40 nm de diamètre et 2  $\mu\text{m}$  de long après différents traitements thermiques. a) Superposition des courbes de refroidissement de 1 Hz à  $10^3$  Hz après le traitement thermique à 80 °C (R 80 °C). b) Superposition des courbes R 80 °C de  $10^4$  Hz à  $10^6$  Hz. c) Superposition des courbes R 100 °C de 1 Hz à  $10^3$  Hz. d) Superposition des courbes R 100 °C de  $10^4$  Hz à  $10^6$  Hz. e) Superposition des courbes R 140 °C de 1 Hz à  $10^3$  Hz. f) Superposition des courbes R 140 °C de  $10^4$  Hz à  $10^5$  Hz. g) Superposition des courbes R 180 °C de 1 Hz à  $10^3$  Hz. h) Superposition des courbes R 180 °C de  $10^4$  Hz à  $10^5$  Hz.

Les dépendances en température des temps de relaxation ( $\tau$ ) des trois processus de relaxation sont représentées dans la figure 5-12. Les temps de relaxation après le traitement thermique à 80 °C ne sont pas représentés car les positions des pics de relaxation en température sont identiques à celles du traitement thermique à 60 °C pour toutes les fréquences. Le troisième processus de relaxation est faiblement détecté à 1 Hz et 10 Hz après le traitement thermique à 100 °C dans notre gamme de fréquence, par conséquent, la dépendance en température des temps de relaxation de ce processus sera obligatoirement linéaire.

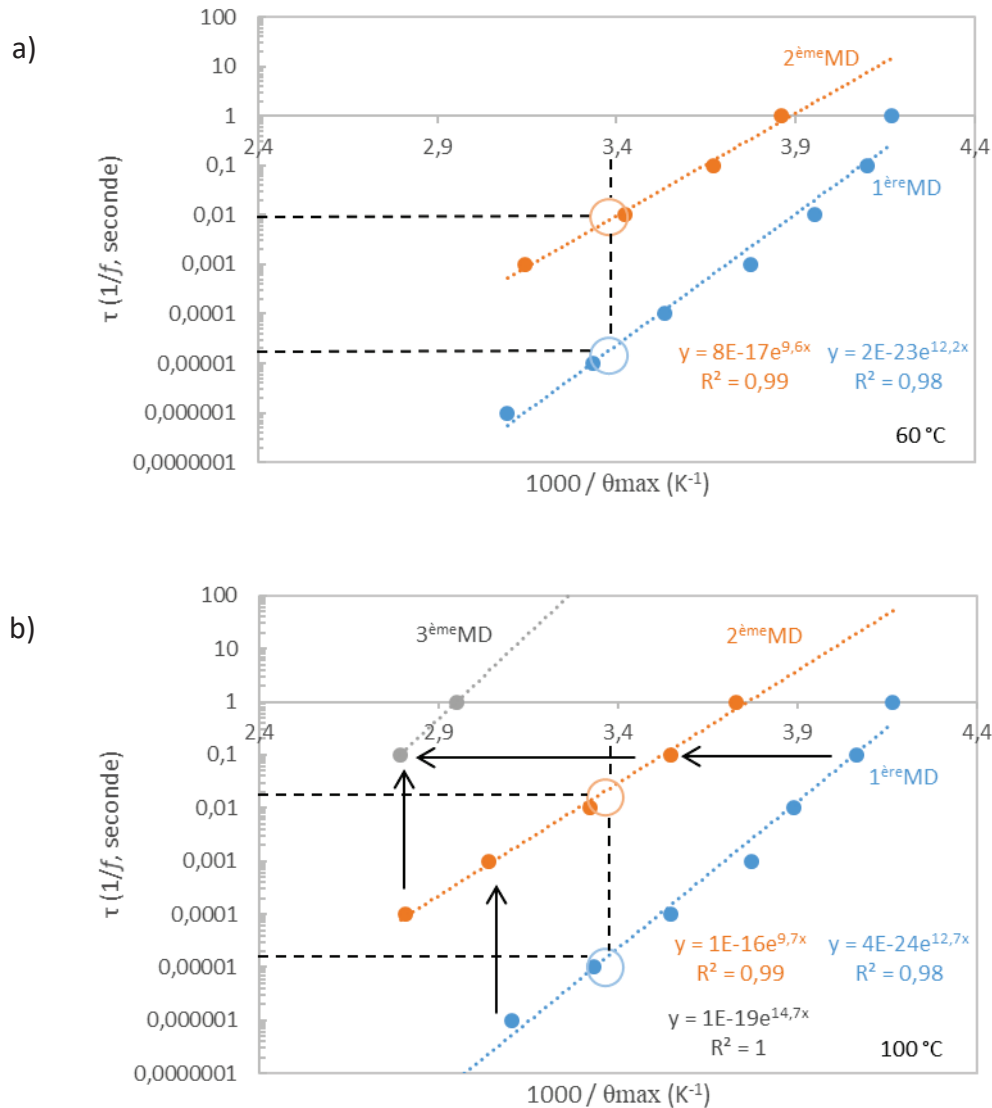


Figure 5-12. Dépendance en température des temps de relaxation dans le confinement à 40 nm de diamètre et 2 μm de long après différents traitements thermiques. a) Dépendance en température des processus de relaxation après le traitement thermique à 60 °C. b) Dépendance en température des processus de relaxation après le traitement thermique à 100 °C.

Les temps de relaxation en fonction de la température sont similaires (cercles) jusqu’au traitement thermique à 100 °C pour le premier et le deuxième processus de relaxation détectés. A une position en température du pic de relaxation similaire, les relaxations deviennent plus lentes du premier au deuxième puis au troisième processus de relaxation détecté (flèches verticales). Les positions des pics de relaxation en température avec des temps de relaxation similaires augmentent du premier au deuxième puis au troisième processus de relaxation (flèches horizontales).

Le premier processus de relaxation détecté possède un pic dans toute la gamme de fréquence étudiée. Ce processus de relaxation possède les temps de relaxation les plus rapides, par conséquent, il s’agit de la population de conformation la plus stable et la plus repliée détectée diélectriquement. De plus, précédent l’hypothèse que le confinement

protégeait la protéine vis-à-vis de la dénaturation thermique a été émise, pour ces raisons nous supposons que ce processus de relaxation correspond à un pentamère de type natif (pentamère replié).

Le deuxième processus de relaxation est détecté dans les mêmes plages de température (après le traitement thermique à 60 °C jusqu'au traitement thermique à 100 °C) que le premier processus de relaxation dans les basses fréquences. Les temps de relaxation de ce processus sont plus lents que ceux du premier, par conséquent, l'environnement des dipôles est moins compact dans ce processus de relaxation permettant ainsi des mouvements à plus grande échelle. Nous supposons donc qu'il s'agit d'un état plus déplié, soit un pentamère partiellement replié.

Le troisième processus de relaxation détecté dans les basses fréquences à hautes températures possède les temps de relaxation les plus lents, par conséquent il s'agit de la population de conformation la plus dépliée. Ces résultats tendent à suggérer que ce processus de relaxation détecte les conformations d'intermédiaires d'assemblage et/ou de monomères repliés.

Si le même raisonnement que pour les deux autres confinements est suivi, alors le premier processus de relaxation détecté et associé à l'état pentamérique de type natif correspond à MD<sub>1</sub>, le deuxième processus de relaxation détecté correspondant à l'état pentamérique partiellement déplié soit à MD<sub>3</sub> et le troisième processus de relaxation détecté associé aux intermédiaires d'assemblage et aux monomères repliés représente MD<sub>2</sub>.

Dans la troisième partie, les différents confinements seront comparés pour déterminer si la longueur du confinement permet d'enrichir et/ou sélectionner différentes populations conformationnelles.

### 5.3 Comparaison

Toutes les expériences de CtxB<sub>5</sub> sont réalisées dans des conditions nanoconfinées et faiblement hydratées. Par conséquent, le nanoconfinement, la déshydratation et/ou la combinaison des deux pourraient impacter les conformations protéiques détectées (124). Tous les confinements comparés dans ce chapitre possèdent un diamètre de confinement de 40 nm de diamètre mais une longueur de confinement différente.

#### 5.3.1 Effet de la déshydratation

Pour rappel, le pourcentage de remplissage volumique des nanopores est de 14,3 % pour tous les confinements, par conséquent, la proportion d'occupation des protéines est la même dans les différents confinements. L'échantillon protéique est toujours soumis 15 minutes à 50 °C et 5 minutes à 30 °C, avant de réaliser les différents traitements thermiques. Après les traitements thermiques, un cycle comprenant une courbe de refroidissement et une courbe de chauffage est réalisé (Chapitre 3).

Pour deux traitements thermiques à 60 °C, le signal diélectrique est reproductible pour tous les confinements : même position du pic du ou des processus de relaxation en

température, même intensité et même largeur de ces pics de relaxation. De plus, aucune différence significative n'est détectée à l'intérieur d'un cycle de température (Fig. 5-13).

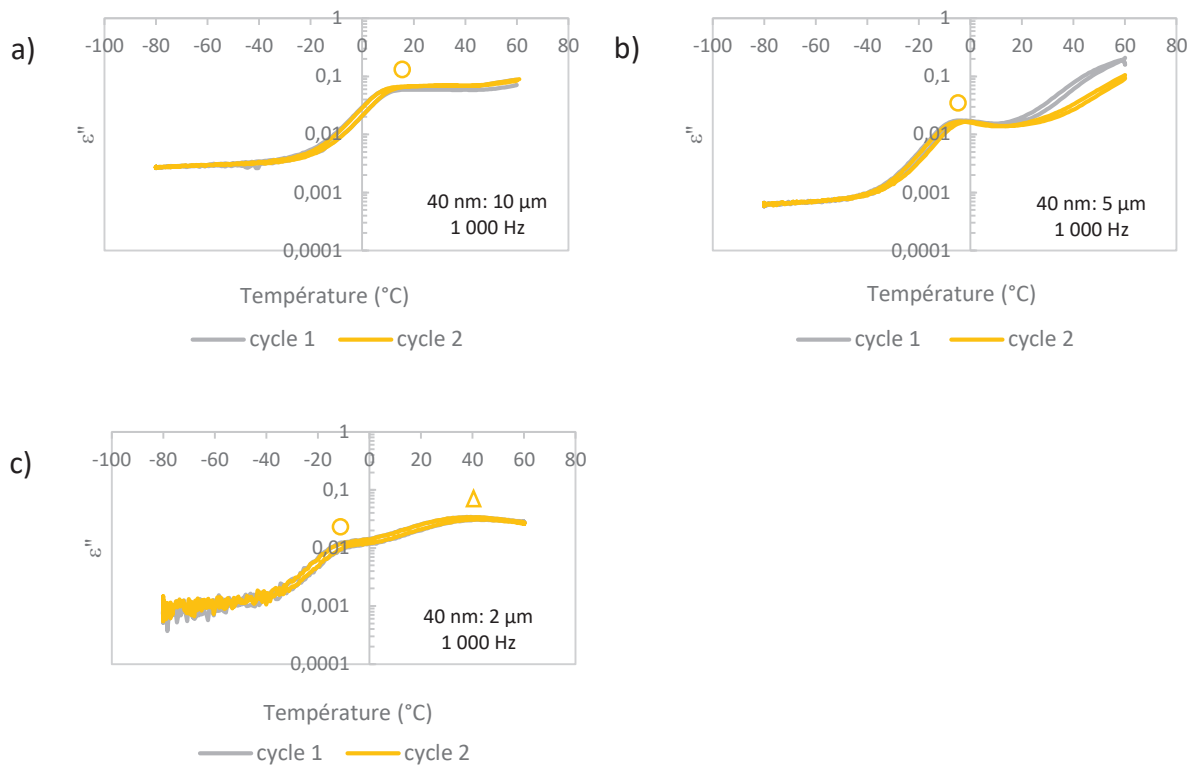


Figure 5-13. Superposition des courbes de refroidissement et de chauffage à la fréquence de  $10^3$  Hz après deux traitements thermiques à  $60$  °C. a) Confinement à  $40$  nm de diamètre et  $10$   $\mu$ m de long. b) Confinement à  $40$  nm de diamètre et  $5$   $\mu$ m de long. c) Confinement à  $40$  nm de diamètre et  $2$   $\mu$ m de long.

L'absence d'hystérésis et la reproduction du signal diélectrique dans un cycle de température et si le même cycle de température est reproduit suggère que l'eau libre est totalement évaporée après trois heures à  $60$  °C dans tous les confinements étudiés. Ainsi, l'échantillon protéique mesuré est composé de protéines, eau liée et PBS. De plus, la population de conformations protéiques détectée dès le traitement thermique à  $60$  °C est stable dans les différents confinements étudiés. Par conséquent, l'état de déshydratation est le même pour les trois confinements.

Pour ces raisons, il est raisonnable de supposer que les conformations non natives détectées suite au dépliement thermique de CtxB<sub>5</sub> proviennent uniquement du confinement et non de la déshydratation de l'échantillon protéique.

### 5.3.2 Effet de la variation de longueur du confinement

La comparaison des signaux diélectriques obtenus en faisant varier la longueur du confinement est représentée dans la figure 5-14. Pour rappel, MD<sub>1</sub> (symbole rond) est associé à une population de conformation de pentamère de type natif, où les dipôles de l'interface N-terminale (interface I<sub>1</sub>) sont déstabilisés. MD<sub>3</sub> (symbole triangle) est associé à une population de pentamère partiellement replié où l'interface principale (interface I<sub>2</sub>) entre des brins beta de monomère adjacent est déstabilisée, MD<sub>2</sub> (symbole étoile) correspond aux intermédiaires

d'assemblage et/ou aux monomères repliés où les dipôles de l'interface de l'hélice alpha centrale (interface I<sub>3</sub>) sont perturbés (124). Dans nos conditions expérimentales (1 Hz à 10<sup>6</sup> Hz) aucun pentamère natif ne devrait être détecté. En effet, à l'état natif uniquement des fluctuations à l'échelle de la nanoseconde (10<sup>9</sup> Hz) associés aux chaînes latérales des acides aminés et aux liaisons peptidiques devraient être présentes (3,124).

Le pH des solutions pour le confinement de 40 nm de diamètre et 10 µm de long, et de 40 nm de diamètre et 5 µm de long est de 6,9, et de 6,8 dans le confinement de 40 nm de diamètre et 2 µm de long. Par conséquent, les mêmes acides aminés (lysine, arginine, acide aspartique et acide glutamique) et les deux extrémités N-terminale et C-terminale sont chargés dans tous les confinements testés.

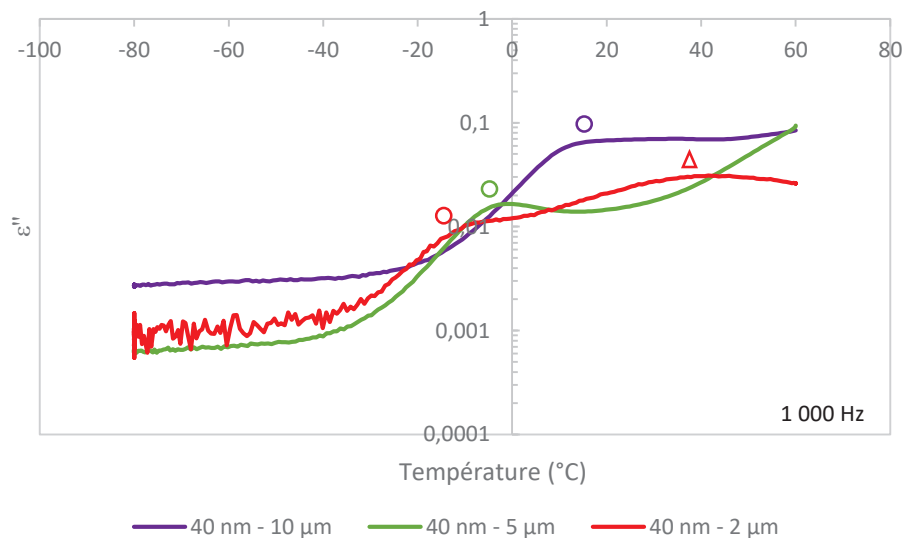


Figure 5-14. Comparaison des processus de relaxation mesurés à la fréquence de 10<sup>3</sup> Hz en faisant varier la longueur du confinement après le deuxième traitement thermique à 60 °C. Superpositions des courbes R2 60 °C pour le confinement à 10 µm de long (courbe violette), 5 µm de long (courbe verte) et 2 µm de long (courbe rouge). Le symbole rond correspond à la population conformationnelle la plus stable dans le confinement.

La température du pic de relaxation associé à la conformation protéique la plus stable (MD<sub>1</sub>) augmente en augmentant la longueur du confinement (Fig. 5-14). Ce processus de relaxation (symbole rond) se situe à une température de -8 °C pour le confinement de 2 µm de long, à -1,5 °C pour le confinement à 5 µm de long et à 13 °C pour le confinement de 10 µm de long. Ces résultats indiquent que les populations conformationnelles associées au processus de relaxation MD<sub>1</sub> sont différentes dans les trois confinements. De plus, la figure 5-15 montre que MD<sub>1</sub> n'est plus détecté après le traitement thermique à 140 °C pour le confinement de 2 µm de long et après celui à 180 °C pour le confinement de 5 µm, alors que ce processus de relaxation est toujours détecté après le traitement thermique à 180 °C pour le confinement de 10 µm de long, suggérant que les différentes conformations enrichies et/ou sélectionnées via le confinement possèdent une stabilité thermique différente.

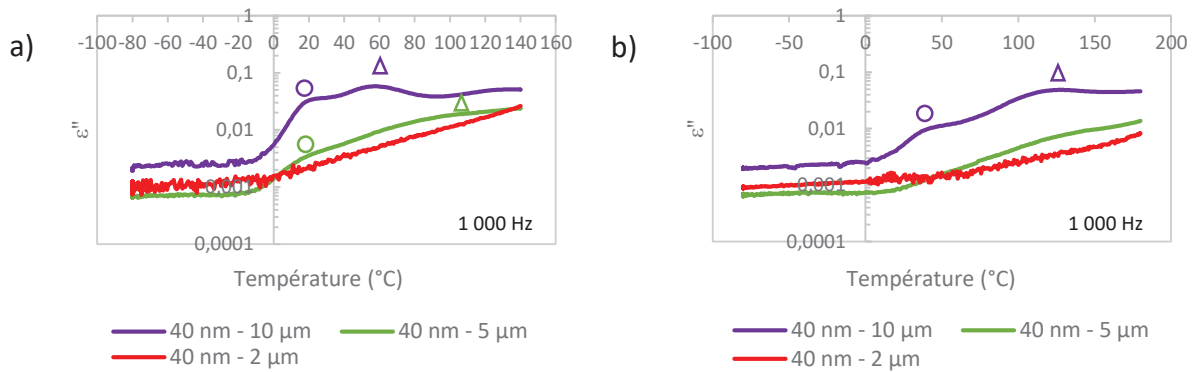


Figure 5-15. Comparaison des processus de relaxation détectés pour différentes longueurs de confinement après les traitements thermiques à 140 °C et 180 °C. a) Représentation des courbes de refroidissement à la fréquence de  $10^3$  Hz après le traitement thermique à 140 °C (R 140 °C). b) Représentation des courbes de refroidissement à la fréquence de  $10^3$  Hz après le traitement thermique à 180 °C (R 180 °C).

Dans la table 5-1, les intensités et les positions des pics de relaxation en température ( $\theta_{max}$ ) du processus de relaxation  $MD_1$  à la fréquence de  $10^3$  Hz sont représentées pour différents confinements en faisant varier la longueur des nanopores.

Cycle		40 nm – 10 $\mu$ m	40 nm – 5 $\mu$ m	40 nm – 2 $\mu$ m
60 °C	$\theta_{max}$ (°C)	13	-1,5	-8
	Intensité	0,059	0,016	0,010
80 °C	$\theta_{max}$ (°C)	13	-1,5	-8
	Intensité	0,055	0,013	0,0080
100 °C	$\theta_{max}$ (°C)	20	5	-8
	Intensité	0,055	0,010	0,0042
140 °C	$\theta_{max}$ (°C)	25	25	X
	Intensité	0,034	0,0035	
180 °C	$\theta_{max}$ (°C)	42	X	X
	Intensité	0,008		

Table 5-1. Intensité et position des pics de relaxation en température associée au processus de relaxation  $MD_1$  à la fréquence de  $10^3$  Hz après différents traitements thermiques pour différents confinements en variant la longueur.

Les intensités du processus de relaxation  $MD_1$  sont plus importantes pour le confinement de 10  $\mu$ m de long que pour les deux autres confinements (5  $\mu$ m et 2  $\mu$ m de long) après le traitement thermique à 60 °C. Par conséquent, la population de conformation pentamérique de type natif, plus précisément la densité volumique de dipôles détectés associés à ces conformations est plus importante dans le confinement de 10  $\mu$ m long que dans les deux autres confinements. Ce résultat est plutôt surprenant, en effet l'intensité est proportionnelle à la densité de dipôle. Or comme le pourcentage de remplissage volumique est le même (14,3 %) dans tous les confinements, les intensités auraient dû être similaire. Comme ce n'est pas le cas, cela suggère un problème d'écoulement dans les nanopores, plus précisément les confinements de 2  $\mu$ m et 5  $\mu$ m de long seraient plus encombrés que celui de 10  $\mu$ m pour un même diamètre.

La largeur du pic de relaxation à la fréquence de  $10^3$  Hz après le traitement thermique à  $60\text{ °C}$  est de  $32\text{ °C}$  pour le confinement de  $2\text{ }\mu\text{m}$  de long, de  $36,5\text{ °C}$  pour le confinement de  $5\text{ }\mu\text{m}$  de long et de  $35\text{ °C}$  pour le confinement de  $10\text{ }\mu\text{m}$  de long. Par conséquent, aucun confinement ne possède une population conformationnelle plus hétérogène que les autres.

En conclusion, la position du pic de relaxation en température associé à une population de pentamère de type natif ( $\text{MD}_1$ ) change en variant la longueur du confinement. Par conséquent, la variation de longueur du confinement permet de cribler différentes conformations protéiques. La variation de la longueur du confinement n'impacte pas l'hétérogénéité de la population conformationnelle dans le nanopore toutefois, elle impacte l'encombrement macromoléculaire.

A la fréquence de  $10^3$  Hz, le processus de relaxation  $\text{MD}_3$  (pentamère partiellement replié) est détecté dès le traitement thermique à  $60\text{ °C}$  pour le confinement de  $2\text{ }\mu\text{m}$  de long et dès celui à  $140\text{ °C}$  pour le confinement de  $5\text{ }\mu\text{m}$  de long. Pour le confinement de  $10\text{ }\mu\text{m}$  de long, même si le pic de relaxation commence à se dessiner dès le traitement thermique à  $80\text{ °C}$ , celui-ci est bien résolu à partir du traitement thermique à  $140\text{ °C}$  (Table. 5-2).

Cycle		40 nm – 10 $\mu\text{m}$	40 nm – 5 $\mu\text{m}$	40 nm – 2 $\mu\text{m}$
60 °C	$\theta_{\text{max}}$ (°C)	X	X	45
	Intensité			0,018
80 °C	$\theta_{\text{max}}$ (°C)	35	X	45
	Intensité	0,010		0,018
100 °C	$\theta_{\text{max}}$ (°C)	41	X	56
	Intensité	0,012		0,023
140 °C	$\theta_{\text{max}}$ (°C)	59	106	X
	Intensité	0,019	0,013	
180 °C	$\theta_{\text{max}}$ (°C)	132	X	X
	Intensité	0,036		

Table 5-2. Intensité et position des pics de relaxation en température associée au processus de relaxation  $\text{MD}_3$  à la fréquence de  $10^3$  Hz après différents traitements thermiques pour différents confinements en variant la longueur.

$\text{MD}_3$  est présent dès le traitement thermique à  $60\text{ °C}$  pour le confinement de  $2\text{ }\mu\text{m}$  de long (Fig. 5-14 et table 5-2) comme le processus de relaxation  $\text{MD}_1$ . Précédemment, nous avons émis l'hypothèse que les populations conformationnelles enrichies et/ou sélectionnées via la longueur du confinement possédaient des stabilités thermiques différentes. Par conséquent, une partie de la population de conformation pentamérique repliée ( $\text{MD}_1$ ) présente dès  $60\text{ °C}$  pourrait subir une déstabilisation de l'interface  $I_2$  pendant les trois heures à  $60\text{ °C}$  conduisant à l'obtention d'une population conformationnelle où le pentamère se déplie ( $\text{MD}_3$ ). De plus, l'intensité du pic de relaxation  $\text{MD}_3$  ne varie pas en augmentant la température des traitements thermiques, alors que l'intensité de  $\text{MD}_1$  diminue de moitié après le traitement thermique à  $100\text{ °C}$ , suggérant que  $\text{MD}_1$  n'est pas uniquement transformée en  $\text{MD}_3$ .



L'apparition de MD<sub>3</sub> à la fréquence de 10<sup>3</sup> Hz après le traitement thermique à 140 °C (Fig. 5-15a, table 5-2, Fig. 5-5b) pour le confinement à 5 µm de long coïncide avec la diminution d'intensité de MD<sub>1</sub> à la même température. En effet, entre les traitements thermiques à 100 °C et 140 °C, l'intensité du pic de relaxation MD<sub>1</sub> est divisée par un facteur supérieur à 2 (Table. 5-1). Ce résultat tend à confirmer qu'une partie de la population conformationnelle de pentamère de type natif est dénaturé thermiquement (déstabilisation de l'interface I<sub>2</sub>) pour donner une population de pentamère partiellement replié (MD<sub>3</sub>). La population conformationnelle associée à MD<sub>3</sub> ne serait pas présente initialement dans ce confinement en raison de la stabilité plus élevée de la population conformationnelle enrichie et/ou sélectionnée dans ce confinement.

Pour rappel, après le traitement thermique à 140 °C, la position du pic de relaxation MD<sub>1</sub> en température est aussi décalée de 20 °C. Par conséquent, il est raisonnable de supposer qu'en plus de la création d'une population de pentamère partiellement replié associé à MD<sub>3</sub>, une partie de la population conformationnelle initiale (pentamère de type natif replié) possède une interface I<sub>1</sub> encore plus déstabilisée (sans perturbation des deux autres interfaces) qu'après le traitement thermique à 100 °C, conduisant à un décalage de la position du pic de relaxation en température associé à MD<sub>1</sub>.

De même, dans le confinement de 10 µm de long, le pic de relaxation MD<sub>3</sub> est bien défini après le traitement thermique à 140 °C, ce qui coïncide avec la diminution d'intensité de MD<sub>1</sub>, confirmant qu'une partie de la population conformationnelle MD<sub>1</sub> est déstabilisée et forme la population conformationnelle MD<sub>3</sub>. Toutefois, le processus de relaxation MD<sub>3</sub> commence à être détecté dans ce confinement dès le traitement thermique à 80 °C sans que l'intensité du processus de relaxation MD<sub>1</sub> ne varie (Table. 5-1). Ce résultat suggère la présence d'un seuil de détection maximale, l'intensité du processus de relaxation ne diminue pas même si une partie de la population MD<sub>1</sub> est dépliée en MD<sub>3</sub> jusqu'à ce qu'une valeur inférieure à celle du seuil soit atteinte. Néanmoins, il est nécessaire d'être prudent car la présence d'un seuil maximum de détection en BDS n'est corroborée par aucune étude.

Le processus de relaxation associé à MD<sub>2</sub> n'est pas détecté à la fréquence 10<sup>3</sup> Hz pour les confinements de 2 µm et 5 µm de long car ce processus de relaxation est associé à une population conformationnelle plus dépliée. L'analyse de MD<sub>2</sub> sera réalisée à une fréquence plus faible, à 1 Hz. A la fréquence de 1 Hz, MD<sub>2</sub> est détecté à partir du traitement thermique à 100 °C pour le confinement de 2 µm et 5 µm de long et dès le traitement thermique à 60 °C pour le confinement de 10 µm de long (Fig. 5-10a, Fig. 5-5a, Fig. 4-10a).

Dans le confinement de 2 µm de long, MD<sub>2</sub> (faiblement détecté) commence à apparaître après le traitement thermique à 100 °C, quand l'intensité de MD<sub>1</sub> diminue. Ce processus de relaxation est aussi présent pour les traitements thermiques supérieures à 100 °C, alors que MD<sub>1</sub> et MD<sub>3</sub> ne sont plus détectés, suggérant que la population de conformations d'intermédiaires d'assemblage et/ou de monomères repliés résulte du dépliement des pentamères repliés et potentiellement des pentamères partiellement dépliés. Pour le confinement de 5 µm de long, MD<sub>2</sub> est détecté sans que les intensités de MD<sub>1</sub> et MD<sub>3</sub> varient (Table. 5-3), ce résultat tend à confirmer la présence d'un seuil de détection maximum. En

effet, si la quantité de populations conformationnelles de MD<sub>1</sub> et MD<sub>3</sub> dépassent ce seuil, aucun changement d'intensité ne sera détecté pour ces deux processus de relaxation, même si MD<sub>2</sub> est issu du dépliement de MD<sub>1</sub> et MD<sub>3</sub>. La détection de MD<sub>2</sub> dès 60 °C pour le confinement de 10 µm de long suggère que les populations conformationnelles enrichies et/ou sélectionnées via la longueur du confinement possèdent des stabilités thermiques différentes mais aussi des cinétiques différentes.

Cycle		MD <sub>1</sub>	MD <sub>2</sub>	MD <sub>3</sub>
60 °C	θ <sub>max</sub> (°C)	-29	X	X
	Intensité	0,015		
80 °C	θ <sub>max</sub> (°C)	-29	X	-2
	Intensité	0,014		0,004
100 °C	θ <sub>max</sub> (°C)	-25	53	4
	Intensité	0,010	0,050	0,004
140 °C	θ <sub>max</sub> (°C)	-14	X	34
	Intensité	0,0032		0,006
180 °C	θ <sub>max</sub> (°C)	X	X	X
	Intensité			

Table 5-3. Intensité et position des pics de relaxation en température des trois processus de relaxation détectés dans le confinement à 40 nm de diamètre et 5 µm de long à la fréquence de 1 Hz.

En conclusion, la température provoque la déstabilisation de l'interface I<sub>1</sub> et la formation de pentamère de type natif associé au premier processus de relaxation détecté (MD<sub>1</sub>). Si la température des traitements thermiques augmente, l'interface I<sub>1</sub> des pentamères de type natif devient plus déstabilisée entraînant la 'formation' de nouveaux pentamères de type natif (variation de θ<sub>max</sub> pour MD<sub>1</sub>). Toutefois, l'augmentation de la température des traitements thermiques entraîne aussi la déstabilisation de l'interface I<sub>2</sub> des pentamères de type natif formant ainsi des pentamères partiellement repliés (MD<sub>3</sub>). Les intermédiaires d'assemblage et/ou les monomères repliés sont issus de la déstabilisation des interfaces (I<sub>1</sub>, I<sub>2</sub> et I<sub>3</sub>) dans les pentamères de type natif et potentiellement dans les pentamères partiellement repliés.

Si l'expérience de CtxB<sub>5</sub> dans le confinement de 2 µm de long est reproduite, la position du pic de relaxation MD<sub>1</sub> en température (symbole rond) est identique, ce qui n'est pas le cas pour le pic de relaxation MD<sub>3</sub> (symbole triangle) (Fig. 5-16a). Ces résultats suggèrent que les populations de conformations de pentamères repliés sont des conformations protéiques stables, ce qui n'est pas le cas des populations de pentamères partiellement repliés issus de la perturbation de l'interface I<sub>2</sub> des pentamères repliés. Pour le confinement de 5 µm de long, le pic associé à MD<sub>1</sub> est détecté dans les deux expériences (Fig. 5-16b). Toutefois, la forme spectrale et la position du pic de relaxation MD<sub>1</sub> en température diffère entre les deux expériences (-1,5 °C pour le test 1 et 3 °C pour le test 2), traduisant un enrichissement et/ou une sélection conformationnelle différent. Cet enrichissement et/ou sélection différent peut-être causé par les dimensions du confinement ou par l'échantillon protéique. L'échantillon protéique de départ pourrait avoir les mêmes conformations protéiques, mais l'enrichissement et/ou la sélection de ces conformations seraient différents à cause du

confinement. A contrario, même si aucune différence de fluorescence n'est observée dans l'échantillon protéique entre les deux mesures diélectriques, les conformations de CtxB<sub>5</sub> pourraient être légèrement différentes, entraînant un enrichissement et/ou une sélection différent lors de la deuxième expérience. Pour rappel (Chapitre 4), lors de la reproduction de l'expérience dans le confinement de 10  $\mu\text{m}$  de long, la position du pic de relaxation MD<sub>1</sub> en température est identique (Fig. 4-6). La reproduction de la position du pic de relaxation MD<sub>1</sub> en température dans les confinements de 2  $\mu\text{m}$  et 10  $\mu\text{m}$  de long tend à suggérer que la différence d'enrichissement et/ou de sélection lors de la reproduction de l'expérience dans le confinement de 5  $\mu\text{m}$  de long provient de l'échantillon protéique de départ et non du confinement.

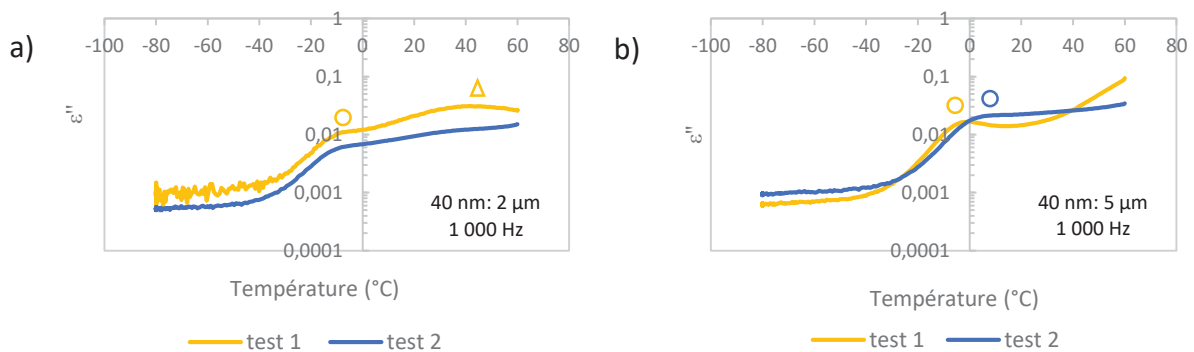


Figure 5-16. Reproduction des expériences des confinements de longueurs différentes à la fréquence de  $10^3$  Hz. a) Expérience dans le confinement à 40 nm de diamètre et 2  $\mu\text{m}$  de long. b) Expérience dans le confinement à 40 nm de diamètre et 5  $\mu\text{m}$  de long.

En 2004, Ravindra et al (149) ont étudié le comportement de la ribonucléase A (RNase A) dans une silice mésoporeuse possédant des pores. Ils ont observé que la RNase A pouvait être dans trois états différents : adsorbée sur la surface externe, encapsulée dans les pores, dans un état libre. Les auteurs ont associé la forme libre à une fraction de protéine en excès qui n'aurait aucune restriction de mobilité, par conséquent, la protéine aurait un comportement similaire à la protéine 'libre'. Leur étude a montré que les protéines encapsulées possédant des restrictions de configuration avaient une stabilité plus importante vis-à-vis du dépliement induit par la température que la protéine dans le solvant en solution. La protéine encapsulée possède une température de dépliement à 90 °C contre une température de dépliement de 62 °C pour la protéine libre et de 52 °C pour la protéine adsorbée en surface.

Les résultats de la combinaison du nanoconfinement et de la spectroscopie diélectrique à large bande sont cohérents avec la détection de protéines encapsulées. En effet, le processus de relaxation correspondant à la conformation protéique la plus stable MD<sub>1</sub> (proche de l'état natif) est détecté jusqu'au traitement thermique à 100 °C au minimum pour tous les confinements étudiés, alors qu'à l'état macroscopique le dépliement du pentamère CtxB<sub>5</sub> s'effectue à 74 °C (127,128).

# Chapitre 6

## Impact du confinement : Variation du diamètre

Les dimensions de nanopores peuvent être modifiées pour changer le confinement via deux paramètres : la longueur et/ou le diamètre. Le chapitre précédent (Chapitre 5) a permis de déterminer que la variation de la longueur du confinement permettait de cribler différentes conformations protéiques. Le chapitre 6 discutera de l'impact du diamètre. Les confinements à 120 nm de diamètre et 2  $\mu\text{m}$  de long et à 20 nm de diamètre et 2  $\mu\text{m}$  de long seront analysés. Puis ces deux confinements seront comparés avec le confinement à 40 nm de diamètre et 2  $\mu\text{m}$  de long (Chapitre 5) pour déterminer si la variation du diamètre permet aussi de cribler différentes conformations protéiques.

Comme pour le chapitre 5, les dimensions du confinement sont vérifiées en utilisant la microscopie électronique à balayage. La fluorescence est utilisée pour contrôler l'état pentamérique de la protéine CtxB<sub>5</sub> avant les mesures de BDS. L'échantillon protéique (composé de CtxB<sub>5</sub>, d'eau et de PBS) est déposé sur la nanomembrane avec le confinement choisi via la technique de la goutte pour un remplissage volumique des nanopores à 14,3 %. L'échantillon protéique est ensuite soumis à différents traitements thermiques de trois heures pour des températures allant de 60 °C à 180 °C. A la fin de chaque traitement thermique, les données de pertes diélectriques sont prélevées jusqu'à la fin des trente minutes à -80 °C et représentées dans les courbes de refroidissement (exemple : R 100 °C). Les courbes de refroidissement seront étudiées en fonction de la température à différentes fréquences. Le ou les pics de relaxation détectés dans les pertes diélectriques en fonction de la température à une fréquence donnée correspondent à la dynamique moléculaire.

### 6.1 Confinement à 120 nm de diamètre et 2 $\mu\text{m}$ de long

#### 6.1.1 Contrôle de l'échantillon protéique

Pour le confinement à 120 nm de diamètre et 2  $\mu\text{m}$  de long, un processus de relaxation commence à être détecté à la fréquence de  $10^3$  Hz pour l'échantillon protéique (Fig. 6-1a). Il est nécessaire d'étudier des fréquences plus basses pour détecter un pic de relaxation (symbole rond) dans la courbe de refroidissement après le traitement thermique à 60 °C (Fig. 6-1b). A 1 Hz, le solvant possède aussi un pic de relaxation à -6 °C (symbole flèche) après le traitement thermique à 60 °C et ce processus de relaxation n'est pas détecté après le traitement thermique à 80 °C (Fig. 6-1b). La membrane vide ne possède aucun signal de perte

diélectrique (pic de relaxation) aux fréquences étudiées. Ces résultats indiquent que le pic de relaxation de l'échantillon protéique est caractéristique des protéines.

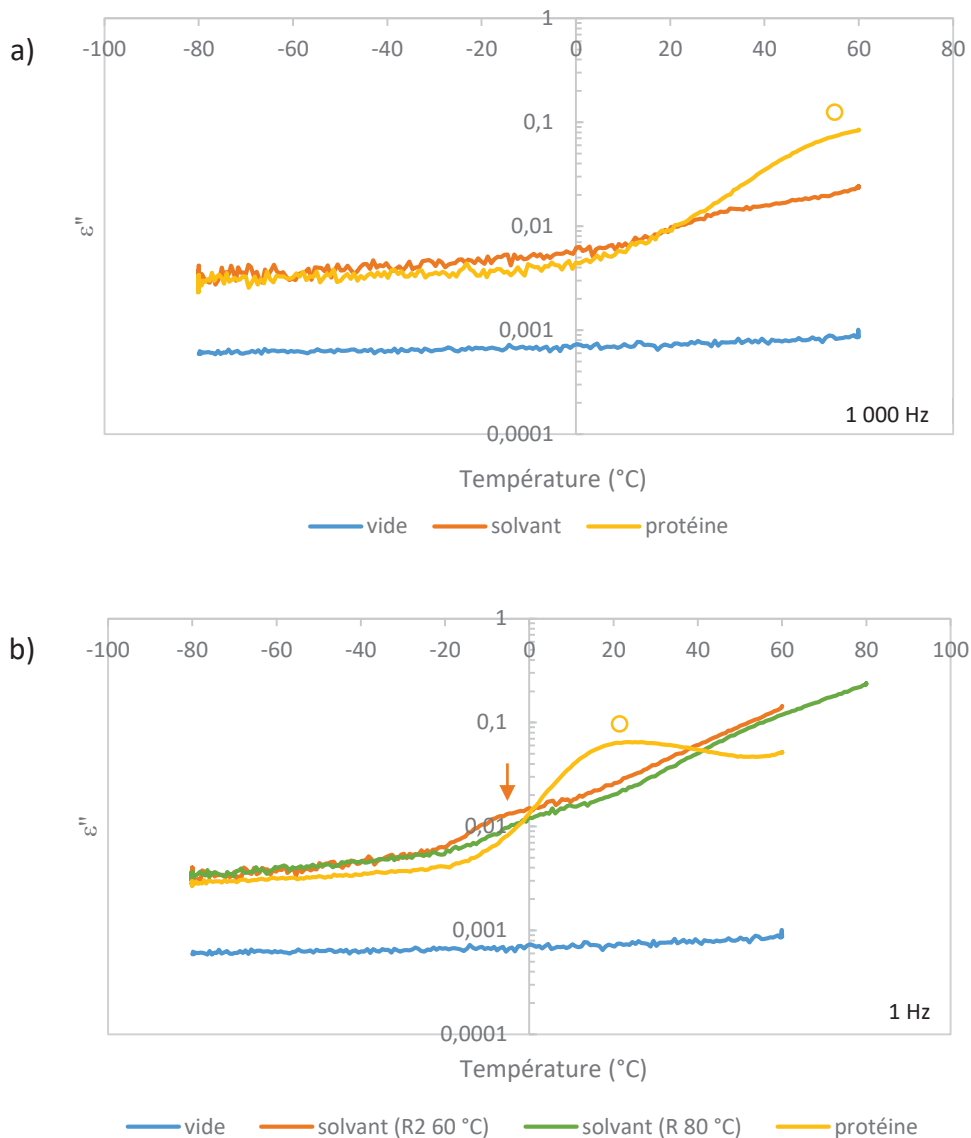


Figure 6-1. CtxB<sub>5</sub> dans le confinement à 120 nm de diamètre et 2 μm de long. a) Superposition des courbes de refroidissement à la fréquence de 10<sup>3</sup> Hz après le deuxième traitement thermique à 60 °C. b) Superposition des courbes de refroidissement à la fréquence de 1 Hz après le deuxième traitement thermique à 60 °C et 80 °C pour le solvant et l'échantillon protéique. Les courbes jaunes correspondent à l'échantillon protéique où 200 μL de solution protéique est déposée sur la nanomembrane. La solution protéique est composée de 3 μL de CtxB<sub>5</sub> à une concentration de 0,55 mg/ml et de 243 μL d'eau déionisée. La concentration protéique finale est de 0,0067 mg/ml. La courbe orange correspond à l'échantillon avec le solvant après le traitement thermique à 60 °C et la courbe verte à l'échantillon du solvant après le traitement thermique à 80 °C. L'échantillon du solvant est composé de 3 μL de PBS (PBS 1X) et de 243 μL d'eau déionisée. La courbe bleue correspond à la membrane vide.

### 6.1.2 Analyse de CtxB<sub>5</sub> à différentes fréquences

Le signal diélectrique varie en fonction de la température mais aussi en fonction de la fréquence (Fig. 3-12). Pour ces raisons, la perte diélectrique est étudiée en fonction de la température à différentes fréquences. La température des pics de relaxation augmente avec

l'augmentation de la fréquence (Fig. 6-2). Pour les fréquences de 1 Hz à  $10^2$  Hz, un seul pic de relaxation (symbole rond) est présent. Le pic apparaît à une température de 21 °C à la fréquence de 1 Hz, à 32 °C pour 10 Hz et à 47 °C pour  $10^2$  Hz. A  $10^3$  Hz, ce pic de relaxation est toujours détectable mais il n'est pas suffisamment résolu pour déterminer la température où l'intensité du pic est maximale. A partir de  $10^4$  Hz, aucun processus de relaxation n'est détecté dans notre gamme de mesure. A  $10^6$  Hz, le signal diélectrique montre une intensité supérieure à celle des autres fréquences, ce qui est spécifique pour une résistance de contact non-négligeable avec les électrodes de mesure conduisant à une augmentation de pertes diélectriques à hautes fréquences. Pour cette raison, les courbes de refroidissement à la fréquence de  $10^6$  Hz, ne seront pas analysées.

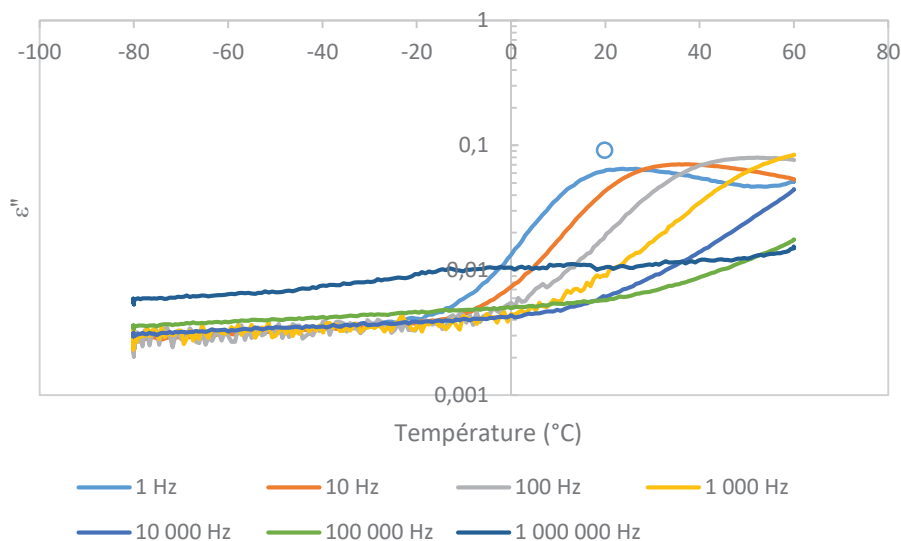
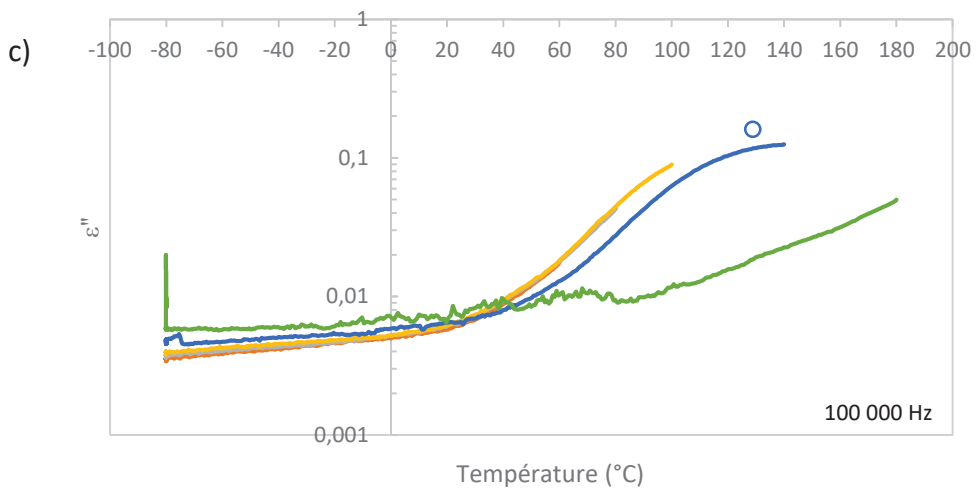
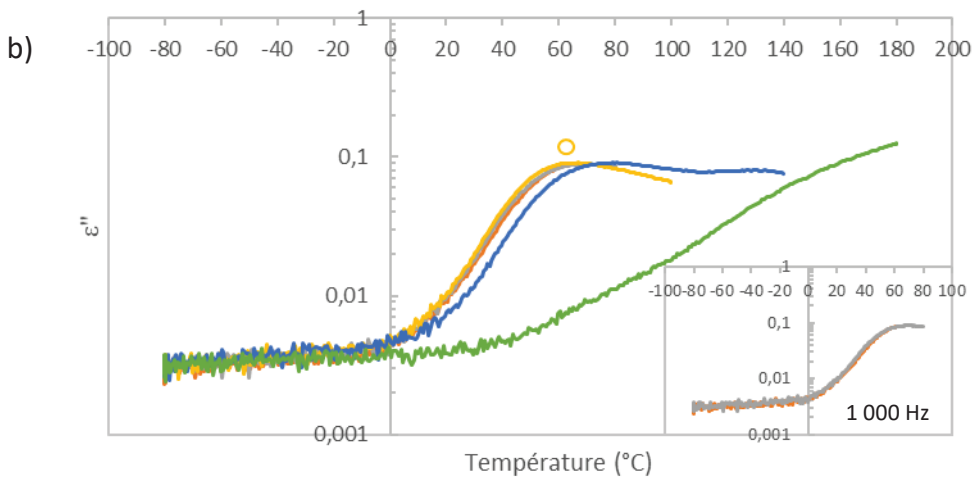
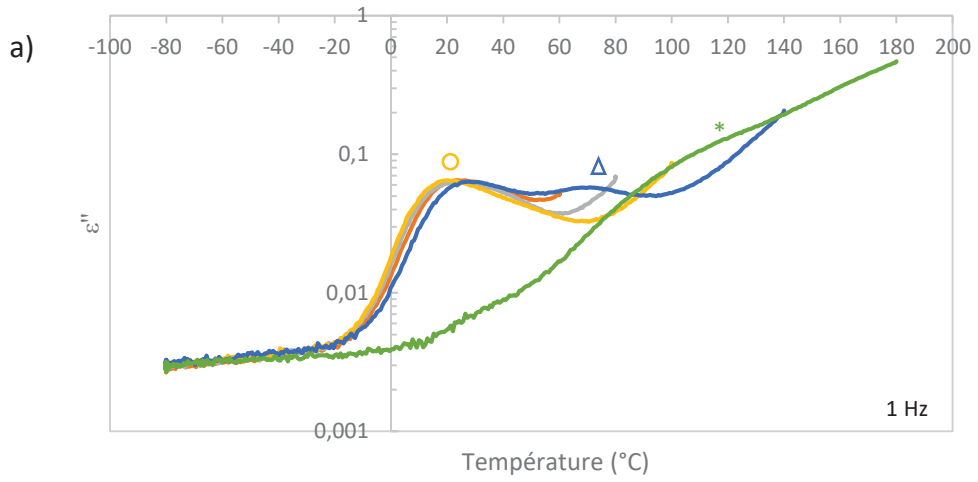


Figure 6-2. Analyse de CtxB<sub>5</sub> à différentes fréquences après le deuxième traitement thermique à 60 °C.

### 6.1.3 Analyse du comportement de CtxB<sub>5</sub> en fonction de la température

Dans la figure 6-3, la dénaturation thermique de la toxine est étudiée. Pour rappel, le solvant montre un pic de relaxation à la fréquence de 1 Hz après les deux traitements thermiques à 60 °C mais pas après celui à 80 °C. La protéine à cette fréquence possède un ou plusieurs processus de relaxation jusqu'au traitement thermique à 180 °C. De même, à la fréquence de  $10^3$  Hz, aucun pic de relaxation du solvant n'est détecté après le deuxième traitement thermique à 60 °C, alors que le pic de relaxation de la toxine commence à apparaître.



*Figure 6-3. Etude du comportement de CxtB<sub>5</sub> après différents traitements thermiques. a) Superposition des courbes de refroidissement de CtxB<sub>5</sub> à la fréquence de 1 Hz. b) Superposition des courbes de refroidissement de la toxine à la fréquence de 10<sup>3</sup> Hz. Insert. Zoom sur les courbes R2 60 °C et R 80 °C à 10<sup>3</sup> Hz. c) Superposition des courbes de refroidissement de la protéine à 10<sup>5</sup> Hz.*

A la fréquence de 1 Hz (Fig. 6-3a), un seul processus de relaxation (symbole rond) est détecté jusqu'au traitement thermique à 100 °C. La position de ce pic de relaxation en température reste inchangée (21 °C) suite aux traitements thermiques de 60 °C à 100 °C. La population conformationnelle détectée est stable jusqu'au traitement thermique à 100 °C. A partir du traitement thermique à 140 °C, ce pic de relaxation change sa position en température (28 °C). Ce résultat suggère un changement de conformation entre les traitements thermiques à 100 °C et 140 °C. Cependant, l'intensité des données de pertes diélectriques ne varie pas jusqu'au traitement thermique à 140 °C. De plus, après les trois heures à 140 °C, un deuxième pic de relaxation apparaît à 73 °C (symbole triangle). Après le traitement thermique à 180 °C, les deux pics de relaxation détectés à 140 °C disparaissent, néanmoins, un troisième pic de relaxation semble être détecté à plus haute température (vers 120 °C). Ce dernier processus de relaxation (symbole étoile) n'est pas assez résolu pour déterminer avec précision sa position en température.

A la fréquence de 10<sup>3</sup> Hz (Fig. 6-3b), le pic de relaxation commence à se dessiner après le traitement thermique à 60 °C et il devient bien défini à partir du traitement thermique à 80 °C (Fig. 6-1 et Insert Fig. 6.3b). Comme à la fréquence de 1 Hz, la position de ce pic de relaxation en température (64 °C) ne varie pas jusqu'au traitement thermique à 100 °C. Après le traitement thermique à 140 °C, la température de ce pic de relaxation augmente à 77 °C, caractéristique d'un changement conformationnel. Après le traitement thermique à 180 °C, aucun processus de relaxation n'est détecté.

A la fréquence de 10<sup>5</sup> Hz (Fig. 6-3c), un seul processus de relaxation (137 °C) est détecté après le traitement thermique à 140 °C.

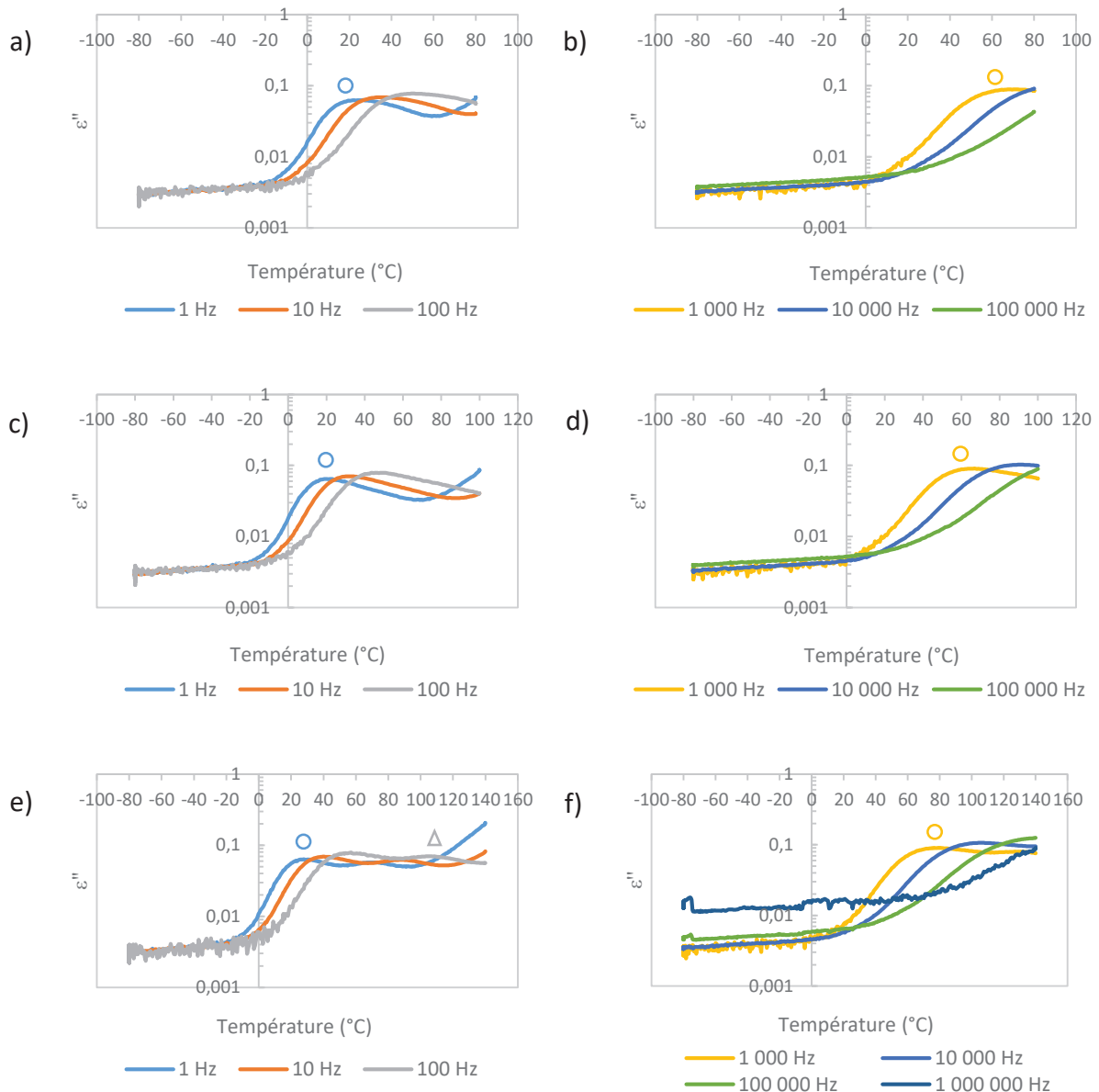
#### 6.1.4 Analyse du comportement de CtxB<sub>5</sub> en fonction de la fréquence

La dépendance en fréquence des processus de relaxation est étudiée en fonction de la température pour les traitements thermiques de 80 °C à 180 °C dans la figure 6-4 et après le traitement thermique à 60 °C dans la figure 6-2. Le premier pic de relaxation détecté couvre les basses fréquences jusqu'au traitement thermique à 80 °C (1 Hz à 10<sup>2</sup> Hz après celui à 60 °C et 1 Hz à 10<sup>3</sup> Hz après 80 °C), principalement les basses fréquences (de 1 Hz à 10<sup>4</sup> Hz) après le traitement thermique à 100 °C, et les basses et hautes fréquences après celui à 140 °C (1 Hz à 10<sup>5</sup> Hz). Le deuxième pic de relaxation détecté après le traitement thermique à 140 °C couvre les basses fréquences (1 Hz à 10<sup>2</sup> Hz) comme le troisième pic de relaxation détecté après le traitement thermique à 180 °C (de 1 Hz à 10<sup>2</sup> Hz).

Pour rappel (Fig. 2-7), une protéine possède des mouvements de structures secondaires de l'ordre de la microseconde à la milliseconde mais aussi des mouvements plus lents sur des plus grandes surfaces : mouvement de repliement, dépliement de domaines, interfaces et structures tertiaires mais aussi des mouvements d'assemblage, dissociation de structures tertiaires. Le premier processus de relaxation détecté dans les basses fréquences



jusqu'au traitement thermique à 80 °C est lié à des mouvements de dépliement de domaines, interfaces et de structures tertiaires (chaînes polypeptidiques) comprenant aussi de potentiels mouvements de dissociation de chaînes (échelle de temps de la seconde à la milliseconde). A partir du traitement thermique à 100 °C, des mouvements de structures secondaires sont aussi détectés. Le deuxième processus de relaxation détecté dans les basses fréquences après le traitement thermique à 140 °C correspond à des mouvements de dépliement (domaines, interfaces, chaînes) et/ou de dissociation de chaînes, comme le troisième processus de relaxation détecté.



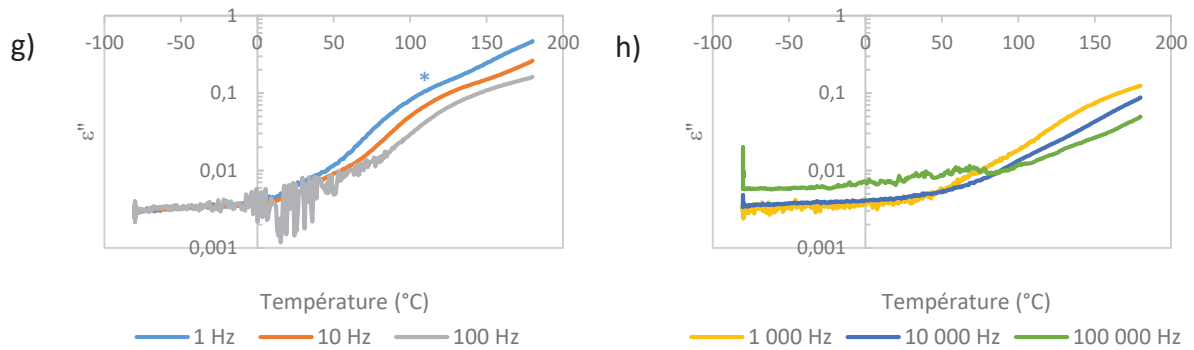


Figure 6-4. Analyse de la dépendance en fréquence des processus de relaxation détectés dans le confinement à 120 nm de diamètre et 2  $\mu\text{m}$  de long après différents traitements thermiques. a) Superposition des courbes de refroidissement de 1 Hz à  $10^2$  Hz après le traitement thermique à 80 °C (R 80 °C). b) Superposition des courbes R 80 °C de  $10^3$  Hz à  $10^5$  Hz. c) Superposition des courbes R 100 °C de 1 Hz à  $10^2$  Hz. d) Superposition des courbes R 100 °C de  $10^3$  Hz à  $10^5$  Hz. e) Superposition des courbes R 140 °C de 1 Hz à  $10^2$  Hz. f) Superposition des courbes R 140 °C de  $10^3$  Hz à  $10^5$  Hz. g) Superposition des courbes R 180 °C de 1 Hz à  $10^2$  Hz. h) Superposition des courbes R 180 °C de  $10^3$  Hz à  $10^5$  Hz.

Le principal pic de relaxation (symbole rond) détecté jusqu'au traitement thermique à 140 °C correspond à la population de conformation protéique la plus stable. De plus, ce processus de relaxation est stable jusqu'au traitement thermique à 100 °C et couvre les basses et hautes fréquences. Par conséquent, d'après les informations obtenues dans les chapitres précédents (Chapitres 4 et 5), ce processus est associé à un pentamère de type natif. Le deuxième processus de relaxation est détecté uniquement après le traitement thermique à 140 °C alors que le troisième est détecté après celui à 180 °C. Par conséquent, il est raisonnable de supposer que le processus de relaxation détecté à 140 °C correspond à un état moins déplié que celui détecté à 180 °C. Pour ces raisons, le troisième processus de relaxation est associé à l'état le plus déplié détecté, soit à des conformations d'intermédiaires d'assemblage et/ou des monomères repliés. Le deuxième processus de relaxation détecté devrait correspondre à un état intermédiaire entre une conformation pentamérique stable (pentamère replié de type natif) et des intermédiaires d'assemblage, nous supposons qu'il s'agit d'une population de conformation pentamérique dépliée ou partiellement dépliée.

En accord avec les chapitres précédents (Chapitres 4 et 5), le premier processus de relaxation correspond à MD<sub>1</sub>, le deuxième processus détecté après le traitement thermique à 140 °C à MD<sub>3</sub> et le troisième détecté après le traitement thermique de 180 °C à MD<sub>2</sub>.

## 6.2 Confinement à 20 nm de diamètre et 2 $\mu\text{m}$ de long

### 6.2.1 Contrôle de l'échantillon protéique

Dans le confinement à 20 nm de diamètre et 2  $\mu\text{m}$  de long, l'échantillon protéique possède un signal diélectrique linéaire dans la gamme de fréquence étudiée après le traitement thermique à 60 °C (Fig. 6-5).

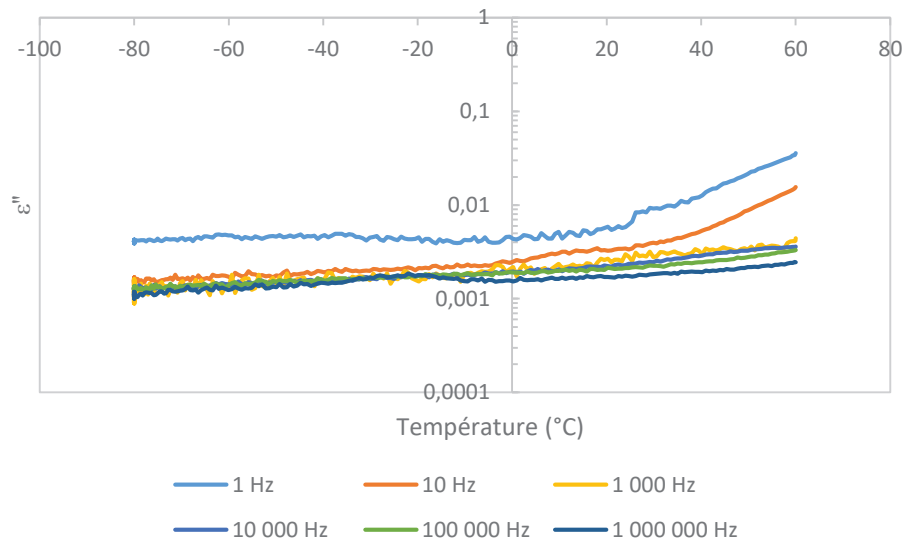


Figure 6-5. Analyse de CtxB<sub>5</sub> dans le confinement à 20 nm de diamètre et 2 μm de long à différentes fréquences après le deuxième traitement thermique à 60 °C. 200 μL de solution composée de 2 μL de toxine à 0,55 mg/ml et 214 μL d'eau déionisée est déposée sur la nanomembrane. La concentration protéique est de 0,0051 mg/ml.

Si l'expérience avec la protéine est reproduite, un léger pic de relaxation à la fréquence de 10<sup>3</sup> Hz est détecté dans ce nouveau test après le traitement thermique à 60 °C. Néanmoins, l'intensité de ce pic de relaxation est similaire à l'intensité de la membrane vide et au test précédent (Fig. 6-6).

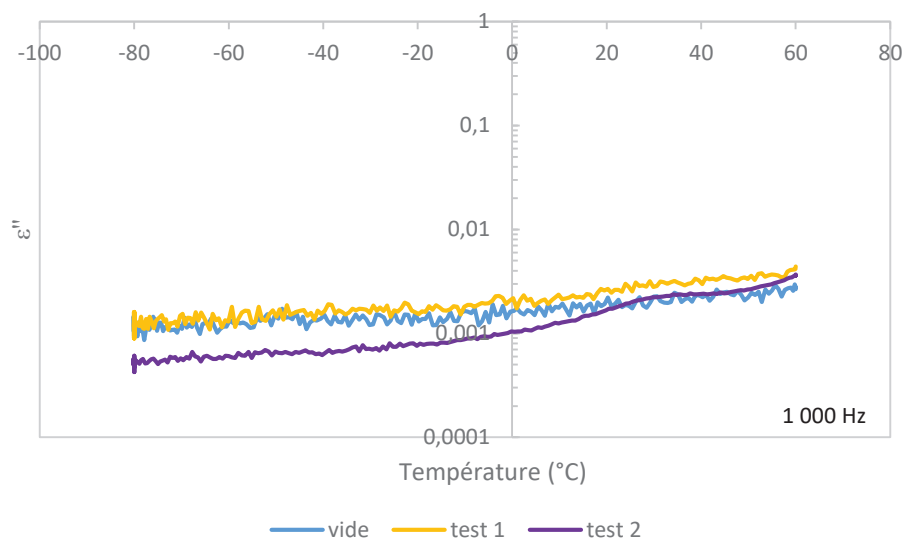


Figure 6-6. Reproduction du signal diélectrique de CtxB<sub>5</sub> à la fréquence de 10<sup>3</sup> Hz dans le confinement à 20 nm de diamètre et 2 μm de long après le traitement thermique à 60 °C. La courbe bleue correspond à la membrane vide. Les courbes jaune et violette à deux expériences identiques avec la toxine.

En diminuant le diamètre du confinement (sans changer le pourcentage de remplissage volumique), un très faible (ou aucun) processus de relaxation est détecté,

suggérant qu'un nombre insuffisant de protéines rentrent dans les nanopores. Deux interprétations sont émises pour expliquer la faible intensité de la dynamique moléculaire dans ce confinement :

- Un encombrement macromoléculaire serait présent à l'entrée des nanopores limitant ainsi l'entrée des protéines (notion d'écoulement à l'échelle nanométrique, plus généralement décrite comme la nanofluidique de l'écoulement).
- La variation du diamètre du confinement entraînerait une sélection conformationnelle et non un enrichissement.

Ce résultat s'expliquerait par la promiscuité entre le diamètre du confinement (20 nm) et le diamètre de la sous-unité B pentamérique de la choléra toxine (6 nm). Un rapport de trois existe entre ces deux diamètres, nous pourrions donc supposer que si deux toxines se fixent à l'entrée du nanopore, la taille du confinement diminuera de 12 nm et ne sera donc plus que de 8 nm, ce qui correspond environ au diamètre de la toxine. Le chapitre 5 a mis en évidence que les confinements de 2  $\mu\text{m}$  et 5  $\mu\text{m}$  de long seraient aussi plus encombrés que celui de 10  $\mu\text{m}$  pour un même diamètre. Par conséquent, il serait intéressant de tester un confinement de 20 nm de diamètre mais avec une longueur de confinement plus grande (exemple : 10  $\mu\text{m}$  de long).

Un autre aspect à prendre en compte est relié avec l'hypothèse émise dans l'introduction, selon laquelle le confinement provoquerait l'enrichissement et/ou la sélection des conformations protéiques. La variation des dimensions du confinement (diamètre et/ou longueur des nanopores) devrait donc permettre d'enrichir et/ou sélectionner différentes populations de conformation et cribler ainsi l'ensemble conformationnel. Dans nos expériences, le pourcentage de remplissage volumique des protéines est le même dans tous les confinements (14,3 %), par conséquent si l'enrichissement conformationnel se produisait, les protéines devraient toujours rentrer dans les nanopores et être en quantité suffisante pour donner un processus de relaxation, car les protéines 'adaptent' leurs conformations (Fig. 2-10). Les hypothèses d'enrichissement et/ou de sélection via les dimensions géométriques du confinement sont en accord avec le fait que le confinement induit par l'encapsulation de protéines dans des micelles inverses de tailles nanométriques est l'un des paramètres qui réduit le nombre de conformations protéiques possibles pour une chaîne polypeptidique (150).

Cette deuxième interprétation suggère que l'effet de confinement sur les conformations protéiques serait différent entre des nanopores et des micelles inverses. L'article de Xu et al en 2017 (137), suggère que le confinement induit par les micelles inverses impacte la conformation protéique et entraîne un enrichissement conformationnel plutôt qu'une sélection de conformation. Si la protéine possède une structure compacte alors le confinement induit par les micelles ne semble pas avoir d'impact sur la conformation protéique. Ce qui n'est pas le cas si la conformation protéique est étendue, où dans ce cas le confinement sphérique des micelles favorise les conformations compactes. Pour information, dans l'étude de Xu et al, les auteurs considèrent qu'une seule protéine de calmoduline est encapsulée dans chaque micelle inverse. Les micelles inverses ont des tailles d'environ 12 nm (121,6 Å), alors que la calmoduline étendue possède une taille de 6 nm contre une taille de 4

nm pour la calmoduline compacte. Cette différence entre sélection et enrichissement pourrait s'expliquer par la dimension 3D des conformations protéiques dans les micelles versus une dimension 'plus' 2D dans les nanopores.

### 6.3 Comparaison

Les trois confinements comparés dans cette partie possèdent tous une longueur de 2  $\mu\text{m}$ , mais un diamètre de confinement différent qui varie entre 20 nm, 40 nm et 120 nm.

Dans le chapitre 5, nous avons supposé que les conformations non natives détectées suite au dépliement de CtxB<sub>5</sub> provenaient uniquement de la longueur du confinement car l'état de déshydratation était le même pour les différents confinements étudiés (partie 5.3.1, Fig. 5-13). La même analyse que précédemment est réalisée pour déterminer si la déshydratation, la variation du diamètre ou la combinaison des deux a un impact sur les conformations protéiques non natives détectées.

#### 6.3.1 Effet de la déshydratation

Pour rappel, le pourcentage de remplissage volumique des nanopores est de 14,3 % pour tous les confinements. De plus, l'échantillon protéique est toujours soumis 15 minutes à 50 °C et 5 minutes à 30 °C, avant les différents traitements thermiques (Chapitre 3). Suite aux différents traitements thermiques, un cycle de température comprenant une courbe de refroidissement et une courbe de chauffage est réalisé.

Comme pour l'étude de l'impact de la longueur, le signal diélectrique est reproductible pour les deux cycles mesurés après les deux traitements thermiques à 60 °C : même position du ou des pics de relaxation en température, de l'intensité et de la largeur des pics détectés (Fig. 6-7). De même, les données de pertes diélectriques ne détectent pas de différence significative à l'intérieur de chaque cycle.

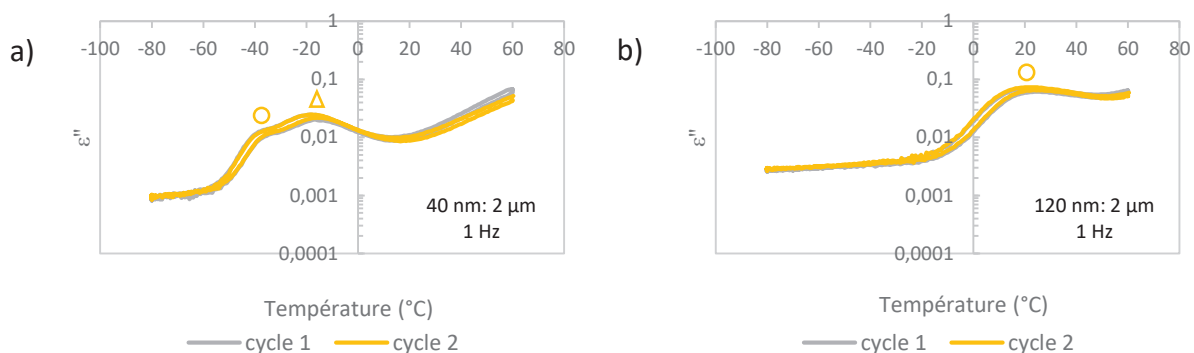


Figure 6-7. Superposition des courbes de refroidissement et de chauffage à la fréquence de 1 Hz après deux traitements thermiques à 60 °C. a) Confinement à 40 nm de diamètre et 2  $\mu\text{m}$  de long. b) Confinement à 120 nm de diamètre et 2  $\mu\text{m}$  de long.

L'absence d'hystérésis et la reproduction du signal diélectrique si un cycle est répété sous-entend que l'état de déshydratation est le même dans les différents confinements étudiés. Par conséquent, il est raisonnable de supposer que les conformations protéiques non natives détectées proviennent uniquement du confinement et non de la déshydratation de l'échantillon protéique.

## 6.3.2 Effet de la variation du diamètre du confinement

La comparaison des positions des processus de relaxation obtenus en faisant varier le diamètre du confinement après le traitement thermique à 60 °C est représentée dans la figure 6-8. Pour rappel, les pentamères natifs ne sont pas détectés dans notre gamme de fréquence (3,124). Par conséquent, MD<sub>1</sub> (symbole rond) est associé à une population de pentamère où l'interface N-terminale (interface I<sub>1</sub>) est déstabilisée, MD<sub>3</sub> (symbole triangle) est associé à une population de pentamère déplié ou partiellement déplié où l'interface principale (interface I<sub>2</sub>) est perturbée et MD<sub>2</sub> est composé d'intermédiaires d'assemblage et/ou de monomères repliés où l'interface de l'hélice centrale (interface I<sub>3</sub>) est déstabilisée.

La longueur des nanopores est identique pour les trois confinements, de même que le pourcentage de remplissage volumique dans les nanopores (14,3 %). La solution de CtxB<sub>5</sub> dans le confinement à 120 nm de diamètre et 2 µm de long possède un pH de 7,2, contre un pH de 6,8 dans le confinement à 40 nm de diamètre et 2 µm de long et un pH de 7 dans le confinement à 20 nm de diamètre et 2 µm de long. Les mêmes acides aminés (lysine, arginine, acide aspartique et acide glutamique) et les deux extrémités N-terminale et C-terminale sont donc chargés dans les trois confinements testés.

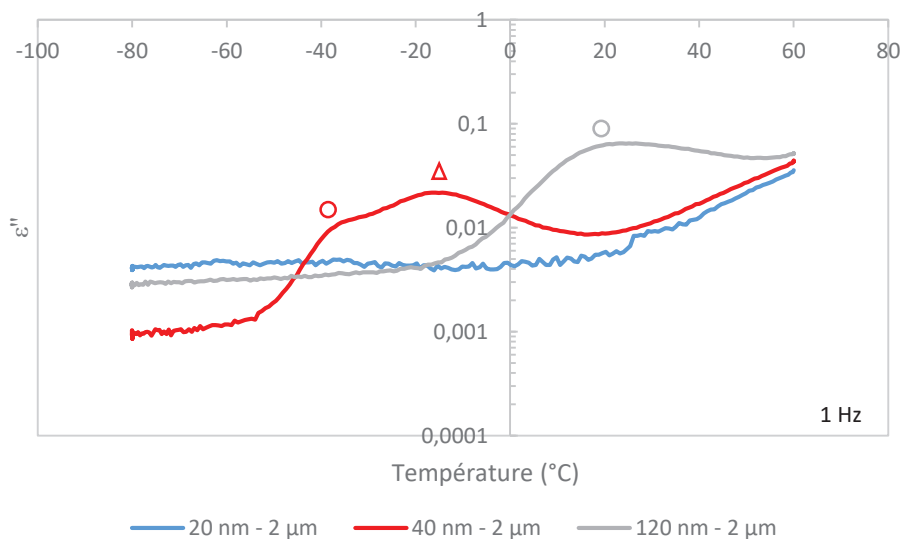


Figure 6-8. Comparaison des processus de relaxation obtenus à la fréquence de 1 Hz en faisant varier le diamètre du confinement après le deuxième traitement thermique à 60 °C. Superposition des courbes R2 60 °C pour les confinements à 20 nm de diamètre (courbe bleu), à 40 nm de diamètre (courbe rouge) et à 120 nm de diamètre (courbe grise). Le symbole rond correspond à la population conformationnelle la plus stable dans le confinement.

La table 6-1 représente les intensités et les positions des pics de relaxation en température ( $\theta_{max}$ ) associées au processus de relaxation MD<sub>1</sub> à la fréquence de 1 Hz pour différents diamètres de confinement.

		120 nm – 2 $\mu$ m	40 nm – 2 $\mu$ m
Cycle 60 °C	$\theta_{\max}$ (°C)	21	-33
	Intensité	0,061	0,011
Cycle 80 °C	$\theta_{\max}$ (°C)	21	-33
	Intensité	0,059	0,0089
Cycle 100 °C	$\theta_{\max}$ (°C)	21	-33
	Intensité	0,061	0,0041
Cycle 140 °C	$\theta_{\max}$ (°C)	28	X
	Intensité	0,060	
Cycle 180 °C	$\theta_{\max}$ (°C)	X	X
	Intensité		

Table 6-1. Intensité et position des pics de relaxation à la fréquence de 1 Hz associée au processus de relaxation MD<sub>1</sub> après différents traitements thermiques pour différents confinements en variant le diamètre.

La température du pic de relaxation associé à la conformation protéique la plus stable (MD<sub>1</sub>) augmente avec l'augmentation du diamètre du confinement (Fig. 6-8 et table 6-1). La position de MD<sub>1</sub> en température est de -33 °C dans le confinement à 40 nm de diamètre et de 21 °C dans le confinement à 120 nm de diamètre, suggérant une sélection et/ou un enrichissement conformationnel différent entre les deux confinements. De plus, la position du pic de relaxation en température ne change pas jusqu'au traitement thermique à 100 °C dans les deux confinements (Table. 6-1), traduisant un comportement protéique similaire. Le processus de relaxation MD<sub>1</sub> n'est plus détecté après le traitement thermique à 140 °C pour le confinement à 40 nm de diamètre alors que ce processus de relaxation est toujours détecté pour le confinement de 120 nm de diamètre (28 °C), suggérant que les conformations protéiques enrichies et/ou sélectionnées ne possèdent pas la même stabilité thermique. Dans le confinement à 120 nm de diamètre, après le traitement thermique à 140 °C, MD<sub>1</sub> est décalé de 7 °C suggérant des pentamères plus déstabilisés que pour les traitements thermiques précédents.

Les intensités du processus de relaxation MD<sub>1</sub> sont plus importantes dans le confinement à 120 nm de diamètre que dans le confinement à 40 nm de diamètre (Table. 6-1). Ainsi, la densité volumique de dipôles détectés est plus importante dans le confinement de 120 nm. Or le pourcentage de remplissage est le même dans les deux confinements étudiés et l'intensité est proportionnelle à la densité de dipôle, par conséquent les intensités du processus de relaxation MD<sub>1</sub> auraient dû être les mêmes après le traitement thermique à 60 °C. Ces données suggèrent une différence d'écoulement dans les nanopores entre ces deux confinements, plus précisément, l'encombrement macromoléculaire doit-être plus important dans le confinement à 40 nm de diamètre. Ceci est cohérent avec une interprétation réalisée précédemment pour expliquer la faible intensité du processus de relaxation mesurée dans le confinement à 20 nm de diamètre.

Le pic de relaxation MD<sub>1</sub> est plus large dans le confinement de 120 nm de diamètre que dans le confinement de 40 nm de diamètre (38 °C versus 21 °C) suggérant une population conformationnelle plus hétérogène. L'hétérogénéité plus importante dans le confinement de 120 nm de diamètre est cohérente avec les deux interprétations mentionnées précédemment

sur l'effet du diamètre. En effet, en augmentant le diamètre de confinement, l'encombrement macromoléculaire devrait être réduit, de même que la sélection conformationnelle.

En conclusion, la variation du diamètre de confinement permet de sélectionner et/ou enrichir différentes populations conformationnelles et devrait donc permettre de cribler l'ensemble conformationnel d'une protéine. L'augmentation du diamètre de confinement tend à réduire l'enrichissement et/ou la sélection conformationnelle. Le diamètre du nanoconfinement impacte l'écoulement dans les nanopores et potentiellement la sélection conformationnelle.

A la fréquence de 1 Hz, le processus de relaxation MD<sub>3</sub> associé à des populations de pentamères partiellement dépliés ou dépliés est détecté dès le traitement thermique à 60 °C pour le confinement à 40 nm de diamètre et après le traitement thermique à 140 °C pour le confinement à 120 nm de diamètre (Table. 6-2).

		120 nm – 2 μm	40 nm – 2 μm
Cycle 60 °C	θ <sub>max</sub> (°C)	X	-14
	Intensité		0,004
Cycle 80 °C	θ <sub>max</sub> (°C)	X	-14
	Intensité		0,011
Cycle 100 °C	θ <sub>max</sub> (°C)	X	-5
	Intensité		0,014
Cycle 140 °C	θ <sub>max</sub> (°C)	73	X
	Intensité	0,005	
Cycle 180 °C	θ <sub>max</sub> (°C)	X	X
	Intensité		

Table 6-2. Intensité et position du processus de relaxation en température associée au processus de relaxation MD<sub>3</sub> à la fréquence de 1 Hz après différents traitements thermiques pour différents confinements en variant le diamètre.

La même analyse que pour le chapitre 5 à la fréquence de 10<sup>3</sup> Hz peut être réalisée à la fréquence de 1 Hz pour le confinement de 40 nm de diamètre et 2 μm de long. MD<sub>3</sub> est présent dès le traitement thermique à 60 °C pour ce confinement (Fig. 6-8 et table 6-2). Précédemment, nous avons observé que la stabilité thermique des protéines était différente suivant le confinement, par conséquent, dans ce confinement l'interface I<sub>2</sub> pourrait se déstabiliser pendant les trois heures à 60 °C entraînant la formation de pentamères partiellement dépliés. De plus, l'intensité du pic de relaxation MD<sub>1</sub> est divisée par deux entre les traitements thermiques à 80 °C et 100 °C, alors que l'intensité de MD<sub>3</sub> ne varie presque pas entre ces deux traitements thermiques, par conséquent il est raisonnable de supposer que les pentamères de type natif (MD<sub>1</sub>) ne sont pas uniquement dépliés en pentamère partiellement repliés (MD<sub>3</sub>).

Pour le confinement à 120 nm de diamètre, le processus de relaxation MD<sub>3</sub> apparaît après le traitement thermique à 140 °C, toutefois l'intensité du pic de relaxation MD<sub>1</sub> ne change pas (Table. 6-1). Comme émis dans le chapitre 5, ce résultat suggère que les pentamères repliés (MD<sub>1</sub>) se déplient en pentamères partiellement repliés (MD<sub>3</sub>) jusqu'à atteindre une valeur inférieure au seuil maximum de détection, où à partir de ce moment



l'intensité associée à MD<sub>1</sub> commencera à diminuer. Pour vérifier cette hypothèse, il serait nécessaire de tester un confinement avec différents pourcentages de remplissage volumique. Si le signal diélectrique du pic de relaxation ne varie pas malgré la variation du nombre de protéines dans les nanopores, cette hypothèse sera validée. Autre explication, la contribution des dipôles serait plus élevée dans le processus de relaxation MD<sub>3</sub> compensant une faible population conformationnelle participant à ce pic. Cette deuxième hypothèse suggère la présence d'un seuil minimum de détection.

Dans le confinement à 40 nm de diamètre, le processus de relaxation MD<sub>2</sub> (faiblement détecté) apparaît après le traitement thermique à 100 °C et coïncide avec la diminution d'intensité du pic de relaxation MD<sub>1</sub>. De plus, ce processus de relaxation est aussi légèrement détecté jusqu'au traitement thermique à 180 °C, alors que les pics de relaxation MD<sub>1</sub> et MD<sub>3</sub> ne le sont plus (Fig. 5-10a). La même observation est faite pour le confinement à 120 nm de diamètre, où MD<sub>2</sub> apparaît après le traitement thermique à 180 °C alors que MD<sub>1</sub> et MD<sub>3</sub> ne sont plus détectés (Fig. 6-3a). Ces résultats suggèrent que les populations de conformations d'intermédiaires d'assemblage et/ou de monomères repliés (MD<sub>2</sub>) résultent du dépliement des pentamères repliés (MD<sub>1</sub>) et partiellement dépliés (MD<sub>3</sub>).

En conclusion, l'augmentation de la température des traitements thermiques peut provoquer une déstabilisation plus importante de l'interface I<sub>1</sub> des pentamères de type natif entraînant un décalage dans la position en température du pic de relaxation MD<sub>1</sub>, une déstabilisation de l'interface I<sub>2</sub> des pentamères de type natif entraînant la formation de pentamères partiellement repliés (MD<sub>3</sub>). De plus, les interfaces I<sub>3</sub> de ces pentamères de type natif (MD<sub>1</sub>) et partiellement dépliés (MD<sub>3</sub>) peuvent être déstabilisées pour donner des intermédiaires d'assemblage et/ou des monomères repliés (MD<sub>2</sub>).

Dans le confinement à 40 nm de diamètre (Fig. 6-9a), la position du pic de relaxation MD<sub>1</sub> (symbole rond) en température est identique si l'expérience est reproduite (-33 °C), alors que la position du pic de relaxation MD<sub>3</sub> (symbole triangle) montre un léger décalage en température (-14 °C versus -17 °C). Si l'expérience est reproduite pour le confinement à 120 nm de diamètre (Fig. 6-9b), la position de MD<sub>1</sub> en température est identique (21 °C). Ces résultats confirment les observations présentées dans le chapitre 5 sur l'étude de la longueur du confinement : les pentamères repliés de type natif sont des conformations protéiques stables, ce qui n'est pas le cas des populations conformationnelles issues de la déstabilisation des interfaces de ces pentamères (MD<sub>3</sub>).

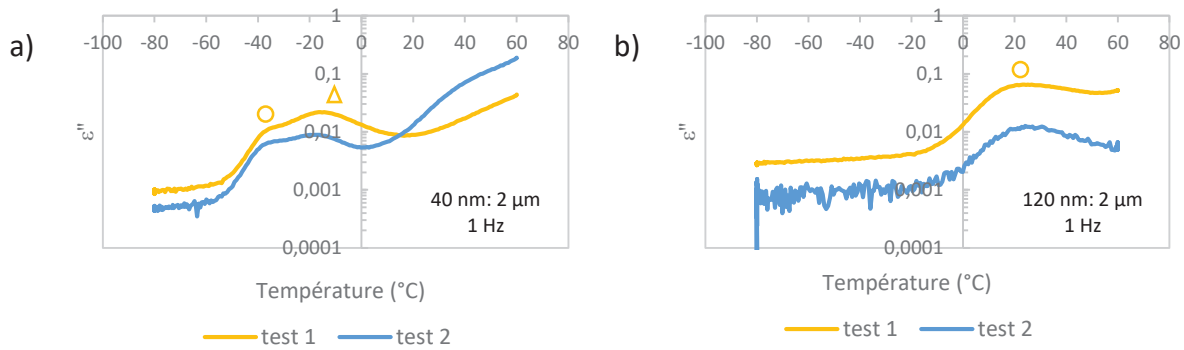


Figure 6-9. Reproduction des expériences des confinements de diamètres différents à la fréquence de 1 Hz. a) Expérience dans le confinement à 40 nm de diamètre et 2  $\mu\text{m}$  de long. b) Expérience dans le confinement à 120 nm de diamètre et 2  $\mu\text{m}$  de long.

Les chapitres 5 et 6 sur la variation de la longueur et du diamètre du confinement ont mis en évidence que les variations des dimensions du confinement (longueur et diamètre) permettaient de sélectionner et/ou d'enrichir différentes populations conformationnelles. La combinaison du nanoconfinement et de la spectroscopie diélectrique à large bande peut donc être utilisée pour cribler différentes conformations protéiques.

L'étude de différents confinements suggère que la variation des dimensions du confinement impacte l'écoulement dans les nanopores. Plus précisément, l'impact de l'encombrement macromoléculaire augmente si le diamètre du confinement et/ou si la longueur de confinement diminue. Pour le confinement de 20 nm de diamètre et 2  $\mu\text{m}$  de long aucun ou un faible processus de relaxation est détecté, par conséquent, en plus d'un impact de type nanofluidique, le diamètre du confinement pourrait potentiellement avoir un effet sur la sélection conformationnelle des protéines. Pour confirmer cette hypothèse, une étude est réalisée sur des nanomembranes avec des nanopores de 15 nm de diamètre et 5  $\mu\text{m}$  de long. Contrairement aux autres nanomembranes utilisées pendant la thèse, les nanomembranes avec ce confinement ont été fabriquées au laboratoire Ingénierie des Matériaux Polymères en utilisant le protocole mis au point pendant la thèse d' Afef Houachtia (86). Les dimensions des nanomembranes sont légèrement différentes de celles produites par InRedox et possèdent une porosité de 15 % au lieu de 12 % (Table. 3-2).

Dans ce confinement, l'échantillon protéique de CtxB<sub>5</sub> présente un processus de relaxation dès le traitement thermique à 60 °C (Fig. 6-10a). L'étude sur ce confinement est dans sa phase préliminaire et les mesures de reproductibilité ne sont pas encore effectuées, de même le contrôle du solvant n'a pas encore été réalisé. Néanmoins, au vue des résultats obtenus dans les chapitres précédents et ce chapitre (Chapitres 4 à 6), il est raisonnable de supposer que le processus de relaxation détecté dans ce confinement provient des dipôles des protéines et non du solvant. De plus, les résultats protéiques de ce confinement sont cohérents avec les résultats de CtxB<sub>5</sub> dans le confinement à 40 nm de diamètre et 5  $\mu\text{m}$  de long. En effet, pour les différentes fréquences étudiées :

- Le comportement du processus de relaxation associé à la conformation protéique la plus stable (symbole rond) change principalement entre les traitements thermiques à 100 °C et 140 °C (Fig. 6-10a).
- Aucun processus de relaxation n'est détecté après le traitement thermique à 180 °C (Fig. 6-10b).

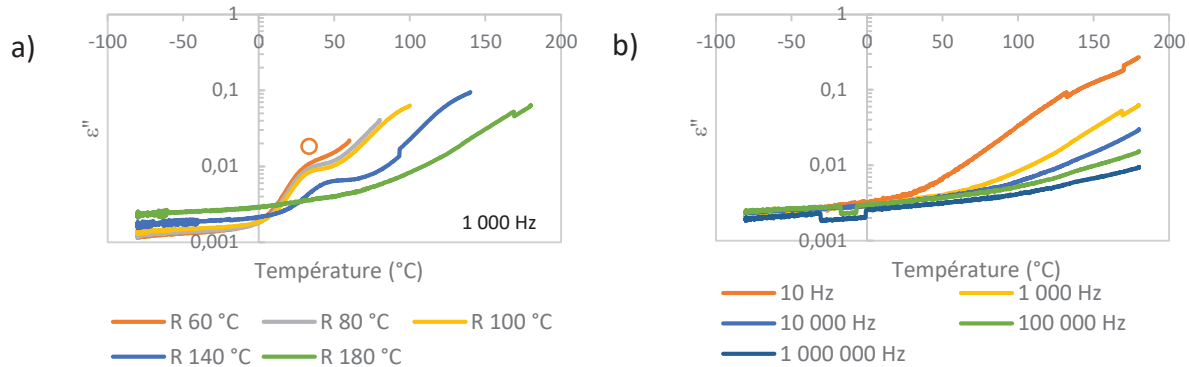


Figure 6-10. Analyse du comportement de CtxB<sub>5</sub> dans le confinement à 15 nm de diamètre et 5  $\mu$ m de long. a) Superposition des courbes de refroidissement mesurées à la fréquence de 10<sup>3</sup> Hz après différents traitements thermiques. b) Superposition des courbes de refroidissement à différentes fréquences après le traitement thermique à 180 °C (R 180 °C). La concentration protéique de la solution déposée sur la nanomembrane est de 0,013 mg/ml. Le pourcentage de remplissage des nanopores est de 10,1 %.

La présence d'un processus de relaxation dans le confinement à 15 nm de diamètre et 5  $\mu$ m de long contredit l'effet de sélection conformationnelle via le diamètre du confinement. En effet, si le diamètre de 20 nm exclu les conformations protéiques, la même chose devrait se produire si le diamètre de confinement diminue. Toutefois, il est nécessaire de faire attention car la longueur du confinement varie aussi entre ces deux expériences (confinement de 20 nm de diamètre et 2  $\mu$ m de long versus confinement de 15 nm de diamètre et 5  $\mu$ m de long), or le chapitre 5 a montré que la variation de longueur impactait l'écoulement dans les nanopores mais aussi l'enrichissement et/ou la sélection conformationnelle.

La position du processus de relaxation en température associé à la population conformationnelle la plus stable (symbole rond) est différente entre le confinement à 40 nm de diamètre et 5  $\mu$ m de long et le confinement à 15 nm de diamètre et 5  $\mu$ m de long. Après le traitement thermique à 60 °C et à la fréquence de 10<sup>3</sup> Hz, le pic de relaxation se situe à -1,5 °C pour le confinement à 40 nm de diamètre contre 34 °C pour le confinement à 15 nm de diamètre (Fig. 6-11), confirmant que la variation du diamètre permet de cribler différentes populations conformationnelles. La largeur du pic de relaxation est similaire dans les deux confinements (37 °C pour le confinement à 15 nm de diamètre et 36,5 °C pour le confinement à 40 nm de diamètre) par conséquent, l'hétérogénéité conformationnelle est similaire. L'intensité du processus de relaxation est plus importante dans le confinement à 40 nm de diamètre que dans celui à 15 nm (0,016 versus 0,0098), ce qui est cohérent avec le fait que le pourcentage de remplissage soit plus élevé dans le confinement à 40 nm (14,3 % versus 10,1 %).

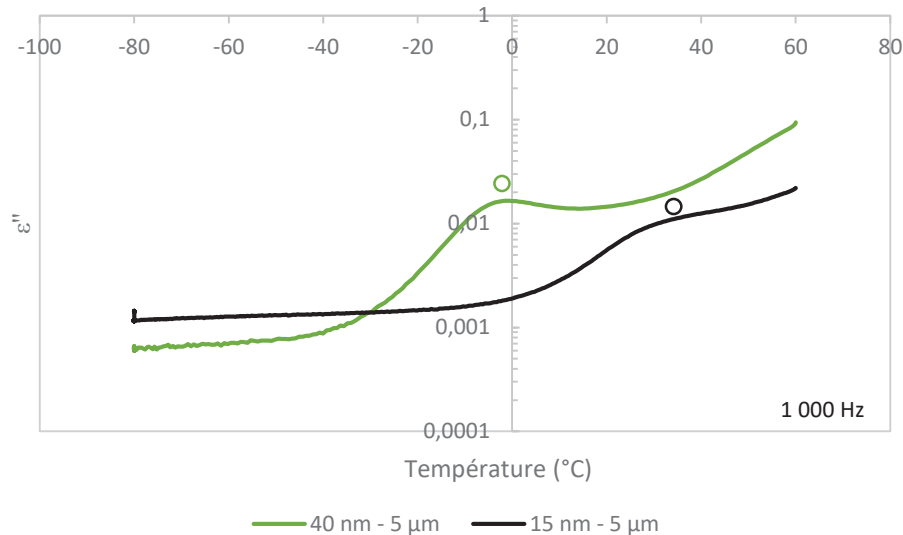


Figure 6-11. Comparaison du signal de CtxB<sub>5</sub> à la fréquence de 10<sup>3</sup> Hz en diminuant le diamètre de confinement après le traitement thermique à 60 °C. Superposition des courbes de refroidissement pour le confinement à 40 nm de diamètre (courbe verte) où le remplissage volumique des nanopores est de 14,3 % et pour le confinement à 15 nm de diamètre (courbe noire) où le remplissage volumique est de 10,1 %.

L'analyse des différents confinements démontre que le choix des dimensions du confinement à la fois via la longueur mais aussi via le diamètre est essentiel pour limiter les problèmes d'encombrements dans les nanopores. L'encombrement macromoléculaire peut impacter les vitesses et les équilibres de réaction (151) mais aussi la structure, dynamique et fonction des protéines (59,152).

Dans la comparaison présentée dans la figure 6-11, le pourcentage de remplissage volumique est de 14,3 % pour le confinement à 40 nm de diamètre, tandis que le pourcentage de remplissage pour le confinement à 15 nm de diamètre est de 10,1 %. Par conséquent, cette variation dans le pourcentage de remplissage et donc dans la quantité de protéine pourrait potentiellement impacter le signal diélectrique, plus précisément la position du pic de relaxation en température, la largeur du pic et son intensité. Pour vérifier cette hypothèse, CtxB<sub>5</sub> dans le confinement à 15 nm de diamètre et 5 μm de long est étudié à un autre pourcentage de remplissage volumique (66 %).

La figure 6-12, montre que pour les deux pourcentages de remplissage volumique dans le confinement à 15 nm de diamètre et 5 μm de long, la position du pic de relaxation en température est la même après le traitement thermique à 60 °C (34 °C). Par conséquent, la population conformationnelle enrichie et/ou sélectionnée via les dimensions du confinement ne change pas malgré l'augmentation du pourcentage de remplissage. La largeur du pic de relaxation est de 33 °C et 37 °C entre les deux pourcentages de remplissage, par conséquent, l'hétérogénéité conformationnelle est similaire. L'intensité du processus de relaxation est plus importante si le pourcentage de remplissage est de 10,1 % (0,0098 versus 0,0076). En augmentant le pourcentage de remplissage, le nombre de protéines présentes dans la solution augmentent, par conséquent plus de protéines devraient rentrer dans les

nanopores. Or comme l'intensité est proportionnelle à la densité de dipôle, en augmentant le pourcentage de remplissage, l'intensité du signal du processus de relaxation aurait aussi dû augmenter. Comme ce n'est pas le cas, ce résultat suggère que l'augmentation du pourcentage de remplissage volumique des nanopores et donc l'augmentation du nombre de protéines pourrait provoquer un effet d'encombrement ou d'embouteillage à l'entrée des nanopores, modifiant le processus d'écoulement.

En conclusion, la différence de pourcentage de remplissage volumique des nanopores n'impacte pas la population conformationnelle sélectionnée et/ou enrichie (position du pic de relaxation en température), l'hétérogénéité de la solution (largueur du pic) mais uniquement le nombre de protéines qui rentrent dans les nanopores (intensité). La comparaison entre les différents confinements via la position des pics de relaxation et la largeur des pics est réalisable malgré les différences dans le pourcentage de remplissage.

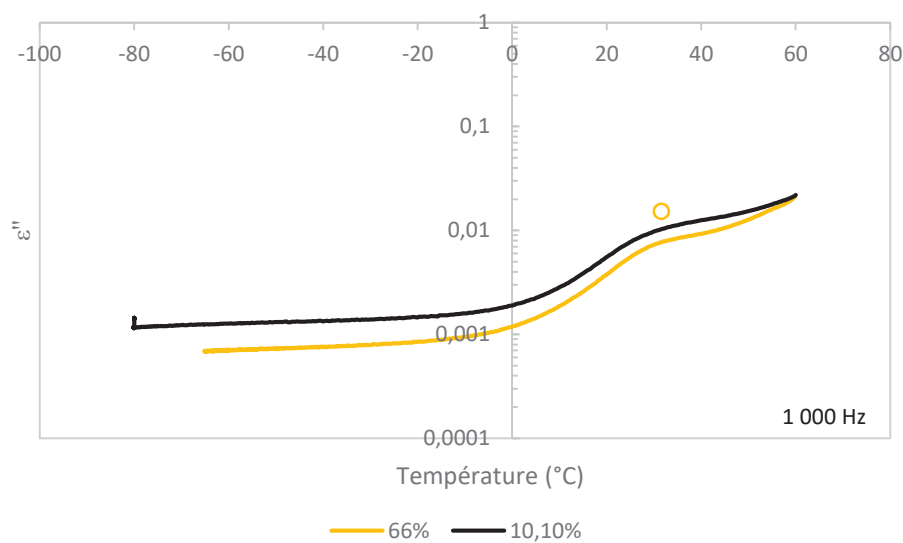


Figure 6-12. Etude de CtxB<sub>5</sub> avec différents pourcentages de remplissage dans le confinement à 15 nm de diamètre et 5  $\mu$ m de long à la fréquence de  $10^3$  Hz après le traitement thermique à 60 °C. La courbe jaune, possède un pourcentage de remplissage volumique des nanopores de 66 %, et la courbe noire un pourcentage de remplissage volumique de 10,1 %. La solution déposée sur la nanomembrane possède une concentration de 0,013 mg/ml pour un pourcentage de remplissage des nanopores de 10,1 % et de 0,035 mg/ml pour un pourcentage de remplissage de 66 %.

Il est intéressant de noter que contrairement au remplissage volumique des nanopores à 10,1 % où la température descend jusqu'à -80 °C, la température de l'expérience avec le remplissage volumique à 66 % s'est arrêtée à -65 °C. Malgré cette différence de température, le processus de relaxation est le même pour les deux expériences. Ces résultats peuvent s'expliquer par le fait qu'à partir d'une certaine température, les protéines dans les nanomembranes sont sûrement 'congelées' suggérant ainsi que même si toutes les expériences réalisées descendent jusqu'à -80 °C, la température minimale pourraient être plus élevée (exemple : -65 °C), et permettre ainsi de diminuer le temps des expériences.

Maintenant que la combinaison du nanoconfinement et de la spectroscopie diélectrique à large bande est validée pour cribler différentes conformations protéiques de

CtxB<sub>5</sub>, le chapitre suivant (chapitre 7) est utilisé pour vérifier que cette technique est applicable sur d'autres protéines et permet notamment d'étudier des variants de séquence.

# Chapitre 7

## Etude sur d'autres protéines

Le chapitre 2 sur l'état de l'art a expliqué que la mesure de la dynamique spatio-temporelle des mouvements collectifs lents dans les protéines était cruciale car il s'agissait de la plage temporelle où se produit le repliement des protéines et les mouvements impliqués dans l'activité biologique. Toutefois, cette plage temporelle permet aussi d'évaluer 'l'état de santé' d'une protéine, car les mutations pathologiques se distinguent des mutations robustes par différents mouvements collectifs (153). Les mutations robustes sont associées à un ou plusieurs changements d'acides aminés dans la séquence protéique n'impactant pas la fonction protéique. Les mutations robustes sont donc des solutions de séquences alternatives fonctionnelles. Inversement, les mutations pathologiques possèdent un ou plusieurs changements d'acides aminés impactant la fonction protéique, celle-ci peut-être altérée ou perdue.

La robustesse des protéines aux mutations ou aux perturbations externes telles que les changements environnementaux repose sur la diversité des structures adoptées par une protéine pour fonctionner (hétérogénéité conformationnelle) (5,154–156). Mais il existe peu de mesures de la diversité structurale dans une protéine ou ses variants de séquence à partir desquelles comparer les perturbations structurales et dynamiques des mutations robustes ou pathologiques pour effectuer un diagnostic de l'état de santé de la protéine (115,157,158).

Luca Ponzonia et Ivet Bahar en 2017 (12) ont montré qu'il était nécessaire de prendre en compte la dynamique structurale de la protéine et non uniquement la séquence et la conformation pour déterminer l'impact des variants sur la fonction biologique.

Le chapitre 4 a démontré que la combinaison du nanoconfinement et de la spectroscopie diélectrique à large bande (BDS) permettait de caractériser une protéine en termes de conformations, stabilités et dynamiques moléculaires notamment grâce à sa sensibilité à la perturbation des interactions moléculaires locales. Cette sensibilité aux perturbations locales pourra être utilisée pour étudier une protéine et ses variants de séquence et ainsi diagnostiquer l'impact des mutations. Ces mesures expérimentales seraient ensuite utilisées pour classer les mutations par rapport à la protéine non-mutée suivant la position des pics de relaxation en température. De plus, le nanoconfinement permet d'étudier la dynamique des protéines à des échelles de longueur comparables aux dimensions géométriques des protéines, ce qui devrait améliorer la sensibilité des mesures diélectriques et aider à détecter l'impact des mutations des protéines. A l'échelle macroscopique, cette

capacité est fortement limitée par les effets de désordre résultant d'une très large distribution des conformations protéiques.

Dans un premier temps, deux toxines similaires : pentamériques avec 83,5 % d'identité de séquence, des structures 3D superposables et des fonctions similaires seront étudiées pour déterminer si la combinaison du nanoconfinement et de la spectroscopie diélectrique à large bande permet de distinguer des processus de relaxation différents entre des variants de séquence et donc de détecter l'impact des mutations sur la dynamique moléculaire. Ensuite, une protéine globalement différente (un monomère) sera étudiée pour déterminer si le nanoconfinement et la spectroscopie diélectrique à large bande est capable de détecter de grandes variations structurales et donc des 'mutations pathologiques'.

### 7.1 Etude d'un variant de séquence : protéine LTB<sub>5</sub>

La sous-unité B pentamérique de la choléra toxine (CtxB<sub>5</sub>, code PDB : 1EEI) et l'entérotoxine thermolabile B pentamérique (code PDB : 1LTR) possèdent toutes les deux 103 d'acides aminés. Sur ces 103 résidus, 86 acides aminés sont identiques et 17 sont différents (résidus en bleu dans la figure 7-1). L'entérotoxine thermolabile B pentamérique (LTB<sub>5</sub>) possède donc 17 mutations d'acides aminés : A1T (le premier acide amine est dans la séquence primaire de LTB<sub>5</sub>, le deuxième acide aminé dans CtxB<sub>5</sub>, le nombre correspond à la position du résidu dans la séquence), S4N, E7D, S10A, Y18H, I20L, L25F, M31L, V38A, S44N, T75A, T80A, I82V, D83E, N94H, S95A, E102A (Fig. 7-1). Les deux toxines pentamériques possèdent une identité de séquence de 83,5 %. Dans la séquence 1LTR, l'acide aminé asparagine (code une lettre : N) en position 103 a été substitué par une lysine (code une lettre : K), par conséquent la mutation du résidu 103 (couleur jaune Fig. 7-1) n'est pas présente entre les deux toxines (159).

```

1EEI  1  10  20  30  40  50
      TPQNITDLCAEYHNTQIHTLNDKIFSYTESLAGKREMAIITFKNGATFQVE
1LTR  APQSITELCSEYHNTQIYTINDKILSYTESMAGKREMIITFKSGATFQVE

      52  60  70  80  90  100
1EEI  VPGSQHIDSQKKAIERMKDTRLRIAYLTEAKVEKLCVWNNKTPHAIAAISMAN
1LTR  VPGSQHIDSQKKAIERMKDTRLRI TYLTE TKIDKLCVWNNKTPNSIAAISMEK

```

Figure 7-1. Alignement de séquence des toxines : CtxB<sub>5</sub> (code PDB :1EEI) et LTB<sub>5</sub> (code PDB :1LTR). En bleu représentation des différentes mutations entre les deux séquences.

Les conformations protéiques des deux toxines LTB<sub>5</sub> et CtxB<sub>5</sub> sont superposables (Fig. 7-2). La différence de RMS (root mean square : moyenne quadratique) pour les atomes du squelette peptidique des sous-unités B pentamériques de la choléra toxine et de l'entérotoxine thermolabile est de 0,55 Å (160). De plus, les deux toxines possèdent des fonctions similaires (77).



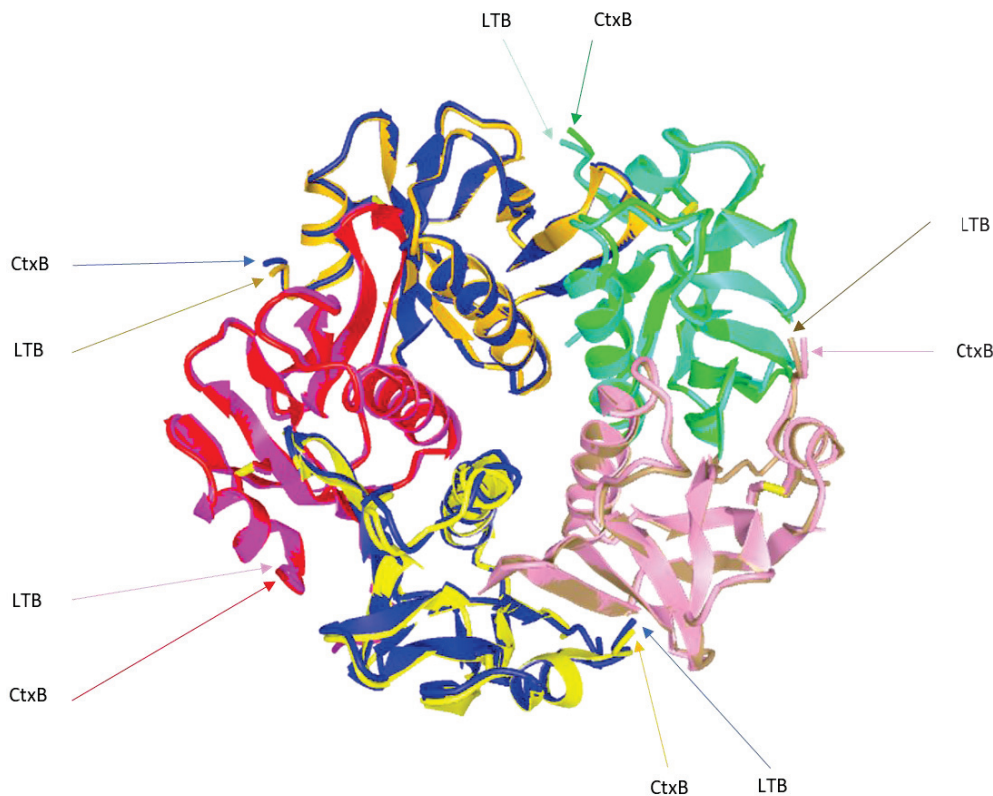


Figure 7-2. Alignement des conformations protéiques de CtxB<sub>5</sub> (code PDB : 1EEI) et LTB<sub>5</sub> (code PDB : 1LTR) avec l'application de visualisation de structure macromoléculaire 3D Cn3D.

L'échantillon protéique est ensuite soumis à différents traitements thermiques de trois heures pour des températures allant de 60 °C à 180 °C. Après chaque traitement thermique l'échantillon est refroidi à -80 °C où la température est stabilisée pendant trente minutes avant de remonter à la température du prochain traitement thermique (Fig. 3-16). A la fin de chaque traitement thermique, les données de pertes diélectriques sont prélevées jusqu'à la fin des trente minutes à -80 °C et représentées dans les courbes de refroidissement (exemple : R 80 °C). Les courbes de refroidissement seront étudiées en fonction de la température à différentes fréquences. Quand un processus de dynamique moléculaire est détecté, un pic de relaxation apparaît dans les spectres de pertes diélectriques (Fig. 3-12).

### 7.1.1 Contrôle de l'échantillon protéique

La sous-unité B pentamérique de l'entérotoxine thermolabile (LTB<sub>5</sub>) est testée dans le confinement à 40 nm de diamètre et 10 µm de long. Aucun processus de relaxation n'est détecté pour la toxine LTB<sub>5</sub> à la fréquence de 10<sup>3</sup> Hz après des traitements thermiques à 60 °C, comme pour les contrôles du solvant et de la membrane vide (Fig. 7-3).

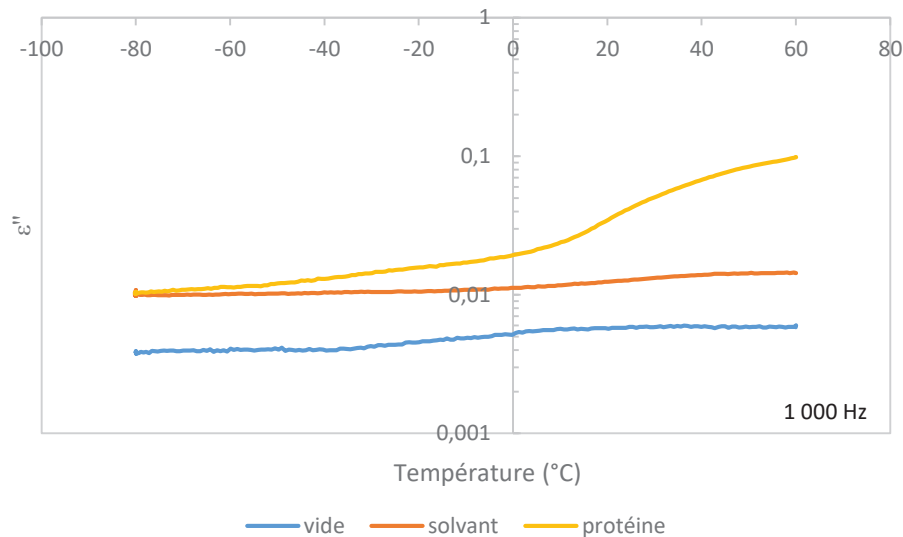


Figure 7-3.  $LTB_5$  dans le confinement à 40 nm de diamètre et 10  $\mu\text{m}$  de long à la fréquence de  $10^3$  Hz après un deuxième traitement thermique à 60 °C. La courbe bleue correspond au test de la membrane vide, la courbe orange au contrôle avec le solvant et la courbe jaune à l'échantillon protéique. 200  $\mu\text{L}$  de solution protéique composée de 2  $\mu\text{L}$  de  $LTB_5$  à une concentration de 2,7 mg/ml et de 212  $\mu\text{L}$  d'eau déionisée, soit une concentration de 0,025 mg/ml est déposée sur la nanomembrane. L'échantillon de solvant est composé de 2  $\mu\text{L}$  de PBS 1X et de 212  $\mu\text{L}$  d'eau déionisée.

L'absence de processus de relaxation pour  $LTB_5$  est reproductible si deux traitements thermiques à 60 °C sont réalisés dans la même expérience (Fig. 7-4a) mais aussi si une deuxième expérience indépendante est réalisée (Fig. 7-4b). Contrairement à  $LTB_5$  après le premier traitement thermique à 60 °C un processus de relaxation associé au solvant est détecté (symbole flèche) pour la courbe de refroidissement mais pas pour la courbe de chauffage, ni après le deuxième traitement thermique à 60 °C (Fig. 7-4c). La reproduction du signal diélectrique lors des deux traitements thermiques à 60 °C indique que l'eau libre est totalement évaporée de l'échantillon protéique après les trois premières heures à 60 °C, ce qui n'est pas le cas pour l'échantillon avec le solvant. La présence d'un processus de relaxation pour le solvant mais pas pour la protéine indique qu'un signal spécifique de la protéine mesure la dynamique moléculaire de la toxine. De plus, les signaux du solvant et de la protéine ne sont pas additifs.

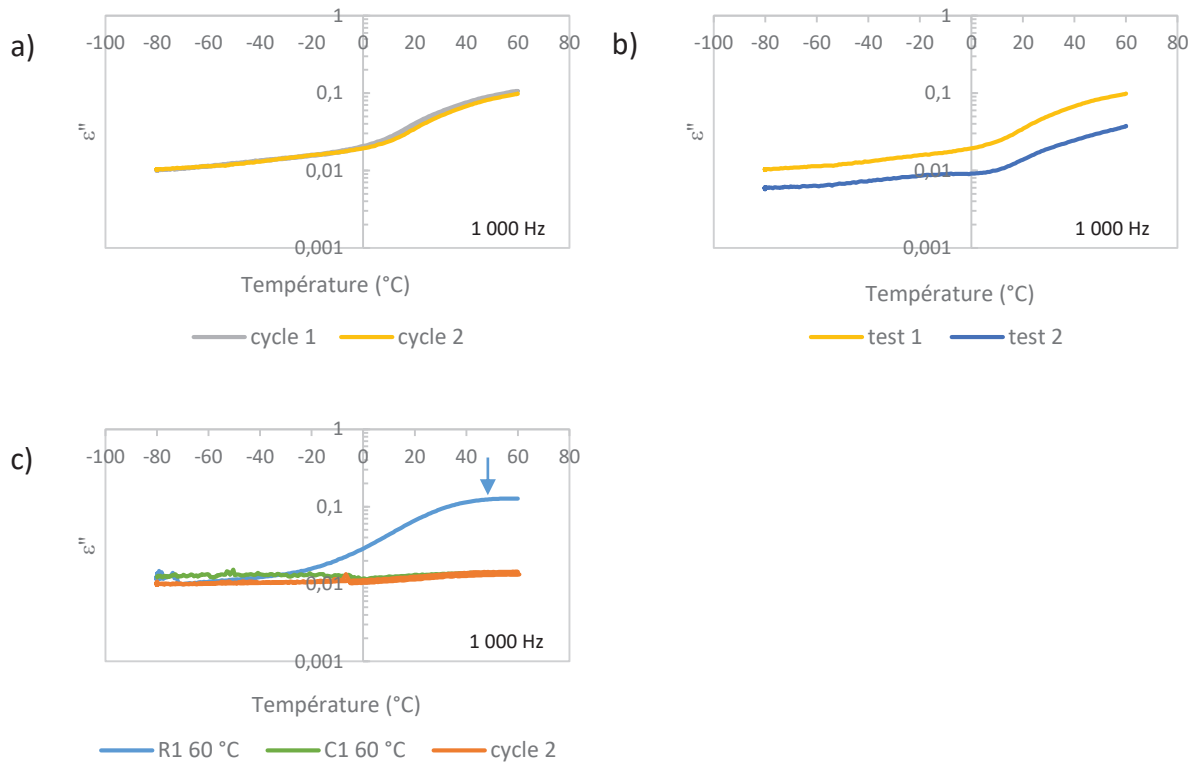


Figure 7-4. Analyse de  $LTB_5$  dans le confinement à 40 nm de diamètre et 10  $\mu\text{m}$  de long à la fréquence de  $10^3$  Hz. a) Analyse du comportement de  $LTB_5$  après deux traitements thermiques à 60 °C. b) Comparaison du signal diélectrique après le deuxième traitement thermique à 60 °C pour deux expériences différentes. c) Superposition des courbes après deux traitements thermiques pour le contrôle avec l'échantillon de solvant.

Cette absence de processus de relaxation après les traitements thermiques à 60 °C est valable pour toute la gamme de fréquence étudiée (Fig. 7-5).

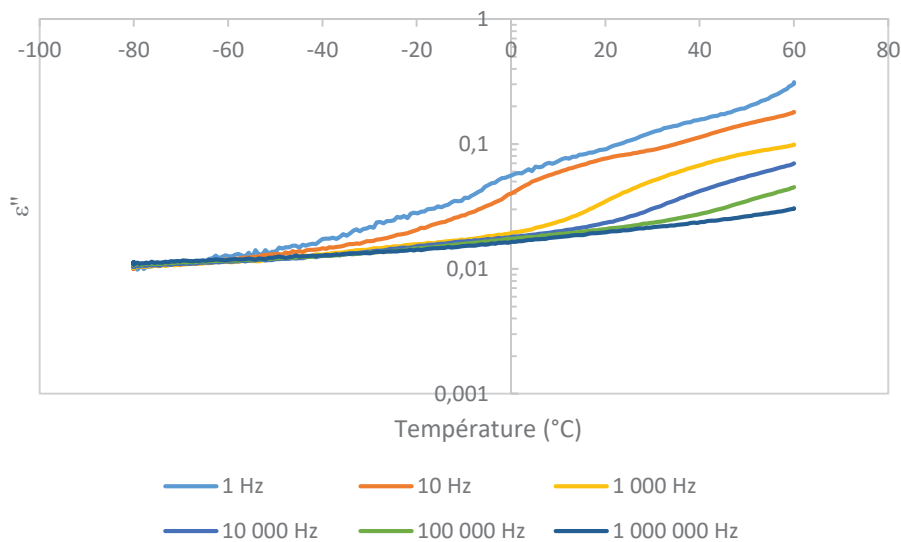


Figure 7-5. Analyse de  $LTB_5$  à différentes fréquences après le deuxième traitement thermique à 60 °C.

Si la température des traitements thermiques augmente à 80 °C, des processus de relaxation commencent à apparaître dans les basses fréquences (Fig. 7-6). A la fréquence de 1 Hz, deux processus de relaxation sont détectés, le premier processus de relaxation (symbole rond) possède une position du pic de relaxation en température à 18 °C contre une température de 48 °C pour le deuxième pic de relaxation (symbole triangle). Le deuxième processus de relaxation n'est pas détecté aux fréquences supérieures alors que le premier processus de relaxation est détecté jusqu'à la fréquence de 10<sup>3</sup> Hz. La détection de processus de relaxation après le traitement thermique à 80 °C suggère que des mouvements de dipôles collectifs dues au commencement du dépliement thermique de LTB<sub>5</sub> et à la perte d'interactions atomiques et moléculaires sont propagés à plus grande échelle.

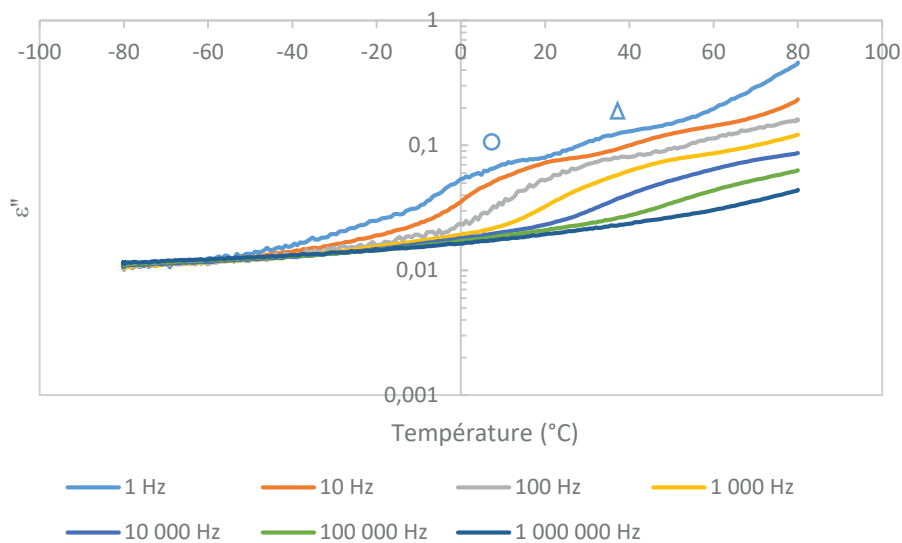


Figure 7-6. Analyse de LTB<sub>5</sub> à différentes fréquences après le traitement thermique à 80 °C.

### 7.1.2 Analyse du comportement de LTB<sub>5</sub> en fonction de la température

Si les mêmes traitements thermiques que pour les chapitres précédents sont réalisés à la fréquence de 10<sup>3</sup> Hz, un processus de relaxation au minimum est détecté jusqu'au traitement thermique à 180 °C pour la toxine (Fig. 7-7). Pour rappel, après le deuxième traitement thermique à 60 °C, aucun processus de relaxation du solvant n'est détecté (Fig. 7-4c et Fig. 7-7). Le premier processus de relaxation (symbole rond) possède une position de son pic de relaxation en température à 57 °C après les traitements thermiques à 80 °C et 100 °C. Après le traitement thermique à 140 °C, la position de ce processus de relaxation est décalée à 76 °C. De plus, un troisième processus de relaxation (symbole étoile) apparaît à 36 °C. Après le traitement thermique à 180 °C, la position du premier processus de relaxation en température est décalée à 62 °C, et un deuxième processus de relaxation (symbole triangle) apparaît avec une position du pic de relaxation à 136 °C.

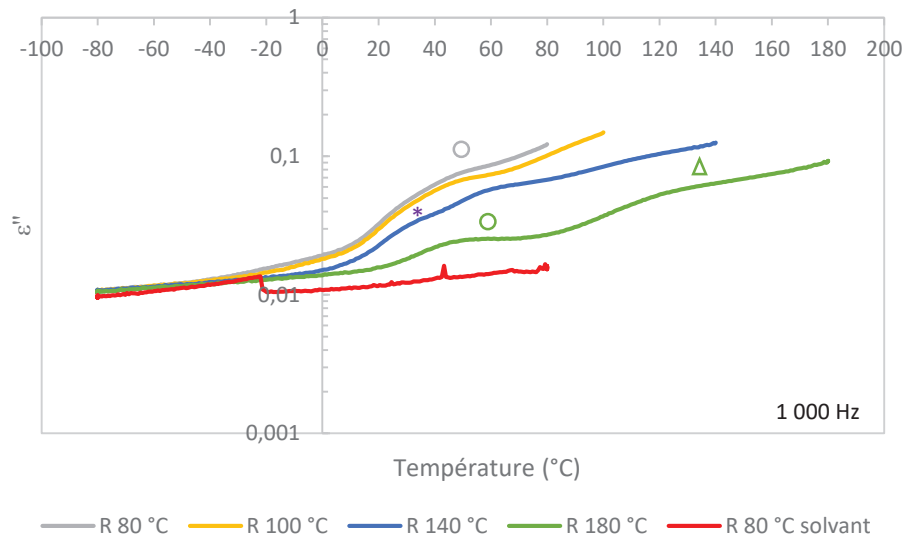


Figure 7-7. Etude du comportement de  $LTB_5$  à la fréquence de  $10^3$  Hz après différents traitements thermiques dans le confinement de 40 nm de diamètre et  $10\ \mu\text{m}$  de long.

### 7.1.3 Analyse du comportement de $LTB_5$ en fonction de la fréquence

Pour rappel, les processus de relaxation commencent à être détectés après le traitement thermique à 80 °C (Fig. 7-5 et Fig. 7-6). Le premier processus de relaxation (symbole rond) est détecté de 1 Hz à  $10^3$  Hz après le traitement thermique à 80 °C, de 1 Hz à  $10^4$  Hz après le traitement thermique à 100 °C, de 1 Hz à  $10^5$  Hz après le traitement à 140 °C et de 1 Hz à  $10^6$  Hz après celui à 180 °C (Fig. 7-8a à 7-8f). Le deuxième processus de relaxation (symbole triangle) est détecté dans les basses fréquences après les traitements à 80 °C (1 Hz), 100 °C et 140 °C (1 Hz à  $10^2$  Hz) et de 1 Hz à  $10^3$  après le traitement thermique à 180 °C. Le troisième processus de relaxation (symbole étoile) est détecté dans les basses et hautes fréquences après le traitement thermique à 140 °C (1 Hz à  $10^4$  Hz).

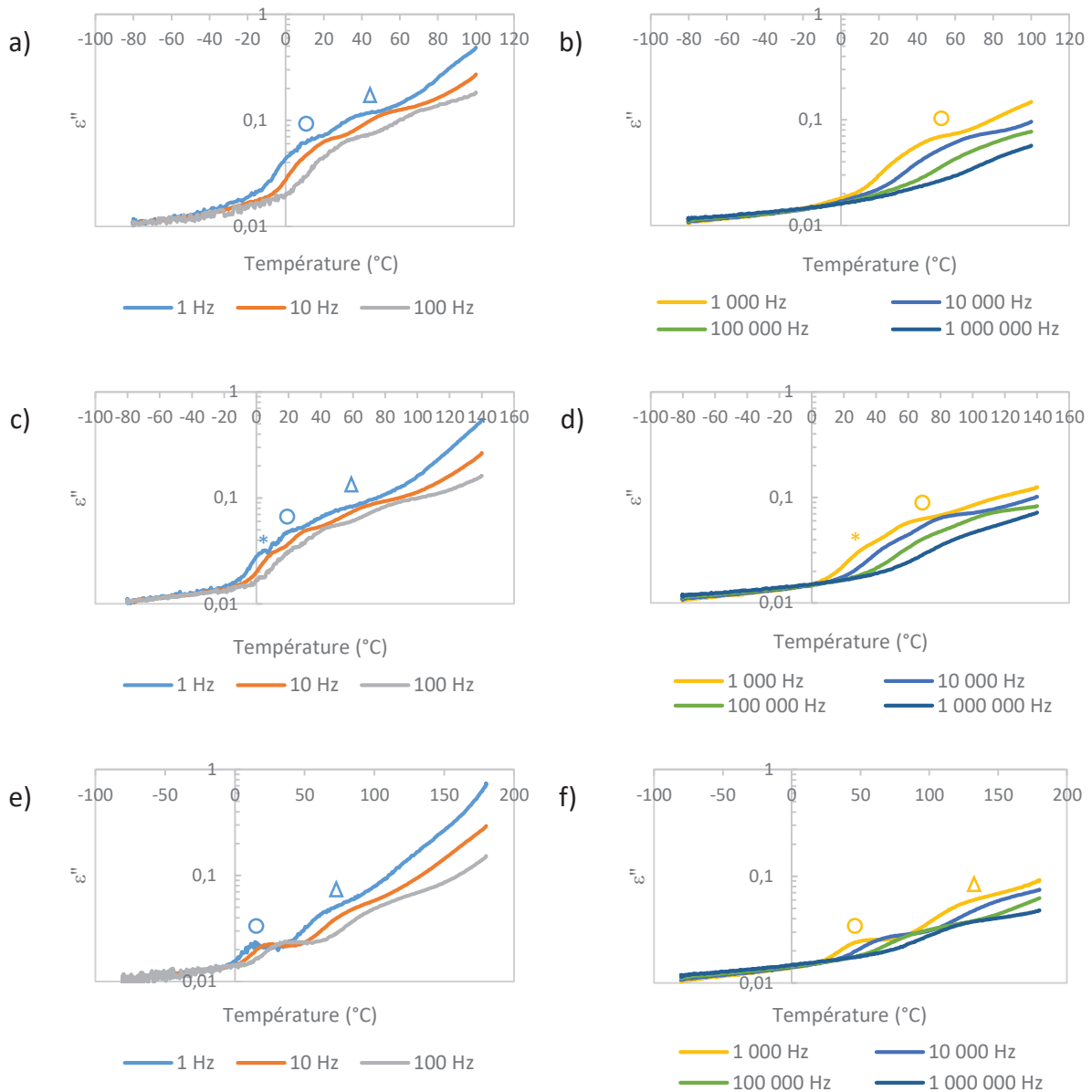


Figure 7-8. Analyse de la dépendance en fréquence des processus de relaxation de  $LTB_5$  dans le confinement de 40 nm de diamètre et 10  $\mu$ m de long après différents traitements thermiques. a) Superposition des courbes de refroidissement de 1 Hz à  $10^2$  Hz après le traitement thermique à 100 °C ( $R$  100 °C). b) Superposition des courbes  $R$  100 °C de  $10^3$  Hz à  $10^6$  Hz. c) Superposition des courbes  $R$  140 °C de 1 Hz à  $10^2$  Hz. d) Superposition des courbes  $R$  140 °C de  $10^3$  Hz à  $10^6$  Hz. e) Superposition des courbes  $R$  180 °C de 1 Hz à  $10^2$  Hz. f) Superposition des courbes  $R$  180 °C de  $10^3$  Hz à  $10^6$  Hz.

#### 7.1.4 Comparaison de deux variants de séquence : $LTB_5$ et $CtxB_5$

La combinaison du BDS et du nanoconfinement permet de détecter des processus de relaxation pour les deux toxines :  $CtxB_5$  et  $LTB_5$ . Ces deux toxines possèdent une identité de séquence de 83,5 %, des structures 3D superposables avec des perturbations atomiques locales (RMS : 0,55 Å) et des fonctions similaires. La comparaison des processus de relaxation obtenus pour ces deux toxines après les différents traitements thermiques, permettra de déterminer si le BDS distingue des populations conformationnelles différentes après l'enrichissement et/ou la sélection via le nanoconfinement de 40 nm de diamètre et 10  $\mu$ m de long. Pour simplifier la comparaison le premier processus de relaxation détecté pour  $CtxB_5$  se

nommera  $1P_C$ , celui de  $LTB_5$   $1P_L$  et ainsi de suite pour les deux autres processus de relaxation détectés.

Après le premier traitement thermique à  $60\text{ }^\circ\text{C}$ ,  $CtxB_5$  possède deux pics de relaxation aux positions respectives de  $-14\text{ }^\circ\text{C}$  pour  $1P_C$  (symbole rond) et de  $30\text{ }^\circ\text{C}$  pour  $2P_C$  (symbole triangle) à la fréquence de  $1\text{ Hz}$  ; un pic de relaxation  $1P_C$  avec une position en température à  $13\text{ }^\circ\text{C}$  à la fréquence de  $10^3\text{ Hz}$ , et aucun processus de relaxation à la fréquence de  $10^6\text{ Hz}$  (Fig. 4-8). Pour le même traitement thermique aucun processus de relaxation n'est détecté à ces fréquences pour la toxine  $LTB_5$  (Fig. 7-5). Ces résultats montrent que la toxine  $LTB_5$ , contrairement à la toxine  $CtxB_5$ , n'est pas encore assez perturbée après les traitements thermiques à  $60\text{ }^\circ\text{C}$  (Fig. 7-9) pour pouvoir détecter des mouvements collectifs dipolaires de la microseconde à la seconde ( $10^6\text{ Hz}$  à  $1\text{ Hz}$ ). L'absence de dynamique moléculaire pour  $LTB_5$  dans la gamme de fréquence étudiée après les traitements thermiques à  $60\text{ }^\circ\text{C}$  suggère qu'à l'état natif, les protéines ont principalement des mouvements de chaînes latérales à l'échelle de la nanoseconde ( $10^9\text{ Hz}$ ) et donc indétectables dans la gamme de fréquence de nos expériences (3). Les fréquences  $1\text{ Hz}$  et  $10^3\text{ Hz}$  seront utilisées pour la comparaison des processus de relaxation pour les deux toxines après les différents traitements thermiques.

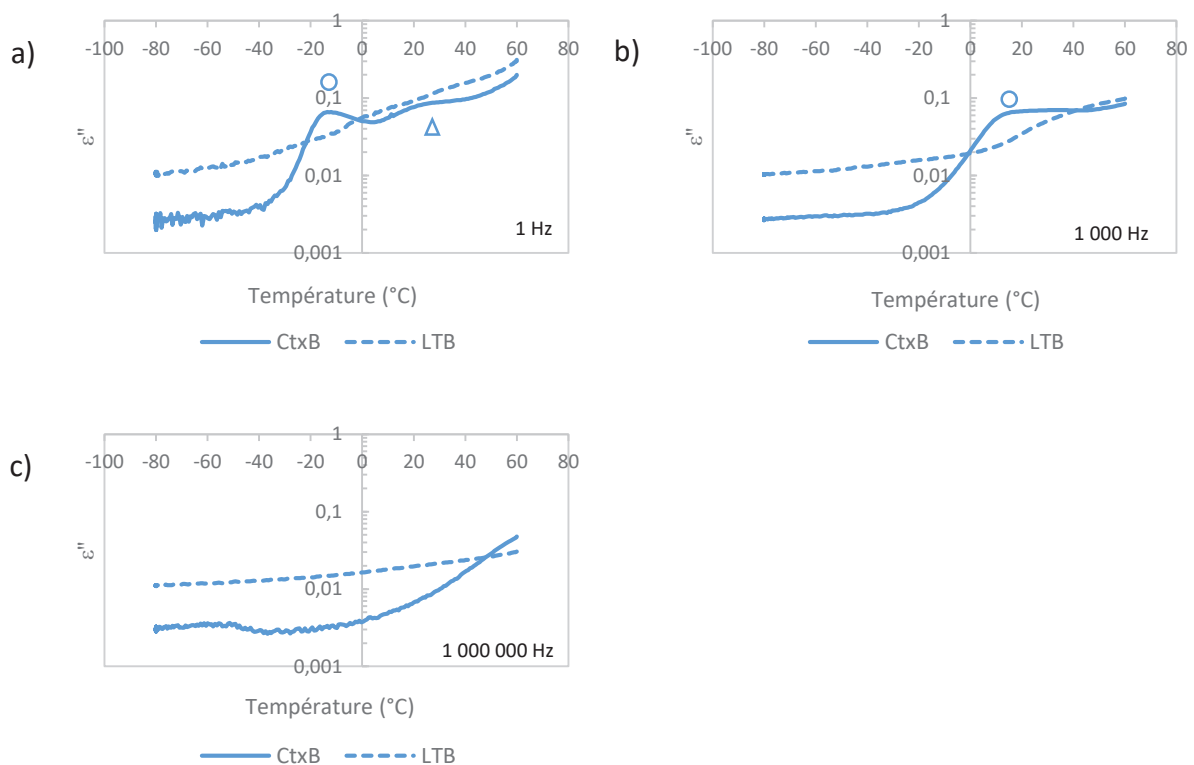


Figure 7-9. Comparaison de la réponse diélectrique pour  $CtxB_5$  et  $LTB_5$  dans le confinement à  $40\text{ nm}$  de diamètre et  $10\text{ }\mu\text{m}$  de long après le traitement thermique à  $60\text{ }^\circ\text{C}$ . a) Comparaison à la fréquence de  $1\text{ Hz}$ . b) Comparaison à  $10^3\text{ Hz}$ . c) Comparaison à  $10^6\text{ Hz}$ .

Après le traitement thermique à  $80\text{ }^\circ\text{C}$  deux processus de relaxation sont détectés à la fréquence de  $1\text{ Hz}$  pour la toxine  $LTB_5$  (Fig. 7-6). Les positions des pics de relaxation en température sont à  $18\text{ }^\circ\text{C}$  pour le premier processus de relaxation ( $1P_L$ ) et à  $48\text{ }^\circ\text{C}$  pour le deuxième ( $2P_L$ ). A cette même fréquence, la toxine  $CtxB_5$  possède les mêmes positions en

température que les pics de relaxation détectés après le traitement thermique à 60 °C, soit -14 °C pour 1Pc et 30 °C pour 2Pc. A la fréquence de 10<sup>3</sup> Hz, LTB<sub>5</sub> détecte uniquement le premier pic de relaxation avec une position à 57 °C tandis que CtxB<sub>5</sub> possède toujours le premier processus de relaxation à 13 °C (1Pc) mais un pic supplémentaire correspondant à un troisième processus de relaxation commence à être détecté à une température de 35 °C (3Pc) (Fig. 7-10).

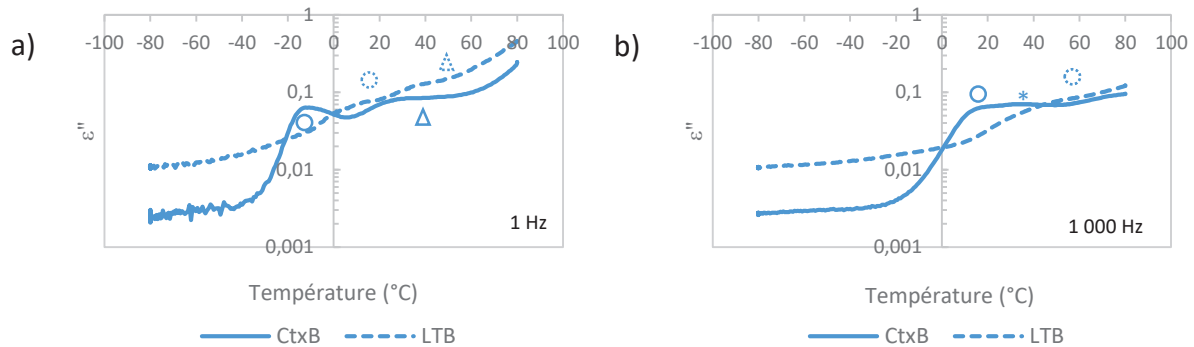


Figure 7-10. Comparaison de la réponse diélectrique pour CtxB<sub>5</sub> et LTB<sub>5</sub> dans le confinement à 40 nm de diamètre et 10 μm de long après le traitement thermique à 80 °C. a) Comparaison à la fréquence de 1 Hz. b) Comparaison à 10<sup>3</sup> Hz.

La position du premier pic de relaxation en température pour les deux toxines (1Pc et 1P<sub>L</sub>, symbole rond) est plus faible pour la toxine CtxB<sub>5</sub> (-14 °C vs 18 °C à la fréquence de 1 Hz et 13 °C vs 57 °C à la fréquence de 10<sup>3</sup> Hz) dans les deux fréquences étudiées. Ces résultats suggèrent que deux zones différentes de toxines et donc des dipôles différents sont perturbés thermiquement dans 1Pc et 1P<sub>L</sub>. Alternativement, les mêmes dipôles peuvent être détectés et les mêmes zones de toxine déstabilisées mais l'environnement autour des dipôles serait différent entraînant une déstabilisation thermique de la zone perturbée à différentes températures pour les deux toxines.

Le deuxième pic de relaxation détecté (symbole triangle) à la fréquence de 1 Hz possède une position en température à 48 °C pour LTB<sub>5</sub> et à 30 °C pour CtxB<sub>5</sub>. Toutefois, le processus de relaxation 2Pc possède un plateau de 30 °C à 48 °C. Ces positions similaires pour 2Pc et 2P<sub>L</sub> suggèrent que ces deux processus de relaxation surveillent les mêmes dipôles, et sondent les mêmes zones de toxines thermiquement déstabilisées mais avec des conséquences différentes sur les conformations des toxines. Ce processus est attribué à des intermédiaires d'assemblage et/ou des monomères repliés pour 2Pc, où l'interface I<sub>3</sub> est sondée (91,124). Ce qui est cohérent avec les données de LTB<sub>5</sub>. En effet, le processus de relaxation 2P<sub>L</sub> est le processus de relaxation le plus déstabilisé comme le démontre la position de ce processus de relaxation en température plus élevée que pour les deux autres processus de relaxation détectés dans LTB<sub>5</sub>. La présence d'un plateau pour 2Pc suggère que la toxine CtxB<sub>5</sub> adopte plus de conformations hétérogènes que LTB<sub>5</sub>.

A la fréquence de 10<sup>3</sup> Hz, un troisième pic de relaxation est présent pour CtxB<sub>5</sub> (symbole étoile). La position de ce pic de relaxation (3Pc) en température est plus faible que celui de 1P<sub>L</sub> à la même fréquence (35 °C pour 3Pc et 57 °C pour 1P<sub>L</sub>) suggérant comme



précédemment que deux zones de toxines différentes sont perturbées ou que les mêmes dipôles sont perturbés mais avec des environnements dipolaires différents entraînant une déstabilisation thermique différente.

Le troisième processus de relaxation (symbole étoile) de CtxB<sub>5</sub> devient mieux défini à la fréquence de 10<sup>3</sup> Hz à mesure que les traitements thermiques augmentent. De plus, ce processus de relaxation est détecté après le traitement thermique à 140 °C à la fréquence de 1 Hz (Fig. 4-10a). Pour la toxine LTB<sub>5</sub>, un troisième processus de relaxation est aussi détecté après le traitement thermique à 140 °C aux fréquences de 1 Hz et 10<sup>3</sup> Hz. La position de 3P<sub>c</sub> en température après le traitement thermique à 140 °C est à 10 °C pour la fréquence de 1 Hz et à 59 °C pour la fréquence de 10<sup>3</sup> Hz. La position de 3P<sub>L</sub> en température après le même traitement thermique est à 5 °C pour la fréquence de 1 Hz et à 36 °C pour la fréquence de 10<sup>3</sup> Hz (Fig. 7-11). Ces résultats suggèrent comme précédemment que des zones différentes sont sondées ou que la même zone est sondée mais avec un environnement dipolaire différent.

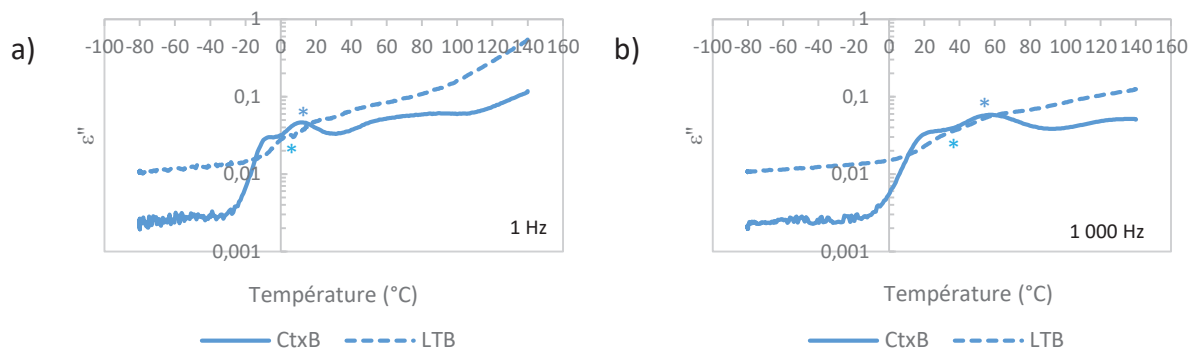


Figure 7-11. Comparaison de la réponse diélectrique pour CtxB<sub>5</sub> et LTB<sub>5</sub> dans le confinement à 40 nm de diamètre et 10 μm de long après le traitement thermique à 140 °C. a) Comparaison à la fréquence de 1 Hz. b) Comparaison à 10<sup>3</sup> Hz.

En conclusion à ce stade, nous supposons que les processus de relaxation 2P<sub>c</sub> et 2P<sub>L</sub> détectent la même zone de toxine déstabilisée, l'interface I<sub>3</sub>. Toutefois, nous ne savons pas avec certitude si les processus de relaxation 1P<sub>c</sub> et 1P<sub>L</sub>, 3P<sub>c</sub> et 3P<sub>L</sub> même 3P<sub>c</sub> et 1P<sub>L</sub> sondent les mêmes zones ou des zones différentes.

La position du pic de relaxation 1P<sub>L</sub> en température (57 °C) après le traitement thermique à 80 °C est proche de la position du pic de relaxation 3P<sub>c</sub> en température (59 °C) après le traitement thermique à 140 °C à la fréquence de 10<sup>3</sup> Hz. Cela suggère que les dipôles thermiques déstabilisés après le traitement thermique à 140 °C pour CtxB<sub>5</sub> et détectés par 3P<sub>c</sub> sont déjà déstabilisés thermique à 80 °C dans LTB<sub>5</sub> et détectés par 1P<sub>L</sub>. Ainsi, les dipôles détectés après le traitement thermique à 80 °C par 1P<sub>L</sub> ne seraient détectés qu'après le traitement thermique à 140 °C par 3P<sub>c</sub> dans CtxB<sub>5</sub>, suggérant que les zones de toxines sondées par ces dipôles sont thermiquement plus résistantes dans CtxB<sub>5</sub> que dans LTB<sub>5</sub>. A 1 Hz, la position en température de 1P<sub>L</sub> est de 18 °C et celle de 3P<sub>c</sub> à 10 °C.

La position du pic de relaxation 3P<sub>c</sub> en température après le traitement thermique à 80 °C est de 35 °C, alors que la position de 3P<sub>L</sub> en température après le traitement thermique à 140 °C est de 36 °C à la fréquence de 10<sup>3</sup> Hz, suggérant que les mêmes dipôles sont

déstabilisés dans les deux toxines mais à une température plus élevée dans LTB<sub>5</sub> que dans CtxB<sub>5</sub>. A la fréquence de 1Hz, 3Pc n'est pas détecté après le traitement thermique à 80 °C, le processus de relaxation est dominé par la conductivité du signal.

La position de 1Pc en température ne possède pas d'équivalent pour la toxine LTB<sub>5</sub> après les différents traitements thermiques aux fréquences de 1 Hz et 10<sup>3</sup> Hz, suggérant que la zone sondée par 1Pc n'est pas déstabilisée dans LTB<sub>5</sub>. Cette zone est associée à une population de conformation de pentamère de type natif, où l'interface I<sub>1</sub> est principalement déstabilisée dans les basses fréquences (91,124). Après le traitement thermique à 180 °C, la position du processus de relaxation 1Pc en température (10 °C) est similaire à la position du processus de relaxation 3P<sub>L</sub> après le traitement thermique à 140 °C (5 °C), suggérant que certains dipôles communs pourraient être déstabilisés après des traitements thermiques à haute température. Par conséquent, malgré des dépliements différents, les deux toxines tendent à se déplier dans des conformations similaires après les traitements thermiques à haute température.

En conclusion, après le traitement thermique à 60 °C, CtxB<sub>5</sub> possède deux processus de relaxation à 1 Hz et un seul processus à 10<sup>3</sup> Hz alors que LTB<sub>5</sub> ne possède aucun pic de relaxation indiquant que les zones dans la toxine LTB<sub>5</sub> sont plus stables. Les pics de relaxation 2Pc et 2P<sub>L</sub> détectent la même zone de la toxine déstabilisée thermiquement mais avec des impacts différents sur la conformation de la toxine (plateau pour CtxB<sub>5</sub>). La position du premier processus de relaxation en température dans LTB<sub>5</sub> (1P<sub>L</sub> après le traitement thermique à 80 °C) est similaire à celle du processus de relaxation 3Pc après le traitement thermique à 140 °C, indiquant que la toxine est plus stable dans CtxB<sub>5</sub>. La position du processus de relaxation 3Pc en température après le traitement thermique à 80 °C est similaire à celle du processus de relaxation 3P<sub>L</sub> après le traitement thermique à 140 °C, indiquant que les zones sondées sont plus stables dans LTB<sub>5</sub>.

L'étude des interactions atomiques entre les acides aminés est réalisée pour déterminer les dipôles pouvant expliquer la différence de signal diélectrique entre les deux toxines. Les interactions atomiques sont déterminées à partir des structures à rayon X avec le code PDB, 1EEI pour CtxB<sub>5</sub> et 1LTR pour LTB<sub>5</sub>. Comme le pic 2Pc est attribué à une population d'intermédiaires d'assemblage (tétramères, trimères, dimères) et/ou de monomères repliés les interactions atomiques impliquant les acides aminés des interfaces sont regardés. La toxine LTB<sub>5</sub> possède les mêmes interfaces que la toxine CtxB<sub>5</sub> à quelques résidus près.

- L'interface I<sub>1</sub> correspondant au domaine N-terminal et est divisée en deux interfaces : I<sub>1a</sub> qui comprend les résidus 1 à 3 de la chaîne E en interactions avec les résidus 92 et 93 de la chaîne adjacente F / I<sub>1b</sub> qui correspond à l'interaction des résidus 1 à 12 de la chaîne E en interactions avec les résidus 28 à 39 de la chaîne F.
- L'interface I<sub>2</sub> correspondant à l'interface principale qui comprend l'interaction entre le brin beta numéro 2 (résidu 25 à 32) et le brin beta numéro 6 (résidu 95 à 103) sur des monomères adjacents : I<sub>2a</sub> comprend les résidus 25 à 32 de la chaîne E en interactions avec les résidus 88, et de 96 à 103 de la chaîne adjacente D / I<sub>2b</sub> comprend les résidus 28 à 39 de la chaîne E en interactions avec les résidus 57 à 73 de la chaîne D / I<sub>2c</sub>

correspond aux résidus 101 à 103 de la chaîne E en interactions avec les résidus 73 à 79 de la chaîne adjacente F.

- L'interface  $I_3$  de l'hélice  $\alpha$  centrale (résidu 60 à 78) est une interface tripartite composée des résidus 61 à 77 de la chaîne E en interactions avec les résidus 63 à 81 de la chaîne D, et les résidus 67 à 81 de la chaîne E en interactions avec les résidus 66 à 77 de la chaîne F.

Basé sur le nombre d'interactions atomiques et de paires d'acides aminés, les interfaces de la toxine LTB<sub>5</sub> sont classées de la même manière que les interfaces de la toxine CxB<sub>5</sub> : les interfaces  $I_{1a}$  et  $I_{2c}$  sont les plus faibles avec respectivement 50 interactions atomiques, 5 paires d'acides aminés et 65 interactions atomiques et 5 paires d'acides aminés. Les interfaces  $I_{2a}$  et  $I_3$  sont les interfaces les plus fortes avec 166 interactions atomiques et 24 paires d'acides aminés et 210 interactions atomiques et 25 paires d'acides aminés. Par conséquent les interfaces  $I_{1b}$  et  $I_{2b}$  sont dites modérées avec 139 interactions atomiques, 18 paires d'acides aminés et 157 interactions atomiques et 19 paires d'acides aminés. Ces résultats suggèrent que les interactions intermoléculaires ne sont pas responsables des différences observées dans le signal diélectrique des deux toxines.

Pour être visible diélectriquement, il est préférable que les paires d'acides aminés et les interactions atomiques se produisent via des acides aminés chargés. Pour rappel, l'interaction entre deux acides aminés chargés donnera un dipôle ionique et l'interaction entre un acide aminé chargé et un acide aminé non chargé donnera un dipôle induit (Partie 3.2.3).

Dans ce cas, la même interprétation que pour CtxB<sub>5</sub> a aussi lieu : les interfaces  $I_{1a}$  et  $I_{2c}$  ne possèdent pas de dipôles ioniques mais deux dipôles induits, ces deux interfaces sont considérées comme invisibles ou faiblement visibles diélectriquement. De même pour l'interface  $I_{2a}$  car celle-ci ne possède aucun dipôle ionique, mais quatre dipôles induits. Les interfaces  $I_{1b}$  et  $I_{2b}$  possèdent respectivement deux et un dipôle ioniques et quelques dipôles induits, par conséquent ces interfaces devraient être visible diélectriquement. L'interface  $I_3$  est l'interface diélectrique la plus puissante car elle possède 14 dipôles ioniques et des dipôles induits.

La toxine LTB<sub>5</sub> possède un dipôle ionique en moins Lys69E-Lys63D (acide aminé lys en position 69 dans la chaîne E interagit avec l'acide aminé lys en position 63 dans la chaîne D) dans l'interface  $I_3$  par rapport à la toxine CtxB<sub>5</sub>. Inversement, la toxine LTB<sub>5</sub> possède un dipôle ionique en plus dans cette même interface (Lys62E-Lys63D). Les voisinages dipolaires (distance seuil de 5 Å entre les résidus chargés) sont similaires entre les deux toxines à l'exception du résidu lysine 62 qui est chimiquement proche du résidu lysine 63 sur la chaîne adjacente dans LTB<sub>5</sub>, introduisant une force de répulsion qui pourrait expliquer la 'perte' du dipôle impliquant le résidu lysine 63 (de la chaîne D) dans LTB<sub>5</sub> (Table. 7-1). A l'exception de ces dipôles, les deux toxines possèdent les mêmes dipôles ioniques intermoléculaires comme le montre la table 7-1. Pour information, le dipôle Glu36E-Lys63D est aussi présent entre la chaîne E et D dans la toxine CtxB<sub>5</sub>, toutefois ce dipôle n'est pas détecté dans les quatre autres

chaînes, par conséquent il n'est pas pris en compte. Néanmoins, la même interprétation que pour le dipôle Lys69E-Lys63D est applicable pour expliquer la 'perte' de ce dipôle dans LTB<sub>5</sub>.

Force des interfaces	Interfaces	Dipôles ioniques LTB <sub>5</sub>	Dipôles ioniques CtxB <sub>5</sub>
Faibles	I <sub>1a</sub>	-	-
	I <sub>2c</sub>	-	-
Modérées	I <sub>1b</sub>	Ala1E-Arg35F Glu11E-Arg35F	Thr1E-Arg35F Glu11E-Arg35F
	I <sub>2b</sub>	Glu29E-Arg67D	Glu29E-Arg67D
Fortes	I <sub>2a</sub>	-	-
	I <sub>3</sub>	<u>Lys62E-Lys63D</u> Glu66E-Lys63D Glu66E-Arg67D Arg67E-Arg67D Lys69E-Arg67D Asp70E-Arg67D Arg73E-Arg67D Arg73E-asp70D  Arg67E-Glu66F Arg67E-Arg67F Arg67E-Lys69F Arg67E-Asp70F Arg67E-Arg73F Asp70E-Arg73F	Glu66E-Lys63D Glu66E-Arg67D Arg67E-Arg67D <u>Lys69E-Lys63D</u> Lys69E-Arg67D Asp70E-Arg67D Arg73E-Arg67D Arg73E-Asp70D  Arg67E-Glu66F Arg67E-Arg67F Arg67E-Lys69F Arg67E-Asp70F Arg67E-Arg73F Asp70E-Arg73F

Table 7-1. Représentation des dipôles ioniques détectés dans les interfaces de LTB<sub>5</sub> et CtxB<sub>5</sub>. La distance seuil pour former un dipôle est de 5 Å.

L'absence de différence significative entre les dipôles ioniques intermoléculaires des deux toxines suggère que la réponse diélectrique des deux toxines devrait être similaire si uniquement ces dipôles étaient détectés. Or les données expérimentales ne sont pas les mêmes entre les deux toxines, suggérant que les dipôles ioniques intramoléculaires pourraient aussi participer au signal diélectrique.

La toxine LTB<sub>5</sub> possède trois dipôles ioniques intramoléculaires supplémentaires (Ala1-Glu7, Lys81-Glu102, Asp83-Glu102) suite aux mutations de A1T (alanine en position 1 dans LTB<sub>5</sub> est mutée en thréonine dans CtxB<sub>5</sub>), E7D, D83E et E102A comme le montre la table 7-2. Dans la structure PDB de LTB<sub>5</sub>, l'asparagine 103 est mutée en lysine, par conséquent le ou les dipôles formés via ce résidu ne sont pas représentés, toutefois, il est raisonnable de supposer que le dipôle Lys81-Asn103 présent dans la toxine CtxB<sub>5</sub> devrait aussi être présent dans la toxine LTB<sub>5</sub>.

Dipôles ioniques LTB <sub>5</sub>	Dipôles ioniques CtxB <sub>5</sub>
<u>Ala1-Glu7</u>	
Glu7-Glu11	Asp7-Glu11
Asp22-Lys23	Asp22-Lys23
Asp22-Lys43	Asp22-Lys43
Asp22-Lys81	Asp22-Lys81
Lys23-Glu79	Lys23-Glu79
Lys23-Lys81	Lys23-Lys81
Glu29-Arg73	Glu29-Arg73
Lys34-Arg35	Lys34-Arg35
Lys34-Glu36	Lys34-Glu36
Arg35-Glu36	Arg35-Glu36
Glu36-Glu51	Glu36-Glu51
Glu36-Lys69	Glu36-Lys69
Glu51-Lys91	Glu51-Lys91
Asp59-Lys62	Asp59-Lys62
Asp59-Lys63	Asp59-Lys63
Lys62-Lys63	Lys62-Lys63
Lys62-Glu66	Lys62-Glu66
Lys63-Glu66	Lys63-Glu66
Lys63-Arg67	Lys63-Arg67
Glu66-Arg67	Glu66-Arg67
Glu66-Lys69	Glu66-Lys69
Glu66-Asp70	Glu66-Asp70
Arg67-Lys69	Arg67-Lys69
Arg67-Asp70	Arg67-Asp70
Lys69-Asp70	Lys69-Asp70
Lys69-Arg73	Lys69-Arg73
Asp70-Arg73	Asp70-Arg73
Glu79-Lys81	Glu79-Lys81
Lys81-Asp83	Lys81-Glu83
<u>Lys81-Glu102</u>	<i>Lys81-Asn103</i>
Asp83-Lys84	Glu83-Lys84
<u>Asp83-Glu102</u>	

Table 7-2. Représentation des dipôles ioniques intramoléculaires détectés dans LTB<sub>5</sub> et CtxB<sub>5</sub>. La distance seuil pour former un dipôle est de 5 Å.

Le dipôle Ala1-Glu7 devrait renforcer la stabilité intramoléculaire de l'extrémité N-terminale LTB<sub>5</sub>, rendant ainsi l'interface I<sub>1</sub> (interface I<sub>1a</sub> et I<sub>1b</sub>) moins prompt à la dissociation. Les dipôles Lys81-Glu102 et Asp83-Glu102 pourraient renforcer la stabilité intramoléculaire de l'hélice  $\alpha$  centrale (interface I<sub>3</sub>) mais aussi de l'extrémité C-terminale (interface I<sub>2c</sub>) en réduisant leurs mobilités.

Grâce au dipôle Ala1-Glu7, la déstabilisation thermique se propagerait uniquement de l'interface I<sub>1a</sub> (acide aminé 1 à 3 de la chaîne E) vers les résidus thréonine 92 et proline 93 et donc à la lysine 91 et au dipôle ionique intramoléculaire Lys91-Glu51 pour la toxine LTB<sub>5</sub>. Dans CtxB<sub>5</sub> la même propagation a lieu, toutefois en raison de l'absence du dipôle Ala1-Glu7, la

perturbation se propagerait aussi dans toute l'interface  $I_1$  (interface  $I_{1a}$  et  $I_{1b}$ ) qui perturberait à son tour l'interface  $I_{2b}$ . Par conséquent, les deux toxines devraient posséder des chemins de dépliement différents.

La perturbation de l'acide aminé Glu51 est propagée au dipôle ionique intramoléculaire Glu36-Lys69 via le résidu acide glutamique 36, mais aussi aux dipôles Asp59-Lys62, Asp59-Lys63 via l'histidine 57 en raison des dipôles induits Glu51-His57 et His57-Asp59 présent dans les deux toxines. Ainsi, la déstabilisation de l'extrémité N-terminale (interface  $I_{1a}$ ) conduit à la déstabilisation intramoléculaire de l'hélice alpha centrale (résidu 60 à 78 impliqué dans l'interface  $I_3$ ) et du brin beta numéro 3 (résidu 35 à 42 impliqué dans l'interface  $I_{2b}$ ).

La présence d'une voie de dépliement supplémentaire à celle via le dipôle Lys91-Glu51 dans CtxB<sub>5</sub> (interfaces  $I_{1a}$  et  $I_{1b}$ ) aurait tendance à provoquer la perturbation des interfaces  $I_2$  et  $I_3$  à une température plus basse dans LTB<sub>5</sub>. Cette hypothèse est cohérente avec la détection du pic de relaxation ( $1P_L$ ) après le traitement thermique à 80 °C dans une position en température similaire à celle du pic de relaxation  $3P_c$  après le traitement thermique à 140 °C. Ces deux processus de relaxation correspondraient à l'interface  $I_2$  et plus particulièrement à l'interface  $I_{2b}$ . Cette hypothèse est moins cohérente avec l'interface  $I_3$  attribuée aux processus de relaxation  $2P_L$  et  $2P_c$ . En effet, la position en température du pic de relaxation  $2P_L$  est à 48 °C après le traitement thermique à 80 °C, contre un plateau de 30 °C à 48 °C pour  $2P_c$  après les traitements thermiques à 60 °C et 80 °C, par conséquent, cette interface ne semble pas plus déstabilisée dans LTB<sub>5</sub> que dans CtxB<sub>5</sub>.

La non-déstabilisation de l'interface  $I_3$  à des températures plus faibles pour LTB<sub>5</sub> pourrait s'expliquer par la présence de deux dipôles intramoléculaires Glu102-Lys81 et Glu102-Asp83 supplémentaires dans LTB<sub>5</sub>. Ces deux dipôles pourraient renforcer la stabilité intramoléculaire de l'extrémité C-terminale (interface  $I_{2c}$ ) mais aussi l'interface  $I_3$  (résidu 61 à 81) de LTB<sub>5</sub>, rendant ces zones plus résistantes thermiquement dans LTB<sub>5</sub> que dans CtxB<sub>5</sub>, ou dans le cas de l'interface  $I_3$ , résistante de manière similaire.

La perturbation de l'extrémité C-terminale se propagerait au résidu Asn103 et au dipôle Asn103-Lys81. Néanmoins en raison des deux dipôles supplémentaires impliquant le résidu Glu102, le dipôle Asn103-Lys81 ne sera pas déstabilisé dans LTB<sub>5</sub>, rendant la zone plus résistante thermique que dans CtxB<sub>5</sub>. Toutefois, la perturbation de l'extrémité C-terminale se propage au résidu arginine 73 dans les deux toxines puis aux dipôles Arg73-Glu29, Arg73-Lys69 et Arg73-Asp70 soit à l'interface  $I_3$  et aux interfaces  $I_{2b}$  et  $I_{2a}$ . De plus, dans LTB<sub>5</sub> contrairement à CtxB<sub>5</sub> le résidu acide glutamique 79 participe à l'interface  $I_{2c}$  entraînant ainsi sa perturbation et la propagation aux dipôles Glu79-Lys23 et Glu79-Lys81. La déstabilisation de l'extrémité C-terminale dans LTB<sub>5</sub> devrait conduire à la déstabilisation des contacts 3D entre deux boucles (résidus 79-81 et 22-24) affaiblissant ainsi les contacts entre l'hélice alpha et le brin beta numéro 2 (résidu 25 à 32), affectant donc l'interface  $I_{2a}$ .

En rendant les extrémités N et C terminal plus stables, les trois dipôles ioniques intramoléculaires devraient bloquer le brin  $\beta$  numéro 6 (Ser95 à Asn 103) dans une position plus rigide qui rendrait l'interface principale (interface  $I_{2a}$  impliquant le brin  $\beta$  numéro 2 d'une

chaîne et le brin beta numéro 6 de la chaîne adjacente) plus stable. L'augmentation de la stabilité de l'interface  $I_{2a}$  (résidu 25 à 32 d'une chaîne et 88, 96-103 de la chaîne adjacente) devrait prévenir de la propagation de la déstabilisation thermique à l'interface  $I_{2b}$  (résidu 28 à 39 d'une chaîne). De plus, le résidu Glu29 présents dans les interfaces  $I_{2a}$  et  $I_{2b}$  forme un dipôle ionique intramoléculaire avec le résidu Arg73, prévenant aussi de la déstabilisation de l'interface  $I_3$  (résidu 61 à 81).

En conclusion, les deux toxines LTB<sub>5</sub> et CtxB<sub>5</sub> soit deux protéines avec une identité de séquence de 83,5 %, des structures 3D superposables et des fonctions similaires possèdent des processus de relaxation différents démontrant la capacité du BDS à distinguer l'impact des variants de séquence sur la dynamique des protéines.

La présence de trois dipôles supplémentaires dans LTB<sub>5</sub> entraîne l'augmentation de la stabilité des interfaces  $I_{1a}$ ,  $I_{1b}$  et  $I_{2c}$ . Le premier pic de relaxation détecté dans CtxB<sub>5</sub> est attribué aux interfaces  $I_{1a}$ ,  $I_{1b}$ ,  $I_{2a}$  et  $I_{2c}$ , le deuxième pic aux interfaces  $I_3$  et le troisième pic aux interfaces  $I_{2b}$ . Dans le cas de LTB<sub>5</sub>, le premier pic de relaxation est attribué aux interfaces  $I_{2b}$ , le deuxième aux interfaces  $I_3$  et le troisième aux interfaces  $I_{2a}$ .

Le BDS permet d'aller dans le détail de la dynamique locale impliquée dans les mécanismes de dépliement sans utiliser de procédures de mutagenèse ou de marquage. Il montre comment les perturbations locales associées aux mutations propagent des impacts à plus grande échelle sur la dynamique de la protéine entière.

Les processus de relaxation 2P<sub>c</sub> et 2P<sub>L</sub> possèdent une position des pics de relaxation en température similaire pour les deux toxines, suggérant la déstabilisation des mêmes dipôles. La présence d'un plateau pour 2P<sub>c</sub> et non pour 2P<sub>L</sub> suggère que la toxine CtxB<sub>5</sub> adopte plus de conformations hétérogènes que LTB<sub>5</sub>. Le processus de relaxation 2P<sub>c</sub> est attribué aux intermédiaires d'assemblage et/ou aux monomères repliés. Or si des intermédiaires d'assemblage sont détectés dans LTB<sub>5</sub>, il serait raisonnable de supposer qu'un plateau devrait aussi être présent, ce qui n'est pas le cas. L'absence d'intermédiaire d'assemblage pour la toxine LTB<sub>5</sub> pourrait s'expliquer de deux manières : la protéine est plus stable est donc peu d'intermédiaires d'assemblage sont détectés ou encodage du cheminement de dépliement dans la séquence avec passage de l'état pentamérique à l'état monomérique sans la formation d'intermédiaire d'assemblage.

Le chapitre 5 a montré que la population conformationnelle CtxB<sub>5</sub> enrichie et/ou sélectionnée possédait des stabilités thermiques différentes suivant la longueur du confinement. La population conformationnelle 1P<sub>c</sub> disparaissait plus tôt dans le confinement à 40 nm de diamètre et 2  $\mu$ m de long que dans les autres confinements testés dont le confinement à 40 nm de diamètre et 10  $\mu$ m de long. Il est raisonnable de supposer que si la toxine LTB<sub>5</sub> est incorporée dans des nanopores à 40 nm de diamètre et 2  $\mu$ m de long, la protéine pourrait aussi avoir une stabilité thermique différente et ainsi permettre une meilleure détection des intermédiaires d'assemblage caractérisés par un plateau si ceux-ci sont présents.

Si un échantillon de toxine LTB<sub>5</sub>, est testé dans le confinement à 40 nm de diamètre et 2  $\mu$ m de long à la fréquence de 1 Hz après un traitement thermique à 60 °C, un premier pic de



relaxation (symbole rond) est détecté avec une position de  $-2\text{ }^{\circ}\text{C}$ , un deuxième processus (symbole triangle) commence à se dessiner à une température de  $23\text{ }^{\circ}\text{C}$ . A cette fréquence et à cette température le solvant possède aussi un pic de relaxation (symbole flèche) à une température de  $-17\text{ }^{\circ}\text{C}$  (Fig. 7-12).

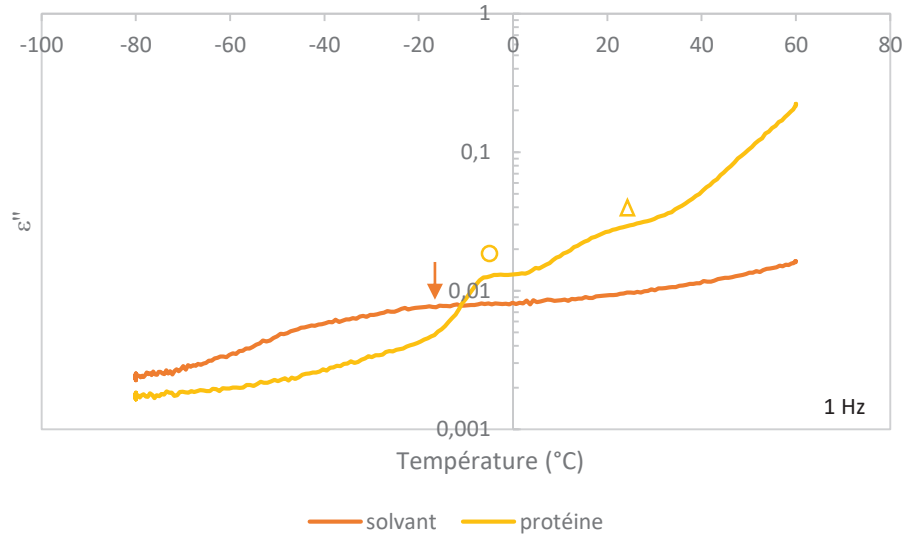


Figure 7-12. Analyse de  $\text{LTB}_5$  dans le confinement à  $40\text{ nm}$  de diamètre et  $2\text{ }\mu\text{m}$  de long à la fréquence de  $1\text{ Hz}$  après le deuxième traitement thermique à  $60\text{ }^{\circ}\text{C}$ .  $200\text{ }\mu\text{L}$  de solution composé de  $2\text{ }\mu\text{L}$  de toxine à  $2,7\text{ mg.ml}$  et  $1\text{ }056\text{ }\mu\text{L}$  d'eau déionisée à une concentration de  $0,0051\text{ mg.ml}$  est déposée sur la nanomembrane. Le contrôle du solvant est composé de  $2\text{ }\mu\text{L}$  de PBS 1X et de  $1\text{ }056\text{ }\mu\text{L}$  d'eau déionisée.

Si la température des traitements thermiques augmente, la position du premier processus de relaxation (symbole rond) en température ne change pas jusqu'au traitement thermique à  $100\text{ }^{\circ}\text{C}$ . Après les traitements thermiques à  $140\text{ }^{\circ}\text{C}$  et  $180\text{ }^{\circ}\text{C}$ , la position de ce processus de relaxation en température est décalée à  $4\text{ }^{\circ}\text{C}$  puis à  $28\text{ }^{\circ}\text{C}$ . La position du deuxième processus de relaxation (symbole triangle) en température est de  $33\text{ }^{\circ}\text{C}$  après les traitements thermiques à  $80\text{ }^{\circ}\text{C}$  et  $100\text{ }^{\circ}\text{C}$  et à  $47\text{ }^{\circ}\text{C}$  après le traitement thermique à  $140\text{ }^{\circ}\text{C}$  (Fig. 7-13). Pour le contrôle du solvant, aucun processus de relaxation n'est détecté après le traitement thermique à  $100\text{ }^{\circ}\text{C}$  (insert Fig. 7-13). Ainsi, les deux processus de relaxation détectés après les différents traitements thermiques pour l'échantillon protéique correspondent bien à la protéine.



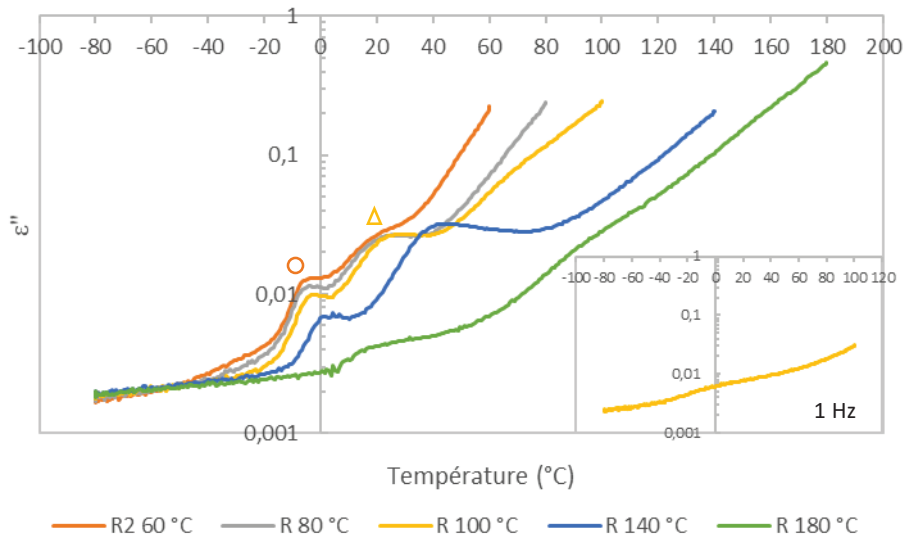


Figure 7-13. Etude de  $LTB_5$  à la fréquence de 1 Hz après différents traitements thermiques dans le confinement à 40 nm de diamètre et 2  $\mu$ m de long. Insert. Courbe de refroidissement à la fréquence de 1 Hz après le traitement thermique à 100 °C pour le solvant.

Le premier processus de relaxation couvre à la fois les basses et hautes fréquences alors que le deuxième processus de relaxation couvre principalement les basses fréquences (Fig. 7-14). La présence d'uniquement deux processus de relaxation quelques soient les fréquences et la température des traitements thermiques associée à la gamme de fréquence et de température de ces processus de relaxation tend à suggérer que le premier processus de relaxation détecté correspond au processus de relaxation  $1P_L$  et le deuxième processus de relaxation à  $2P_L$  détecté dans le confinement à 40 nm de diamètre et 10  $\mu$ m de long. Le processus de relaxation  $2P_L$  associé à une population d'intermédiaires d'assemblage et/ou de monomères repliés dans le confinement à 40 nm de diamètre et 2  $\mu$ m de long est mieux défini que dans le confinement à 40 nm de diamètre et 10  $\mu$ m de long. Toutefois, ce processus de relaxation (symbole triangle) ne semble posséder aucun plateau, comme le montre la forme du pic dans la courbe de refroidissement après le traitement thermique à 140 °C à la fréquence de 1 Hz (Fig. 7-13). Par conséquent, même en diminuant la longueur du confinement aucun plateau n'est présent pour le processus de relaxation  $2P_L$ , suggérant que la toxine  $LTB_5$  ne se déplie pas en intermédiaire d'assemblage. La combinaison du nanoconfinement et de la spectroscopie diélectrique à large bande semble donc capable de suivre les chemins de dépliement des protéines. La différence de chemin de dépliement dans  $LTB_5$  par rapport à  $CtxB_5$  est expliquée par la présence de dipôles ioniques supplémentaires, dont le dipôle à l'extrémité N-terminale (ala1-glu7). Pour valider que cette approche permet de suivre les chemins de dépliement des protéines, une liaison ionique pourrait être ajoutée à l'extrémité N-terminale de  $CtxB_5$  (cross-link). Dans ce cas de figure, les chemins de dépliement devraient être différents pour  $CtxB_5$  avec et sans liaison supplémentaire pour un même confinement.

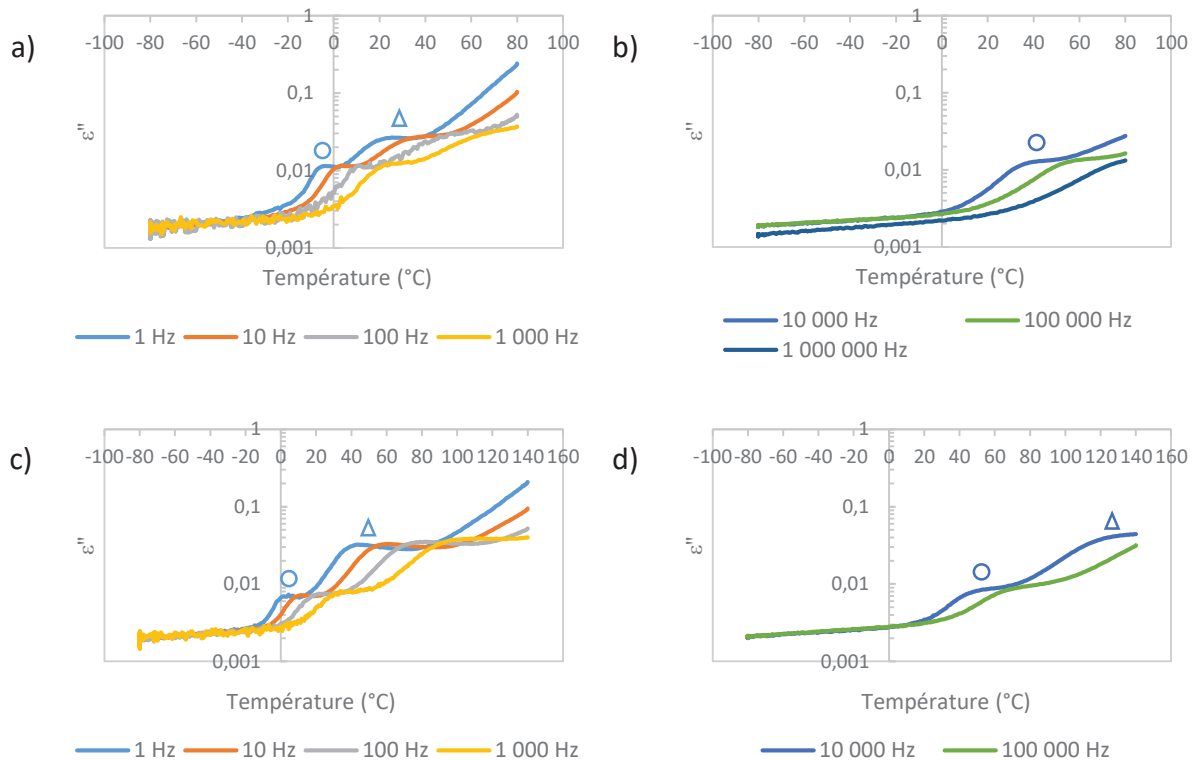


Figure 7-14. Analyse de la dépendance en fréquence des processus de relaxation de  $LTB_5$  dans le confinement à 40 nm de diamètre et 2  $\mu\text{m}$  de long après différents traitements thermiques. a) Superposition des courbes de 1 Hz à la fréquence  $10^3$  Hz après le traitement thermique à 80 °C (R 80 °C). b) Superposition des courbes R 80 °C de  $10^4$  Hz à  $10^6$  Hz. c) Superposition des courbes R 140 °C de 1 Hz à  $10^3$  Hz. d) Superposition des courbes R 140 °C à  $10^4$  Hz et  $10^5$  Hz.

Dans le confinement à 40 nm de diamètre et 2  $\mu\text{m}$  de long, les processus de relaxation associés à  $LTB_5$  sont détectés dès le traitement thermique à 60 °C, alors que ceci sont détectés après le traitement thermiques à 80 °C dans le confinement à 40 nm de diamètre et 10  $\mu\text{m}$  de long. Ainsi comme pour  $CtxB_5$ , les conformations protéiques enrichies et/ou sélectionnées possèdent des stabilités thermiques différentes si la longueur du confinement varie. Si l'expérience est reproduite, la position du premier processus de relaxation en température est reproduite, alors que celle du deuxième processus de relaxation est mieux définie dans la deuxième expérience (Fig. 7-15), suggérant que, comme pour  $CtxB$  les populations conformationnelles issues du dépliement de la toxine sont moins stables (effet de la cinétique).

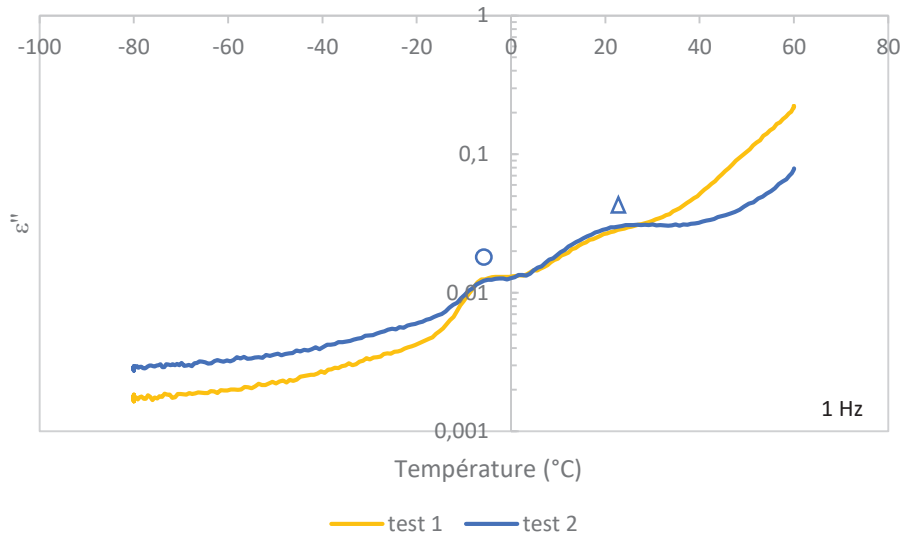


Figure 7-15. Reproduction de l'expérience  $LTB_5$  dans le confinement à 40 nm de diamètre et 2  $\mu$ m de long à la fréquence de 1 Hz après le traitement thermique à 60 °C.

Après avoir déterminé que la dynamique moléculaire des variants de séquence peut être distinguée en combinant le nanoconfinement et la spectroscopie diélectrique à large bande. Une protéine avec une conformation protéique monomérique est étudiée, pour déterminer si des changements conformationnels global pouvant être associés à des mutations de type pathologique peuvent-être détectés.

## 7.2 Etude du lysozyme

Le lysozyme est une protéine monomérique possédant 129 acides aminés (80,81,83) avec une fonction bactéricide (79). La stabilité thermique du lysozyme est étudiée par calorimétrie différentielle à balayage (DSC) et présente un pic endothermique à 74,3 °C (83). Par conséquent, les mêmes traitements thermiques que précédemment sont utilisés pour étudier le lysozyme.

### 7.2.1 Contrôle de l'échantillon protéique

Dans le confinement à 40 nm de diamètre et 5  $\mu$ m de long, l'échantillon protéique de lysozyme possède un signal diélectrique linéaire pour un remplissage volumique des nanopores à 14,3 % après différents traitements thermiques (Fig. 7-16). La méthodologie pour déterminer la quantité de lysozyme dans les nanopores est légèrement différente de la méthodologie utilisée pour les toxines : CtxB<sub>5</sub> et LTB<sub>5</sub>. CtxB<sub>5</sub> et LTB<sub>5</sub> sont des toxines pentamériques possédant plusieurs interfaces ce qui n'est pas le cas pour le lysozyme, par conséquent pour augmenter sa connectivité le volume du lysozyme est multiplié par 5. Malgré, une quantité de monomère de lysozyme plus importante, aucun processus de relaxation n'est détecté après les traitements thermiques à 60 °C et 80 °C (Fig. 7-16).

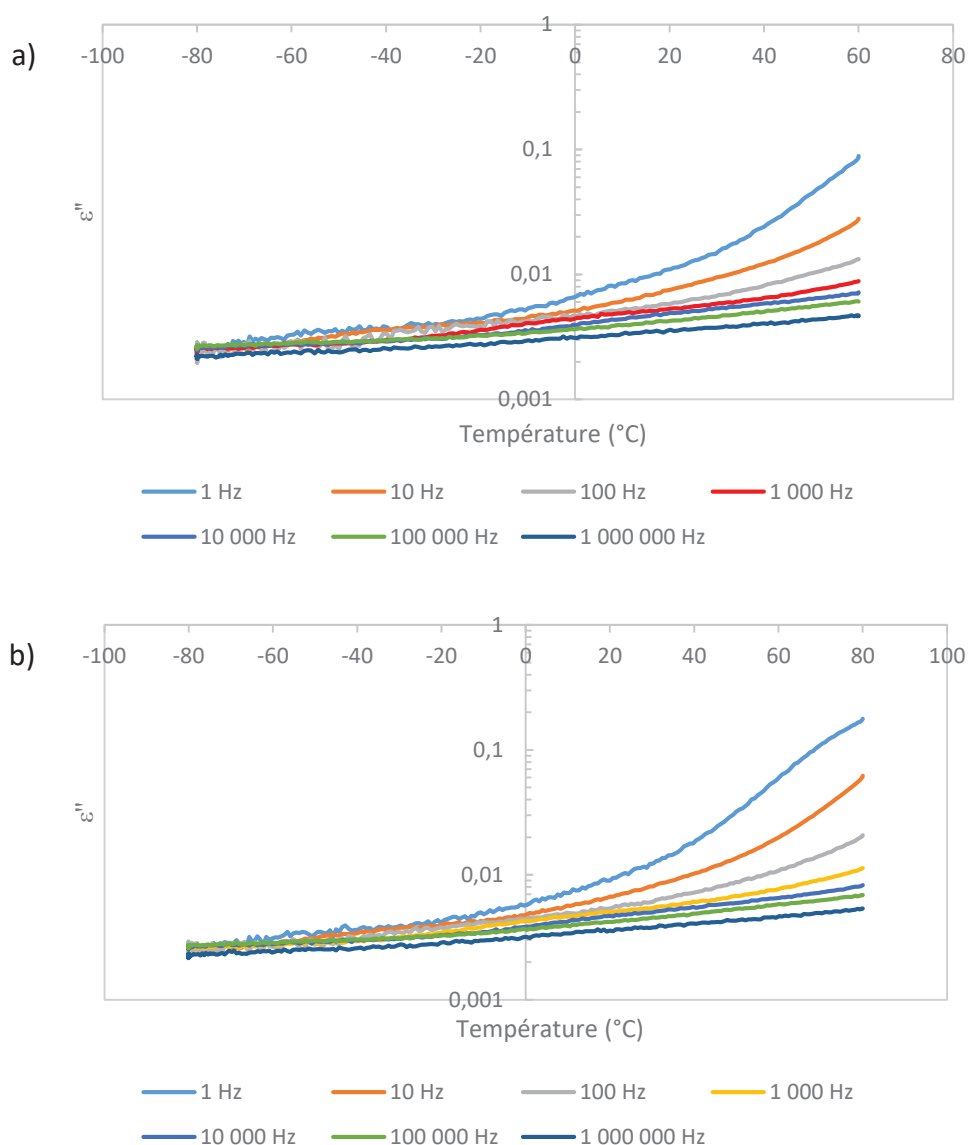


Figure 7-16. Lysozyme dans le confinement à 40 nm de diamètre et 5  $\mu\text{m}$  de long à différentes fréquences. a) Superposition des courbes de refroidissement après le deuxième traitement thermique à 60 °C (R2 60 °C). b) Superposition des courbes de refroidissement après le traitement thermique à 80 °C (R 80 °C). La solution protéique est composée de 10  $\mu\text{L}$  de lysozyme à 2mg/ml et de 316  $\mu\text{L}$  d'eau déionisée. La concentration finale est de 0,061 mg/ml.

D'après l'étude de Lai et al en 2000, le lysozyme possède un pic endothermique à 74,3 °C (83), par conséquent le lysozyme devrait commencer à se déplier et donc être perturbé thermiquement à 80 °C, entraînant la détection de processus de relaxation. Le chapitre précédent (Chapitre 6) a suggéré la présence d'un seuil minimum de détection. Le pourcentage de remplissage volumique des nanopores a donc été augmenté à 66 %, pour déterminer si le remplissage volumique à 14,3 % se situe en dessous du seuil de détection. Si le pourcentage de remplissage volumique des nanopores augmente, un processus de relaxation (symbole rond) est détecté (Fig. 7-17). La position de ce pic de relaxation en température se situe à 15°C. Pour plus de précision, le même échantillon de lysozyme aliquoté est utilisé dans les deux expériences de la figure 7-17, soit pour le remplissage volumique à 14,3 % mais aussi à 66%. Toutefois, dans les mêmes conditions, le contrôle du solvant possède

aussi un processus de relaxation à 15 °C. La présence d'un pic de relaxation à la même position si le pourcentage de remplissage est de 66 %, à la fois pour l'échantillon de lysozyme et pour le contrôle du solvant, interroge sur l'origine de ce pic. Le lysozyme est-il vraiment détecté ?

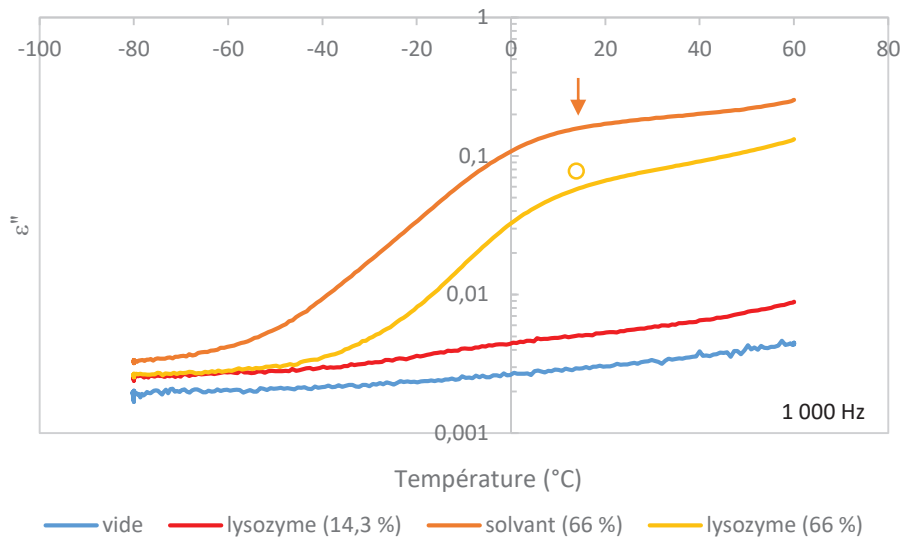


Figure 7-17. Lysozyme dans le confinement à 40 nm de diamètre et 5  $\mu\text{m}$  de long à la fréquence de  $10^3$  Hz après le traitement thermique à 60 °C. La courbe bleue correspond à la membrane vide. La courbe rouge au remplissage volumique des nanopores à 14,3 % de lysozyme et la courbe jaune au remplissage volumique des nanopores à 66 % de lysozyme. Pour la courbe jaune, l'échantillon protéique est composé de 50  $\mu\text{L}$  de lysozyme à 2 mg/ml et de 340  $\mu\text{L}$  d'eau déionisée. La concentration protéique de l'échantillon est de 0,26 mg/ml. La courbe orange correspond au contrôle du solvant avec 50  $\mu\text{L}$  de PBS et 340  $\mu\text{L}$  d'eau déionisée.

L'échantillon protéique et le solvant sont regardés à différentes fréquences après le traitement thermique à 60 °C (Fig. 7-18). A la fréquence de 1 Hz, le premier processus de relaxation (symbole rond) se situe à la même température (-27 °C) dans les deux expériences présentées dans la figure 7-18a. Ce processus de relaxation n'est pas détecté à la fréquence de  $10^5$  Hz (Fig. 7-18b). A la fréquence de 1 Hz, un deuxième processus de relaxation (symbole triangle) est aussi présent pour l'échantillon protéique et celui du solvant. Toutefois, la position de ce deuxième processus de relaxation en température est différente entre les deux expériences : 43 °C pour l'échantillon de lysozyme et 51 °C pour l'échantillon avec le solvant (Fig. 7-18a).

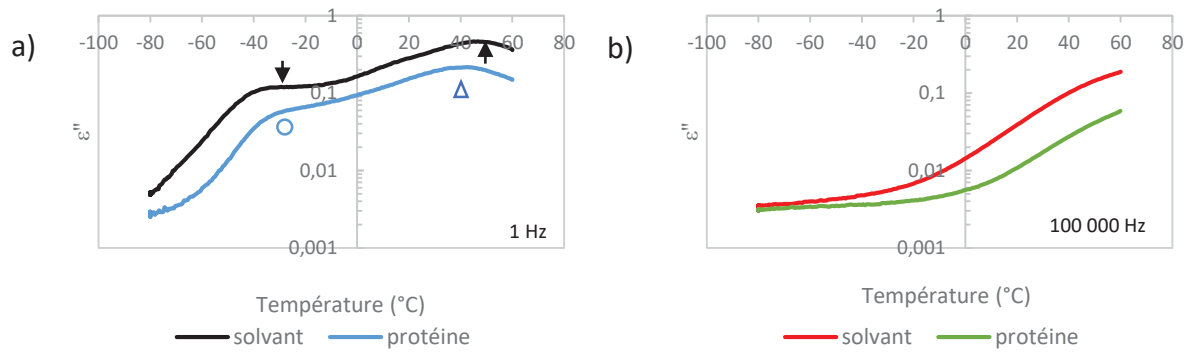
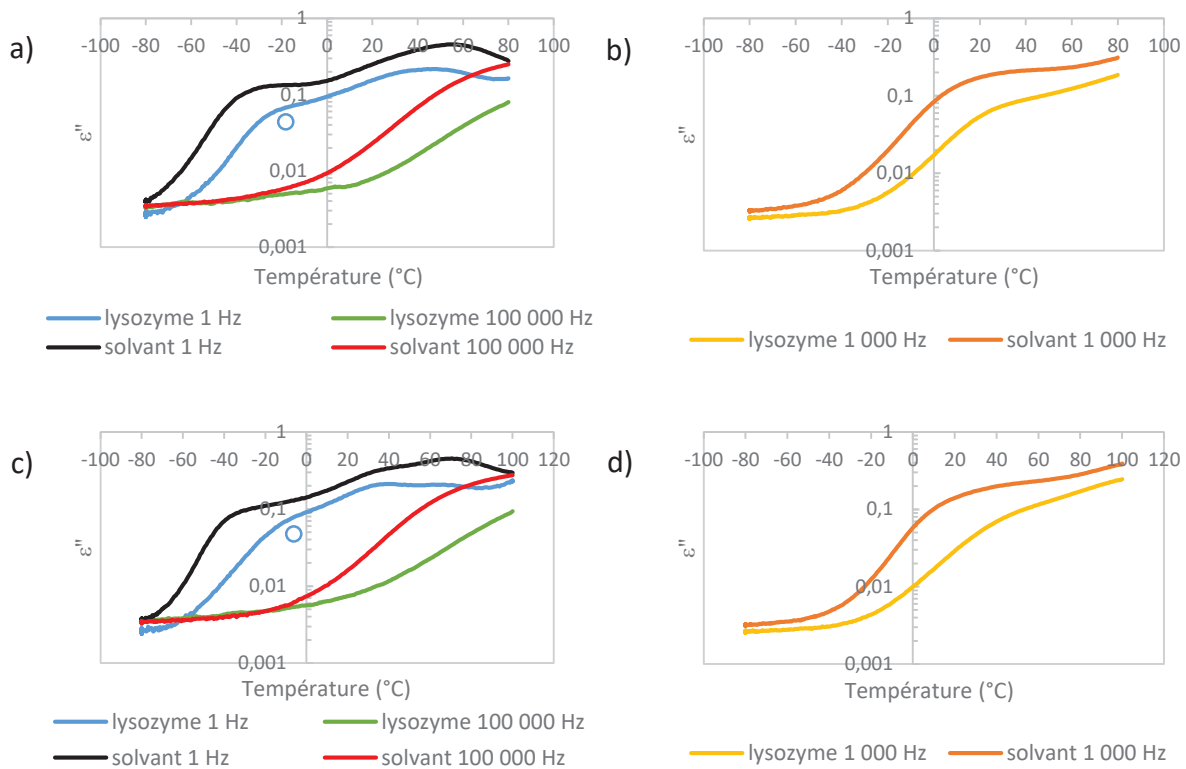


Figure 7-18. Comparaison des signaux diélectriques du solvant et du lysozyme après le deuxième traitement thermique à 60 °C dans le confinement à 40 nm de diamètre et 5 µm de long. a) Comparaison des courbes à la fréquence 1 Hz. b) Comparaison à 10<sup>5</sup> Hz.

L'analyse après le traitement thermique à 60 °C à différentes fréquences de l'échantillon protéique et du solvant ne permet pas de déterminer si les processus de relaxation détectés dans l'échantillon protéique sont bien associés à la protéine, plus précisément à sa conformation. Pour essayer de répondre à cette question, l'échantillon de lysozyme et le solvant sont soumis à différents traitements thermiques (Fig. 7-19).



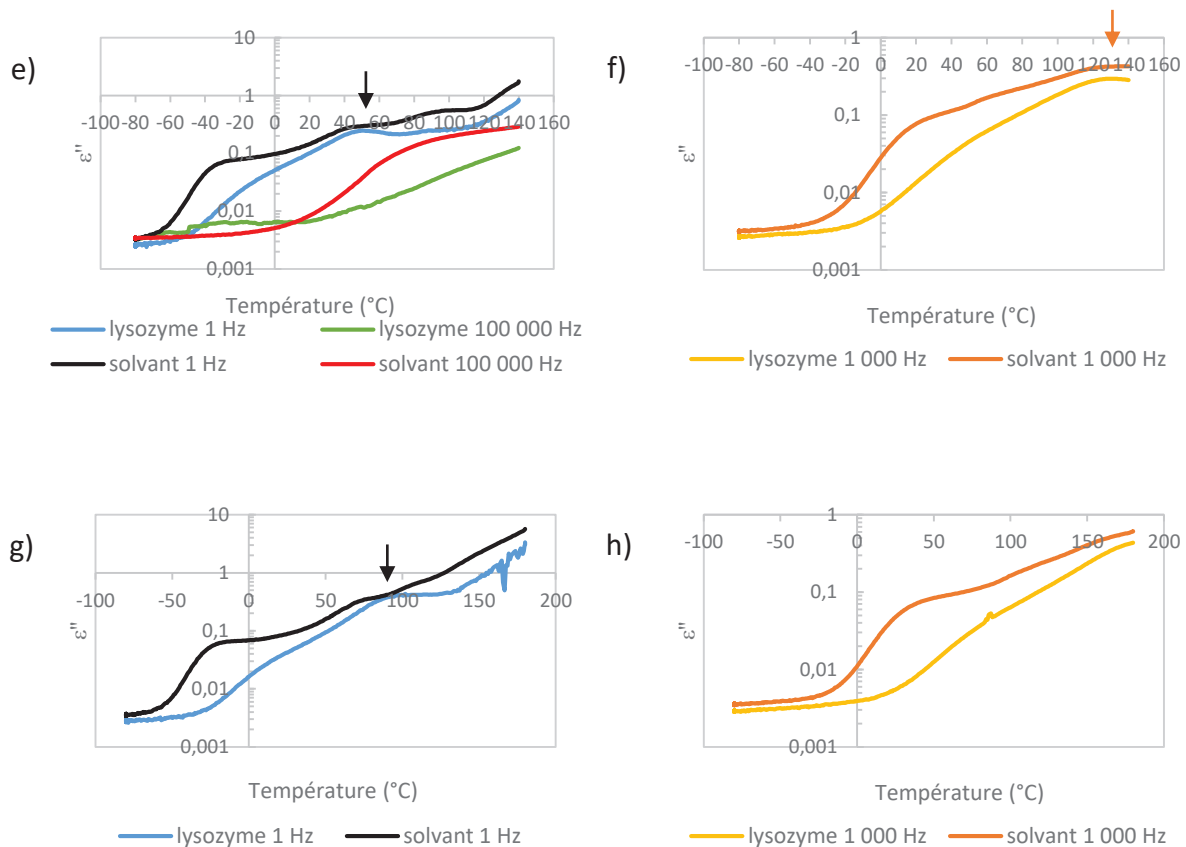


Figure 7-19. Comparaison des processus de relaxation du solvant et du lysozyme à différentes fréquences pour le confinement à 40 nm de diamètre et 5  $\mu$ m de long. a) Comparaison des courbes  $R$  80 °C à 1 Hz et 10<sup>5</sup> Hz. b) Comparaison des courbes  $R$  80 °C à 10<sup>3</sup> Hz. c) Comparaison des courbes  $R$  100 °C à 1 Hz et 10<sup>5</sup> Hz. d) Comparaison des courbes  $R$  100 °C à 10<sup>3</sup> Hz. e) Comparaison des courbes  $R$  140 °C à 1 Hz et 10<sup>5</sup> Hz. f) Comparaison des courbes  $R$  140 °C à 10<sup>3</sup> Hz. g) Comparaison des courbes  $R$  180 °C à 1 Hz. h) Comparaison des courbes  $R$  180 °C à 10<sup>3</sup> Hz.

A la fréquence de 1 Hz, l'échantillon de lysozyme possède deux processus de relaxation jusqu'au traitement thermique à 100 °C puis un seul après le traitement thermique à 140 °C. La température du premier processus de relaxation (symbole rond) augmente avec l'augmentation des traitements thermiques jusqu'à ne plus être détectée au traitement thermique de 140 °C. Après le traitement thermique à 80 °C, le pic de relaxation se situe à -21 °C puis à -10 °C après le traitement thermique à 100 °C. La position du deuxième processus de relaxation en température ne change pas jusqu'au traitement thermique à 100 °C (43 °C). Après le traitement thermique à 140 °C, la position de ce pic de relaxation en température est décalée à 53 °C puis à 103 °C après le traitement thermique à 180 °C.

A cette fréquence, le solvant possède aussi deux processus de relaxation jusqu'au traitement thermique à 100 °C mais trois processus de relaxation après le traitement thermique à 140 °C et deux processus de relaxation après le traitement thermique à 180 °C. La position du premier processus de relaxation en température change entre les traitements thermiques à 60 °C et 80 °C (-27 °C après le traitement à 60 °C et -25 °C après celui à 80 °C). Toutefois contrairement au lysozyme, ce premier processus de relaxation est détecté jusqu'au traitement thermique à 180 °C. De plus, la position de ce processus de relaxation en température ne change pas entre les traitements thermiques à 80 °C et 140 °C (-25 °C). Le

deuxième processus de relaxation du solvant est détecté jusqu'à 140 °C. En plus de disparaître plus tôt vis-à-vis de l'échantillon protéique (140 °C versus 180 °C), contrairement à l'échantillon protéique, la position de ce deuxième processus de relaxation du solvant en température change après tous les traitements thermiques. Enfin, contrairement au lysozyme, après les traitements thermiques à 140 °C et 180 °C, un troisième processus de relaxation associé au solvant est détecté (59 °C après le traitement thermique à 140 °C et 79 °C après celui à 180 °C).

A la fréquence  $10^3$  Hz, l'échantillon de lysozyme détecte uniquement le premier processus de relaxation jusqu'au traitement thermique à 100 °C. La température de ce processus de relaxation augmente avec l'augmentation de la température des traitements thermiques. Après le traitement thermique à 140 °C, le premier processus de relaxation n'est plus détecté, toutefois, un deuxième processus de relaxation est détecté à 131 °C. Pour le contrôle du solvant, la température du premier processus de relaxation détecté jusqu'au traitement thermique à 100 °C augmente avec l'augmentation des traitements thermiques. Néanmoins, dans le cas du solvant ce premier processus de relaxation est aussi détecté après les traitements thermiques à 140 °C et 180 °C. Comme pour la protéine après le traitement thermique à 140 °C, un deuxième pic de relaxation est détecté pour l'échantillon du solvant.

A la fréquence de  $10^5$  Hz, aucun processus de relaxation n'est détecté pour le lysozyme, dans la gamme de température des traitements thermiques. Toutefois, dans le cas du solvant le premier processus de relaxation commence à être détecté après le premier traitement thermique à 80 °C. Ce processus de relaxation possède un pic de relaxation à 88 °C après le traitement thermique à 100 °C et à 96 ° après le traitement thermique à 140 °C.

La table 7-3, récapitule la position du premier processus de relaxation en température à la fois pour l'échantillon de lysozyme mais aussi pour celui du solvant aux trois fréquences étudiées. De même la table 7-4, regroupe la position du deuxième processus de relaxation en température détecté à la fois pour la protéine et le solvant.

Traitement thermique	1 Hz		1 000 Hz		100 000 Hz	
	Lysozyme	Solvant	Lysozyme	Solvant	Lysozyme	Solvant
60 °C	-27 °C	-27 °C	15 °C	15 °C		
80 °C	-21 °C	-25 °C	32 °C	26 °C		Commence à être détecté
100 °C	-10 °C	-25 °C	47 °C	31 °C		88 °C
140 °C		-25 °C		31 °C		96 °C
180 °C		-14 °C		59 °C	*	*

Table 7-3. Table récapitulative de la position du premier processus de relaxation en température détecté pour l'échantillon de lysozyme et pour l'échantillon du solvant. L'analyse à la fréquence de  $10^5$  Hz après le traitement thermique à 180 °C n'est pas réalisée, car la courbe de l'expérience avec le lysozyme n'est pas interprétable.



	1 Hz		1 000 Hz	
Traitement thermique	Lysozyme	Solvant	Lysozyme	Solvant
60 °C	43 °C	51 °C		
80 °C	43 °C	56 °C		
100 °C	43 °C	73 °C		
140 °C	53 °C	103 °C	131 °C	131 °C
180 °C	103 °C			

Tableau 7-4. Table récapitulative de la position du deuxième processus de relaxation en température détecté pour l'échantillon du lysozyme et pour l'échantillon du solvant.

Les figures 7-19 et la table 7-3 montrent que le premier processus de relaxation détecté pour l'échantillon de protéine et pour le solvant diffère à la fois vis-à-vis de sa résistance aux différents traitements thermiques mais aussi vis-à-vis des fréquences de détection. Par conséquent, il est raisonnable de supposer que le premier processus de relaxation détecté est bien caractéristique des conformations protéiques. Toutefois, la réponse diélectrique du lysozyme doit résulter d'un couplage protéine-PBS ou protéine-PBS-eau liée.

Pour le deuxième processus de relaxation détecté dans l'échantillon de lysozyme et du solvant, il est plus difficile d'affirmer que ce processus de relaxation est bien associé à la conformation protéique du lysozyme. Toutefois, la position de ce processus de relaxation en température et sa détection jusqu'au traitement thermique à 180 °C dans le cas de l'échantillon protéique suggèrent que ce deuxième processus de relaxation devrait aussi être associé à la conformation protéique (Fig. 7-19 et table 7-4).

Dans le cas, où les deux processus de relaxation détectés pour l'échantillon de lysozyme correspondent à la détection de la conformation protéique vis-à-vis des dipôles protéines-PBS ou protéine-PBS-eau, une interprétation sur la présence de deux pics de relaxation est réalisée. Le premier processus de relaxation du lysozyme est détecté à la fois dans les basses et hautes fréquences (1 Hz à  $10^5$  Hz) alors que le deuxième processus de relaxation du lysozyme est détecté dans les basses fréquences (jusqu'à  $10^3$  Hz). De plus, la position du deuxième processus de relaxation en température est plus importante que celle du premier processus de relaxation pour tous les traitements thermiques. Nous présumons donc que le deuxième processus de relaxation détecté possède un environnement moins compact que le premier permettant ainsi des mouvements à plus grande échelle. Par conséquent, il est raisonnable de supposer que le premier processus de relaxation détecté correspond à la conformation protéique la plus stable et la plus repliée, soit un monomère de type natif et que le deuxième processus de relaxation correspond à un état plus déplié que le premier. Au vu de la conformation monomérique du lysozyme, cet état correspond soit à une population de conformation intermédiaire, soit à une population de monomère déplié. Dans les chapitres précédents (Chapitre 4 à 6), nous avons émis l'hypothèse que les monomères dépliés ne sont pas détectés dans cette gamme de fréquence, par conséquent il est raisonnable de supposer que le deuxième processus de relaxation détecté correspondrait à une population de conformation intermédiaire du lysozyme, ce qui est en accord avec des études sur le lysozyme (80,82,161).

La présence de pic de relaxation dans l'échantillon du solvant peut s'expliquer par la nature de celui-ci. En effet, le solvant est composé de PBS, soit de différents sels et d'eau : NaCl, KCl,  $\text{KH}_2\text{PO}_4$  et  $\text{Na}_2\text{HPO}_4$ ,  $\text{H}_2\text{O}$ . Par conséquent, différents dipôles peuvent être présents dans le solvant et donner un ou plusieurs processus de relaxation. Le changement de position des processus de relaxation en température et l'apparition d'un troisième processus de relaxation pourrait s'expliquer par la réorganisation des dipôles lors des différents traitements thermiques.

Si la même expérience est reproduite avec l'échantillon de lysozyme, la position du premier processus de relaxation en température ne varie pas (Fig. 7-20a). Dans ce confinement et avec un remplissage volumique des nanopores à 66%, les données de pertes diélectriques ne détectent pas de différence significative entre les processus de refroidissement et de chauffage (Fig. 7-20b). De même, si deux traitements thermiques identiques sont réalisés dans une même expérience, la position du pic de relaxation en température ne change pas (Fig. 7-20c).

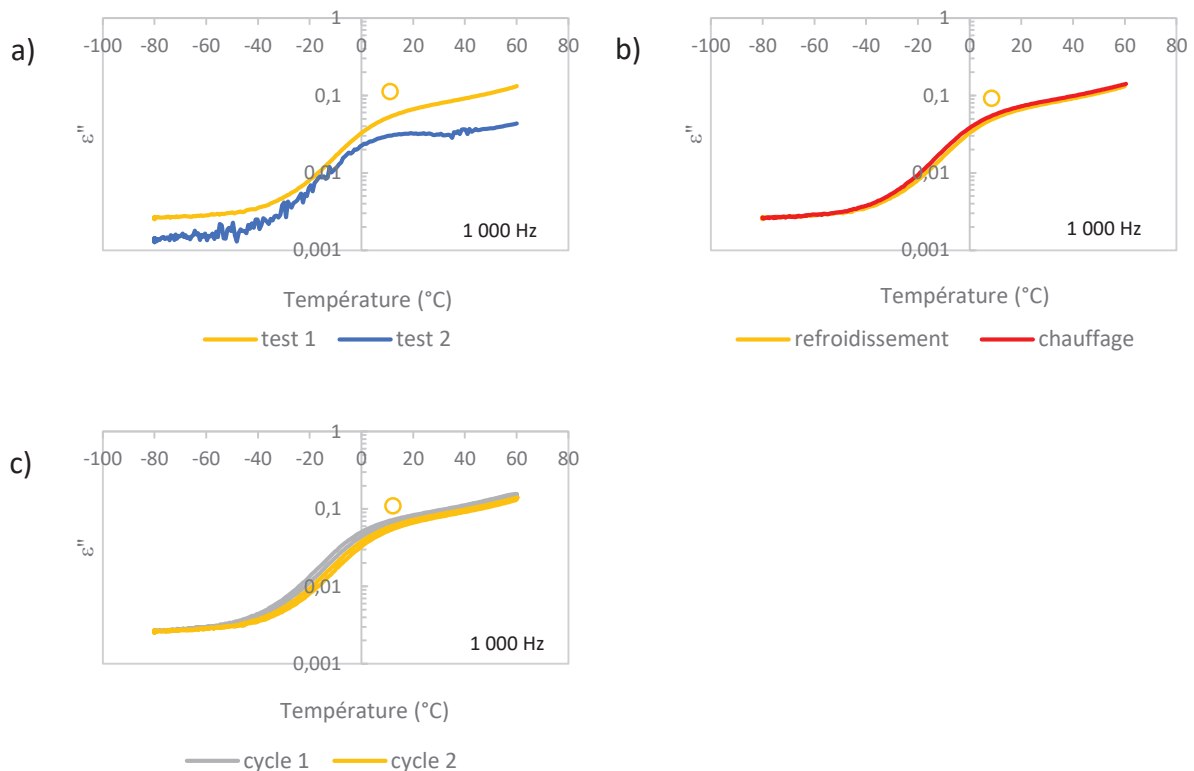


Figure 7-20. Analyse du lysozyme dans le confinement à 40 nm de diamètre et 5  $\mu\text{m}$  de long à la fréquence de  $10^3$  Hz. a) Comparaison des courbes  $R_2$  60 °C pour deux expériences différentes. b) Analyse du deuxième cycle à 60 °C pour le lysozyme. c) Superposition de deux cycles à 60 °C dans une même expérience.

En conclusion, ce chapitre a permis de montrer que la combinaison du nanoconfinement et de la spectroscopie diélectrique à large bande permettait de détecter l'impact des variants de séquence, donc des changements locaux avec l'étude de  $\text{LTB}_5$  mais aussi des changements structuraux plus importants lors de l'étude du lysozyme. L'étude de

différentes protéines a mis en évidence que la combinaison du nanoconfinement et de la spectroscopie diélectrique à large bande permet de détecter l'impact des mutations sur la conformation protéique via la dynamique moléculaire.

# Chapitre 8

## Discussion, Conclusions, Perspectives

Les protéines sont caractérisées par un ensemble de conformations protéiques fournissant une même fonction. Ces différentes conformations protéiques sont définies par des interactions moléculaires différentes et donc des dynamiques distinctes avec des mouvements de structures secondaires, tertiaires et quaternaires et de dépliement et/ou repliement différents. Les différentes conformations d'une protéine possèdent donc des dynamiques moléculaires distinctes. Ces alternatives conformationnelles et dynamiques, donnent au système une plasticité structurale qui sous-tend sa durabilité fonctionnelle car elles permettent de combiner robustesse et adaptabilité et de diminuer la fragilité du système comparé à une solution structurale unique. Au jour d'aujourd'hui, différentes techniques expérimentales et théoriques existent pour caractériser une conformation protéique et sa dynamique. Toutefois, aucune technique ne permet de cribler l'ensemble conformationnel d'une protéine et la dynamique individuelle de chaque conformation. L'objectif de cette thèse est de déterminer l'applicabilité d'une approche combinant le nanoconfinement et la spectroscopie diélectrique à large bande pour cribler les différentes dynamiques d'une protéine et en inférer les différentes conformations protéiques qui constituent l'ensemble conformationnel d'une protéine. Le nanoconfinement devrait permettre de réduire l'hétérogénéité conformationnelle d'une population de protéine et la spectroscopie diélectrique à large bande devrait mesurer la dynamique moléculaire des protéines nanoconfinées.

Cette approche combinant le nanoconfinement et la spectroscopie diélectrique à large bande n'a jamais été utilisée sur des protéines par conséquent, cette thèse exploratoire vise à 'débroussailler le terrain' et répondre à plusieurs questions pour établir son utilité pour l'étude des protéines.

La première question est : cette approche est-elle applicable aux protéines ? Les protéines possèdent-elles un signal diélectrique caractéristique de leurs conformations ? Les travaux réalisés dans ce mémoire démontrent que la combinaison du nanoconfinement et de la spectroscopie diélectrique à large bande permet de caractériser une protéine en termes de conformation, stabilité et dynamique moléculaire. L'analyse des données de spectroscopie diélectrique à différentes températures et à différentes fréquences permet de suivre les différentes dynamiques associées aux différentes conformations protéiques adoptées au cours du dépliement thermique à travers la mesure des fluctuations dipolaires de l'échelle de la microseconde à la seconde. Ces échelles de fluctuations correspondent à des mouvements

collectifs allant des mouvements de structures secondaires à des mouvements de dépliement et repliement ou encore des mouvements de chaînes lors de processus d'assemblage et de dissociation protéique.

Pour comprendre les signaux diélectriques obtenus via cette approche, plusieurs paramètres sont utilisés : la position du pic en température et en fréquence, la largeur du pic et son intensité. Ces signaux diélectriques sont ensuite analysés de manière non conventionnelle par rapport aux études 'traditionnelles' de spectroscopie diélectrique à large bande. En effet, les signaux ne sont pas analysés en termes de relaxations  $\beta$  et  $\alpha$  qui rapportent la dynamique d'un polymère. Avec pour la relaxation  $\beta$  des fluctuations locales, c'est à dire des mouvements de petites amplitudes **détectés à hautes fréquences et à basse température et pour la relaxation  $\alpha$  des mouvements collectifs à plus large échelle, comme le mouvement de domaines, détectés à plus basse température et à plus basses fréquences.** Nos mesures détectent des dynamiques à haute température à des fréquences du MHz au Hz, qui correspondent aux échelles de temps des mouvements connus pour la dynamique des protéines tels que les mouvements de chaîne ou de large domaine au-delà de la milliseconde et des mouvements de structures secondaire de la microseconde à la milliseconde. Nous avons donc choisi d'interpréter nos résultats en terme de dynamique des structures protéiques. De plus, nous proposons une approche inspirée de la spectroscopie de fluorescence où les propriétés spectrales sont reliées à l'environnement des sondes fluorescentes et en particulier à leur confinement local. Dans une protéine de structure native, le fluorophore est confiné (entouré de voisins), alors que dans une protéine dépliée il sera exposé au solvant. Ces deux états de confinement modifient la réponse spectrale car ils affectent les mouvements et l'orientation des sondes. De même, la relaxation moléculaire mesurée en BDS est due à la fluctuation des dipôles, fluctuation qui dépend comme pour le fluorophore de leur environnement. Les résultats de BDS montrent bien que les caractéristiques spectrales des signaux diélectriques (position en fréquence et en température, largeur des pics, intensité) se modifient lors du dépliement thermique des protéines, soutenant l'analogie avec la spectroscopie de fluorescence. Dans nos expériences en BDS, plus un pic est détecté à haute température à une fréquence donnée, plus la dynamique mesurée est due à des mouvements de large amplitude correspondant à un état déplié. Les mouvements de petites amplitudes comme ceux des chaînes latérales ne sont pas détectés aux fréquences étudiées. De même, les pics observés reflètent la dynamique des domaines dépliés dans la protéine, les domaines toujours repliés ne peuvent pas être détectés aux fréquences étudiées du fait de leurs mouvements trop rapides et de faibles amplitudes. Les différents pics détectent les dynamiques de différents sets de dipôles reflétant ainsi des zones de dépliement distinctes dans la toxine. Nous avons en plus eu l'avantage que les protéines étudiées ont des structures atomiques connues nous permettant d'inférer l'ensemble des dipôles électrostatiques présent dans la structure native, ensemble constitué de toutes les charges positives et négatives à proximité les unes des autres. Les dipôles électrostatiques sont les principaux contributeurs du signal diélectrique dans un milieu ionisé et sont aussi déterminants pour la structure d'une protéine. En combinant les échelles de la dynamique des protéines avec les réseaux de dipôles électrostatiques inférés des structures

atomiques, nous avons proposé une approche permettant d'assigner les zones de la protéine aux différents pics de relaxation.

Les expériences effectuées avec le lysozyme conforte ce type d'analyse puisqu'elles montrent qu'à nombre similaire de dipôles, une protéine ayant un réseau de dipôles électrostatiques 'connecté' aura un signal diélectrique dans les fréquences mesurées ici, plus facilement détectable (exemple : CtxB<sub>5</sub>) qu'une protéine ayant des dipôles spatialement distribués (exemple : Lysozyme). Du fait probablement de leurs motions collectives qui contribuent de façon coopérative au signal diélectrique. Une préparation d'échantillon adéquate peut s'avérer nécessaire pour détecter un signal diélectrique en compensant la plus faible réponse de dipôles électrostatiques distribués spatialement dans la structure.

Il reste encore beaucoup à explorer pour comprendre comment les dipôles contribuent au signal diélectrique d'une protéine, nos analyses non conventionnelles offrent une perspective à poursuivre pour la déconvolution des signaux diélectriques en termes structuraux. Ces analyses visant à établir la relation entre caractéristiques spectrales et structures/dynamique des protéines sont soutenues par le fait que les signaux diélectriques sont détectés dans les mêmes gammes de fréquences que celles rapportées par des approches de dynamique moléculaire par simulation ou par des approches expérimentales pour la dynamique des protéines.

Ayant montré que le BDS en condition nanoconfinée permettait de suivre des conformations protéiques différentes via leurs dynamiques respectives, nous avons alors envisagé que le BDS puisse permettre l'étude de l'impact de modifications de séquence sur une conformation protéique ou sa dynamique.

La deuxième question est donc : deux protéines de séquences et structures proches ont-elles néanmoins, des signaux diélectriques distincts du fait de l'impact des mutations sur la dynamique des protéines ? Cet aspect est très intéressant, car en plus de posséder un ensemble conformationnel pour une même fonction, les protéines peuvent subir des mutations impactant leurs fonctions, sans qu'il soit facile de remonter à la cause de la perturbation fonctionnelle. Pourtant diagnostiquer la cause est essentielle pour proposer des moyens de corriger la perturbation et rétablir une fonction. Pour répondre à cette question, deux toxines (CtxB<sub>5</sub> et LTB<sub>5</sub>) possédant un pourcentage d'identité de séquence élevée (83,5 %), des conformations et fonctions similaires sont étudiées. Ces toxines possèdent des signaux diélectriques différents, démontrant que cette approche est capable de détecter des différences de dynamique moléculaire entre des variants de séquence. De plus, ces deux toxines possèdent des chemins de dépliage différents détectés en utilisant cette approche. En effet, suivant la température des traitements thermiques, le nombre de processus de relaxation détectés et la largeur de ces pics varient entre les deux toxines. La présence de trois dipôles ioniques intramoléculaires supplémentaires dans la toxine LTB<sub>5</sub> (Ala1-Glu7, Lys81-Glu102 et Asp83-Glu102) modifie la stabilité locale des extrémités N et C terminal et protège la toxine de la dissociation thermique en renforçant ces interfaces. Le rôle du dipôle supplémentaire localisé dans l'extrémité N-terminale pourrait être confirmé en cross-linkant le N-terminal de la cholera toxine et en répétant l'expérience en diélectrique. La cholera toxine cross-linkée devrait avoir une extrémité N-terminale plus stable et un comportement diélectrique similaire à celui de LTB<sub>5</sub>.

Une fois que différentes conformations protéiques (pentamères dépliés, intermédiaires d'assemblage, séquence variant) ont pu être détectées par cette approche, la dernière question envisagée est : peut-on détecter différentes conformations protéiques en modifiant les dimensions du confinement ?

Pour répondre à cette question la protéine (CtxB<sub>5</sub>) est testée dans différents confinements. Les travaux de ce mémoire, démontrent que les variations des dimensions de nanoconfinement à la fois en termes de diamètre et de longueur permettent de cribler différentes conformations protéiques caractérisées par des positions en température et en fréquence des pics de relaxation différentes ainsi que par des stabilités thermiques différentes.

Les dimensions du confinement visent à privilégier l'entrée dans le nanopore de certaines conformations parmi l'ensemble conformationnel, opérant ainsi une sélection conformationnelle. Maintenant, il est aussi possible que les dimensions du confinement refaçonnent les différentes conformations qui sont dans la goutte en une seule conformation confinée, opérant alors une homogénéisation conformationnelle. Comparez à une mesure macroscopique, dans le cas premier cas (Fig. 2-10a), une seule conformation homogène sera enrichie et le même nombre de protéines qu'en macroscopique sera détecté. Les conformations enrichies ne reflèteront donc peut-être pas les conformations réelles. Dans le deuxième cas (Fig. 2-10b), moins de protéines seront détectées du fait de la sélection de certaines conformations et l'exclusion d'autres. Dans ces deux cas de figure, on s'attend alors à avoir un spectre moins large mais avec une diminution d'intensité du signal par rapport à la situation macroscopique pour la sélection conformationnelle contre une plus haute intensité et une même surface sous la courbe pour l'enrichissement conformationnel. On distinguera donc la sélection du cas de l'enrichissement, même si la sélection conformationnelle mène aussi à un enrichissement de certaines conformations au détriment des autres.

Pour déterminer si les dimensions du confinement permettent un enrichissement et/ou une sélection des populations conformationnelles, idéalement la protéine CtxB<sub>5</sub> devrait être testée dans des conditions faiblement hydratée et macroscopique. La comparaison de l'intensité et de la largeur du pic de relaxation pour la protéine dans ces deux conditions devrait permettre de déterminer si le confinement provoque un enrichissement ou une sélection conformationnelle. Malheureusement, les échantillons de CtxB<sub>5</sub> dans des conditions faiblement hydratées en condition macroscopique (pas déposés dans des nanopores), entraîne un court-circuit et ne peuvent donc pas être étudiés en BDS à ce stade. Cela montre que la déshydratation en condition macroscopique n'a pas le même impact sur la structure protéique que la déshydratation en condition nanoscopique, phénomène qui sera intéressant d'étudier dans le futur. Pour pouvoir comparer les réponses diélectriques en conditions déshydratées en nanoconfinement et en macroscopique il faudrait mettre en place une nouvelle cellule de détection permettant de régler les conditions d'hydratation.

Une autre solution pour déterminer l'impact du nanoconfinement est d'augmenter le diamètre des nanopores. Dans le cas de l'enrichissement, les intensités des pics de relaxation devraient être plus élevées et les pics plus fins pour les confinements de plus petits diamètres. Comme la sélection exclue des conformations protéiques, en augmentant le diamètre de confinement le nombre de conformation protéique exclu devrait diminuer et donc l'intensité du pic de relaxation devrait augmenter. Les résultats montrent qu'en augmentant le diamètre



du confinement, l'intensité du pic de relaxation augmente ainsi le confinement provoquerait une sélection conformationnelle. Toutefois, les pics de relaxation détectés dans les différents confinements ne se chevauchent pas avec les pics détectés dans le confinement de diamètre large (120 nm), ce qui suggère que les conformations détectées avec les pores de petit diamètre ne sont pas détectées avec les pores de diamètre plus large. Or si différentes populations conformationnelles sont sélectionnées par le diamètre des nanopores, il est raisonnable de supposer qu'en augmentant le diamètre ces différentes conformations devraient être détectées et donc les pics de relaxation devraient se chevaucher en température. Comme ce n'est pas le cas, il est possible que le confinement provoque aussi un enrichissement conformationnel, en façonnant des conformations protéiques dans les pores de petit diamètre. Les résultats suggèrent qu'une combinaison de sélection et d'enrichissement se produit.

De plus, en modifiant les dimensions du confinement via le diamètre et la longueur, l'écoulement des protéines dans les nanopores diffère aussi. L'impact de l'écoulement dans les nanopores via le diamètre s'explique par les dimensions des nanopores et les dimensions des protéines étudiées. En effet, si les protéines possèdent des dimensions similaires à celles du confinement, alors l'entrée des protéines dans les nanopores devrait être limitée suite à un encombrement macromoléculaire. Parallèlement, si le diamètre de confinement est plus important que celui des protéines alors les protéines rentreront plus facilement. L'entrée des protéines dans les nanopores est déterminée par la taille des protéines mais aussi par leurs conformations (formes). Pour un même diamètre de confinement, plus la taille des protéines est petite plus le nombre de protéines pouvant entrer dans les nanopores augmentent. De même, suivant la conformation de la protéine plus ou moins de protéines pourront rentrer dans les nanopores. Suivant la taille et la forme des protéines et leurs contacts, les 'vides' présents entre les protéines seront différents entraînant des vitesses d'écoulement potentiellement différentes et laissant ainsi passer différentes conformations, suivant un processus similaire à ceux observés pour la dynamique des systèmes granulaires jusqu'aux possibilités de bouchons dit 'jamming'.

Le dernier aspect mis en évidence dans ce travail, correspond à la présence d'une dynamique moléculaire pour le solvant. Le solvant est composé d'eau et de différents sels. Dans les confinements de 40 nm de diamètre et 10  $\mu\text{m}$  de long, 40 nm de diamètre et 2  $\mu\text{m}$  de long et 120 nm de diamètre et 2  $\mu\text{m}$  de long aucun processus de relaxation du solvant n'est détecté après six heures à 60 °C. En revanche, un signal diélectrique du solvant est détecté pour le contrôle du lysozyme jusqu'au traitement thermique à 180 °C. Cela pourrait s'expliquer par la concentration de sel beaucoup plus élevée dans le contrôle du lysozyme du fait d'un pourcentage de remplissage de 66% au lieu de 14,3 %. Cette importante quantité de sel pourrait entraîner la nucléation des sels et la formation de large réseau cristallin présentant des dynamiques lentes et donc détectées expérimentalement. Toutefois, le fait de détecter aussi une dynamique moléculaire jusqu'à 180 °C dans le confinement de 40 nm de diamètre et 5  $\mu\text{m}$  de long alors que le pourcentage de remplissage est de 14,3 % suggère que les dimensions du confinement jouent aussi un rôle sur la formation de réseau cristallin. Il faudra poursuivre ces études pour éclaircir les conditions les plus adaptées pour éviter la



contribution des dipôles du solvant dans le signal diélectrique qui complique l'analyse des données protéiques.

L'avantage de la combinaison du nanoconfinement et de la spectroscopie diélectrique à large bande est qu'une très faible quantité de protéine est utilisée. De plus, cette approche permet de détecter l'impact des mutations sur la dynamique moléculaire des protéines à l'échelle locale et globale mais aussi de cribler différentes conformations protéiques et leurs dynamiques respectives de l'échelle locale (exemple : mouvement des structures secondaires à la fréquence de  $10^3$  Hz) à l'échelle globale (exemple : mouvement de chaîne à la fréquence de 1 Hz). Les dimensions et formes des protéines étudiées influenceront sur les choix des dimensions des confinements à utiliser pour cribler ces conformations. Le choix de la longueur et du diamètre de confinement mais aussi du pourcentage de remplissage volumique des nanopores est essentiel pour cribler différentes conformations protéiques et 'limiter les problèmes' d'écoulement dans les nanopores.

L'étude des protéines via le nanoconfinement et la spectroscopie diélectrique à large bande est un développement récent qui a mis en évidence de nombreuses perspectives. Une perspective non explorée reste la question de la fonctionnalité des conformations détectées en diélectrique. Les nanopores sont utilisés comme biocapteur et permettent de conserver la fonction d'une protéine, ce qui tend à suggérer que les conformations étudiées au cours de cette thèse sont fonctionnelles. Dans le cas des toxines, les nanopores pourraient être recouverts de bicouche lipidique contenant les récepteurs des toxines GM<sub>1</sub> puis les toxines ajoutées pour effectuer la mesure diélectrique. Les signaux pourraient être comparés avec les expériences sans GM<sub>1</sub> afin de déterminer si les toxines adoptent les mêmes conformations dans les deux cas suggérant le maintien de la fonction dans le nanopore.

La spectroscopie diélectrique à large bande mesure des fluctuations moléculaires impliquant principalement les acides aminés chargés, par conséquent le remplissage de nanopores sous champ électrique devrait aussi permettre de cribler différentes conformations protéiques. L'application d'un champ électrique favorisera l'orientation des dipôles, par conséquent l'analyse de différents champs électriques devrait favoriser différentes orientations de dipôles et donc la détection de différentes dynamiques reflétant différentes conformations protéiques.

Les travaux de ce mémoire ont démontré que différentes conformations globulaires peuvent être détectées, par conséquent il est nécessaire de savoir si d'autres protéines, comme les protéines fibreuses, les protéines membranaires, les protéines désordonnées ... peuvent aussi être étudiées via cette approche. Il est raisonnable de supposer que si les protéines étudiées possèdent un nombre suffisant de dipôles et d'interactions via les acides aminés chargés alors une dynamique moléculaire devrait être détectée.

# Références

1. Berg JM, Tymoczko JL, Gatto GJ, Stryer L. Biochemistry Eighth Edition. W. H. Freeman and Company;
2. Garrett RH, Grisham CM. Biochemistry Sixth Edition. Cengage Learning Custom Publishing;
3. Muñoz V, Cerminara M. When fast is better: protein folding fundamentals and mechanisms from ultrafast approaches. *Biochem J*. 2016 01;473(17):2545–59.
4. Best RB, Lindorff-Larsen K, DePristo MA, Vendruscolo M. Relation between native ensembles and experimental structures of proteins. *Proc Natl Acad Sci U S A*. 2006 Jul 18;103(29):10901–6.
5. Parisi G, Zea DJ, Monzon AM, Marino-Buslje C. Conformational diversity and the emergence of sequence signatures during evolution. *Curr Opin Struct Biol*. 2015 Jun;32:58–65.
6. Bonomi M, Heller GT, Camilloni C, Vendruscolo M. Principles of protein structural ensemble determination. *Curr Opin Struct Biol*. 2017;42:106–16.
7. Javier Zea D, Miguel Monzon A, Fornasari MS, Marino-Buslje C, Parisi G. Protein conformational diversity correlates with evolutionary rate. *Mol Biol Evol*. 2013 Jul;30(7):1500–3.
8. Monzon AM, Juritz E, Fornasari MS, Parisi G. CoDNaS: a database of conformational diversity in the native state of proteins. *Bioinforma Oxf Engl*. 2013 Oct 1;29(19):2512–4.
9. Juritz EI, Alberti SF, Parisi GD. PCDB: a database of protein conformational diversity. *Nucleic Acids Res*. 2011 Jan;39(Database issue):D475-479.
10. Rossjohn J, Raja SM, Nelson KL, Feil SC, van der Goot FG, Parker MW, et al. Movement of a loop in domain 3 of aerolysin is required for channel formation. *Biochemistry (Mosc)*. 1998 Jan 13;37(2):741–6.
11. Spyrakakis F, BidonChanal A, Barril X, Luque FJ. Protein flexibility and ligand recognition: challenges for molecular modeling. *Curr Top Med Chem*. 2011;11(2):192–210.
12. Ponzoni L, Bahar I. Structural dynamics is a determinant of the functional significance of missense variants. *Proc Natl Acad Sci U S A*. 2018 17;115(16):4164–9.
13. Henzler-Wildman KA, Lei M, Thai V, Kerns SJ, Karplus M, Kern D. A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature*. 2007 Dec 6;450(7171):913–6.

14. Henzler-Wildman KA, Thai V, Lei M, Ott M, Wolf-Watz M, Fenn T, et al. Intrinsic motions along an enzymatic reaction trajectory. *Nature*. 2007 Dec 6;450(7171):838–44.
15. Wolf-Watz M, Thai V, Henzler-Wildman K, Hadjipavlou G, Eisenmesser EZ, Kern D. Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair. *Nat Struct Mol Biol*. 2004 Oct;11(10):945–9.
16. Lange OF, Lakomek N-A, Farès C, Schröder GF, Walter KFA, Becker S, et al. Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution. *Science*. 2008 Jun 13;320(5882):1471–5.
17. Fuentes EJ, Der CJ, Lee AL. Ligand-dependent dynamics and intramolecular signaling in a PDZ domain. *J Mol Biol*. 2004 Jan 23;335(4):1105–15.
18. Smock RG, Gierasch LM. Sending signals dynamically. *Science*. 2009 Apr 10;324(5924):198–203.
19. Aharoni A, Gaidukov L, Khersonsky O, McQ Gould S, Roodveldt C, Tawfik DS. The “evolvability” of promiscuous protein functions. *Nat Genet*. 2005 Jan;37(1):73–6.
20. Fraser JS, van den Bedem H, Samelson AJ, Lang PT, Holton JM, Echols N, et al. Accessing protein conformational ensembles using room-temperature X-ray crystallography. *Proc Natl Acad Sci U S A*. 2011 Sep 27;108(39):16247–52.
21. Zhao D, Liu S, Gao Y. Single-molecule manipulation and detection. *Acta Biochim Biophys Sin*. 2018 Mar 1;50(3):231–7.
22. Ashkin A, Dziedzic JM, Bjorkholm JE, Chu S. Observation of a single-beam gradient force optical trap for dielectric particles. *Opt Lett*. 1986 May 1;11(5):288.
23. Neuman KC, Nagy A. Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nat Methods*. 2008 Jun;5(6):491–505.
24. Neuman KC, Block SM. Optical trapping. *Rev Sci Instrum*. 2004 Sep;75(9):2787–809.
25. Moffitt JR, Chemla YR, Smith SB, Bustamante C. Recent advances in optical tweezers. *Annu Rev Biochem*. 2008;77:205–28.
26. Fazal FM, Block SM. Optical tweezers study life under tension. *Nat Photonics*. 2011 May 31;5:318–21.
27. Svoboda K, Schmidt CF, Schnapp BJ, Block SM. Direct observation of kinesin stepping by optical trapping interferometry. *Nature*. 1993 Oct 21;365(6448):721–7.
28. Gelles J, Landick R. RNA polymerase as a molecular motor. *Cell*. 1998 Apr 3;93(1):13–6.
29. Kellermayer MS, Smith SB, Bustamante C, Granzier HL. Mechanical fatigue in repetitively stretched single molecules of titin. *Biophys J*. 2001 Feb;80(2):852–63.

30. Bergeron J, Zehtabi-Oskuie A, Ghaffari S, Pang Y, Gordon R. Optical trapping of nanoparticles. *J Vis Exp JoVE*. 2013 Jan 15;(71):e4424.
31. Pang Y, Gordon R. Optical trapping of a single protein. *Nano Lett*. 2012 Jan 11;12(1):402–6.
32. Geisse NA. AFM and combined optical techniques. *Mater Today*. 2009 Jul;12(7–8):40–5.
33. Hinterdorfer P, Baumgartner W, Gruber HJ, Schilcher K, Schindler H. Detection and localization of individual antibody-antigen recognition events by atomic force microscopy. *Proc Natl Acad Sci U S A*. 1996 Apr 16;93(8):3477–81.
34. Allen S, Chen X, Davies J, Davies MC, Dawkes AC, Edwards JC, et al. Detection of antigen-antibody binding events with the atomic force microscope. *Biochemistry (Mosc)*. 1997 Jun 17;36(24):7457–63.
35. Moy VT, Florin EL, Gaub HE. Intermolecular forces and energies between ligands and receptors. *Science*. 1994 Oct 14;266(5183):257–9.
36. Whited AM, Park PS-H. Atomic force microscopy: a multifaceted tool to study membrane proteins and their interactions with ligands. *Biochim Biophys Acta*. 2014 Jan;1838(1 Pt A):56–68.
37. Carvalho FA, Martins IC, Santos NC. Atomic force microscopy and force spectroscopy on the assessment of protein folding and functionality. *Arch Biochem Biophys*. 2013 Mar;531(1–2):116–27.
38. Oberhauser AF, Marszalek PE, Erickson HP, Fernandez JM. The molecular elasticity of the extracellular matrix protein tenascin. *Nature*. 1998 May 14;393(6681):181–5.
39. Rief M, Pascual J, Saraste M, Gaub HE. Single molecule force spectroscopy of spectrin repeats: low unfolding forces in helix bundles. *J Mol Biol*. 1999 Feb 19;286(2):553–61.
40. Ando T, Kodera N, Takai E, Maruyama D, Saito K, Toda A. A high-speed atomic force microscope for studying biological macromolecules. *Proc Natl Acad Sci U S A*. 2001 Oct 23;98(22):12468–72.
41. Sasmal DK, Pulido LE, Kasal S, Huang J. Single-molecule fluorescence resonance energy transfer in molecular biology. *Nanoscale*. 2016 Dec 8;8(48):19928–44.
42. Miller H, Zhou Z, Shepherd J, Wollman AJM, Leake MC. Single-molecule techniques in biophysics: a review of the progress in methods and applications. *Rep Prog Phys Phys Soc G B*. 2018;81(2):024601.
43. Götz M, Wortmann P, Schmid S, Hugel T. A Multicolor Single-Molecule FRET Approach to Study Protein Dynamics and Interactions Simultaneously. *Methods Enzymol*. 2016;581:487–516.

44. Deniz AA, Laurence TA, Beligere GS, Dahan M, Martin AB, Chemla DS, et al. Single-molecule protein folding: diffusion fluorescence resonance energy transfer studies of the denaturation of chymotrypsin inhibitor 2. *Proc Natl Acad Sci U S A*. 2000 May 9;97(10):5179–84.
45. Zhao Y, Terry D, Shi L, Weinstein H, Blanchard SC, Javitch JA. Single-molecule dynamics of gating in a neurotransmitter transporter homologue. *Nature*. 2010 May 13;465(7295):188–93.
46. Zhao Y, Terry DS, Shi L, Quick M, Weinstein H, Blanchard SC, et al. Substrate-modulated gating dynamics in a Na<sup>+</sup>-coupled neurotransmitter transporter homologue. *Nature*. 2011 Jun 2;474(7349):109–13.
47. Serghei A, Zhao W, Miranda D, Russell TP. Curie transitions for attograms of ferroelectric polymers. *Nano Lett*. 2013 Feb 13;13(2):577–80.
48. Houachtia A, Alcouffe P, Boiteux G, Seytre G, Gérard J-F, Serghei A. Nanofluidics Approach to Separate between Static and Kinetic Nanoconfinement Effects on the Crystallization of Polymers. *Nano Lett*. 2015 Jul 8;15(7):4311–6.
49. Soubias O, Gawrisch K. Rhodopsin-lipid interactions studied by NMR. *Methods Enzymol*. 2013;522:209–27.
50. Lazzara TD, Lau KHA, Knoll W, Janshoff A, Steinem C. Macromolecular shape and interactions in layer-by-layer assemblies within cylindrical nanopores. *Beilstein J Nanotechnol*. 2012;3:475–84.
51. de la Escosura-Muñiz A, Chunglok W, Surareungchai W, Merkoçi A. Nanochannels for diagnostic of thrombin-related diseases in human blood. *Biosens Bioelectron*. 2013 Feb 15;40(1):24–31.
52. de la Escosura-Muñiz A, Espinoza-Castañeda M, Chamorro-García A, Rodríguez-Hernández CJ, de Torres C, Merkoçi A. In situ monitoring of PTHLH secretion in neuroblastoma cells cultured onto nanoporous membranes. *Biosens Bioelectron*. 2018 Jun 1;107:62–8.
53. Jackson SE. How do small single-domain proteins fold? *Fold Des*. 1998;3(4):R81-91.
54. Lindorff-Larsen K, Piana S, Dror RO, Shaw DE. How fast-folding proteins fold. *Science*. 2011 Oct 28;334(6055):517–20.
55. Fomina M, Schirò G, Cupane A. Hydration dependence of myoglobin dynamics studied with elastic neutron scattering, differential scanning calorimetry and broadband dielectric spectroscopy. *Biophys Chem*. 2014 Jan;185:25–31.
56. Khodadadi S, Pawlus S, Sokolov AP. Influence of hydration on protein dynamics: combining dielectric and neutron scattering spectroscopy data. *J Phys Chem B*. 2008 Nov 13;112(45):14273–80.

57. Jansson H, Bergman R, Swenson J. Relation between solvent and protein dynamics as studied by dielectric spectroscopy. *J Phys Chem B*. 2005 Dec 22;109(50):24134–41.
58. Jansson H, Bergman R, Swenson J. Role of solvent for the dynamics and the glass transition of proteins. *J Phys Chem B*. 2011 Apr 14;115(14):4099–109.
59. Schirò G, Cupane A, Vitrano E, Bruni F. Dielectric relaxations in confined hydrated myoglobin. *J Phys Chem B*. 2009 Jul 16;113(28):9606–13.
60. Schirò G, Fomina M, Cupane A. Communication: Protein dynamical transition vs. liquid-liquid phase transition in protein hydration water. *J Chem Phys*. 2013 Sep 28;139(12):121102.
61. Lindorff-Larsen K, Best RB, Depristo MA, Dobson CM, Vendruscolo M. Simultaneous determination of protein structure and dynamics. *Nature*. 2005 Jan 13;433(7022):128–32.
62. Karplus M, McCammon JA. Molecular dynamics simulations of biomolecules. *Nat Struct Biol*. 2002 Sep;9(9):646–52.
63. Eisenmesser EZ, Bosco DA, Akke M, Kern D. Enzyme dynamics during catalysis. *Science*. 2002 Feb 22;295(5559):1520–3.
64. Rasmussen BF, Stock AM, Ringe D, Petsko GA. Crystalline ribonuclease A loses function below the dynamical transition at 220 K. *Nature*. 1992 Jun 4;357(6377):423–4.
65. Spangler BD. Structure and function of cholera toxin and the related *Escherichia coli* heat-labile enterotoxin. *Microbiol Rev*. 1992 Dec;56(4):622–47.
66. De Wolf MJ, Dams E, Dierick WS. Interaction of a cholera toxin derivative containing a reduced number of receptor binding sites with intact cells in culture. *Biochim Biophys Acta*. 1994 Sep 8;1223(2):296–305.
67. De Wolf MJ, Dierick WS. Regeneration of active receptor recognition domains on the B subunit of cholera toxin by formation of hybrids from chemically inactivated derivatives. *Biochim Biophys Acta*. 1994 Sep 8;1223(2):285–95.
68. McCann JA, Mertz JA, Czworkowski J, Picking WD. Conformational changes in cholera toxin B subunit-ganglioside GM1 complexes are elicited by environmental pH and evoke changes in membrane structure. *Biochemistry (Mosc)*. 1997 Jul 29;36(30):9169–78.
69. Tsai B, Rapoport TA. Unfolded cholera toxin is transferred to the ER membrane and released from protein disulfide isomerase upon oxidation by Ero1. *J Cell Biol*. 2002 Oct 28;159(2):207–16.
70. Hirst T. Biogenesis of Cholera toxin and Related oligomeric Enterotoxins. In: *Handbook of Natural Toxins*.

71. De Wolf M, Van Dessel G, Lagrou A, Hilderson HJ, Dierick W. Structural features of the binding site of cholera toxin inferred from fluorescence measurements. *Biochim Biophys Acta*. 1985 Nov 29;832(2):165–74.
72. Lesieur C, Cliff MJ, Carter R, James RFL, Clarke AR, Hirst TR. A kinetic model of intermediate formation during assembly of cholera toxin B-subunit pentamers. *J Biol Chem*. 2002 May 10;277(19):16697–704.
73. Ruddock LW, Ruston SP, Kelly SM, Price NC, Freedman RB, Hirst TR. Kinetics of acid-mediated disassembly of the B subunit pentamer of *Escherichia coli* heat-labile enterotoxin. Molecular basis of pH stability. *J Biol Chem*. 1995 Dec 15;270(50):29953–8.
74. Mudrak B, Kuehn MJ. Heat-labile enterotoxin: beyond G(m1) binding. *Toxins*. 2010;2(6):1445–70.
75. Qadri F, Svennerholm A-M, Faruque ASG, Sack RB. Enterotoxigenic *Escherichia coli* in developing countries: epidemiology, microbiology, clinical features, treatment, and prevention. *Clin Microbiol Rev*. 2005 Jul;18(3):465–83.
76. Sixma TK, Pronk SE, Kalk KH, Wartna ES, van Zanten BA, Without B, et al. Crystal structure of a cholera toxin-related heat-labile enterotoxin from *E. coli*. *Nature*. 1991 May 30;351(6325):371–7.
77. Zrimi J, Ng Ling A, Giri-Rachman Arifin E, Feverati G, Lesieur C. Cholera toxin B subunits assemble into pentamers--proposition of a fly-casting mechanism. *PloS One*. 2010 Dec 21;5(12):e15347.
78. Dyawanapelly S, Mehrotra P, Ghosh G, Jagtap DD, Dandekar P, Jain R. How the surface functionalized nanoparticles affect conformation and activity of proteins: Exploring through protein-nanoparticle interactions. *Bioorganic Chem*. 2019;82:17–25.
79. Masschalck B, Michiels CW. Antimicrobial properties of lysozyme in relation to foodborne vegetative bacteria. *Crit Rev Microbiol*. 2003;29(3):191–214.
80. Dobson CM, Evans PA, Radford SE. Understanding how proteins fold: the lysozyme story so far. *Trends Biochem Sci*. 1994 Jan;19(1):31–7.
81. Meersman F, Atilgan C, Miles AJ, Bader R, Shang W, Matagne A, et al. Consistent picture of the reversible thermal unfolding of hen egg-white lysozyme from experiment and molecular dynamics. *Biophys J*. 2010 Oct 6;99(7):2255–63.
82. Kazmirski SL, Daggett V. Non-native interactions in protein folding intermediates: molecular dynamics simulations of hen lysozyme. *J Mol Biol*. 1998 Dec 4;284(3):793–806.
83. Lai B, Cao A, Lai L. Organic cosolvents and hen egg white lysozyme folding. *Biochim Biophys Acta*. 2000 Nov 30;1543(1):115–22.



84. Cao A, Wang G, Tang Y, Lai L. Linear correlation between thermal stability and folding kinetics of lysozyme. *Biochem Biophys Res Commun*. 2002 Mar 8;291(4):795–7.
85. Poinern GEJ, Ali N, Fawcett D. Progress in Nano-Engineered Anodic Aluminum Oxide Membrane Development. *Mater Basel Switz*. 2011 Feb 25;4(3):487–526.
86. Houachtia A. Dielectric investigations of attograms and zeptograms of matter. Université de Lyon;
87. Stępniewski WJ, Nowak-Stępniewska A, Presz A, Czujko T, Varin RA. The effects of time and temperature on the arrangement of anodic aluminum oxide nanopores. *Mater Charact*. 2014 May;91:1–9.
88. Stępniewski WJ, Bojar Z. Synthesis of anodic aluminum oxide (AAO) at relatively high temperatures. Study of the influence of anodization conditions on the alumina structural features. *Surf Coat Technol*. 2011 Oct;206(2–3):265–72.
89. Belwalkar A, Grasing E, Van Geertruyden W, Huang Z, Misiolak WZ. Effect of Processing Parameters on Pore Structure and Thickness of Anodic Aluminum Oxide (AAO) Tubular Membranes. *J Membr Sci*. 2008 Jul;319(1–2):192–8.
90. Schönhals A, Kremer F. *Broadband Dielectric Spectroscopy*.
91. Pacini L, Bourgeat L, Serghei A, Lesieur C. Analysis of Nanoconfined Protein Dielectric Signals Using Charged Amino Acid Network Models. *Aust J Chem* [Internet]. 2020 [cited 2020 Jun 17]; Available from: <http://www.publish.csiro.au/?paper=CH19502>
92. Williams R. *An Introduction to Fluorescence Spectroscopy*. Perkin-Elmer; 1981.
93. Sauer M, Hofkens J, Enderlein J. *Basic Principles of Fluorescence Spectroscopy*. Wiley online Library; 2011.
94. Zrimi J, Ng Ling A, Giri-Rachman Arifin E, Feverati G, Lesieur C. Cholera toxin B subunits assemble into pentamers--proposition of a fly-casting mechanism. *PLoS One*. 2010 Dec 21;5(12):e15347.
95. Li S, Li D. Investigation on the pH-dependent binding of benzocaine and lysozyme by fluorescence and absorbance. *Spectrochim Acta A Mol Biomol Spectrosc*. 2011 Nov;82(1):396–405.
96. Fischer H, Polikarpov I, Craievich AF. Average protein density is a molecular-weight-dependent function. *Protein Sci Publ Protein Soc*. 2004 Oct;13(10):2825–8.
97. Serghei A, Zhao W, Wei X, Chen D, Russell TP. Nanofluidics with phase separated block-copolymers: Glassy dynamics during capillary flow. *Eur Phys J Spec Top*. 2010 Oct;189(1):95–101.
98. Smiszek-Lindert WE, Kamińska E, Minecka A, Heczko D, Madejczyk O, Tarnacka M, et al. Studies on dynamics and isomerism in supercooled photochromic compound



- Aberchrome 670 with the use of different experimental techniques. *Phys Chem Chem Phys PCCP*. 2018 Jul 14;20(26):18009–19.
99. Barends TRM, Foucar L, Ardevol A, Nass K, Aquila A, Botha S, et al. Direct observation of ultrafast collective motions in CO myoglobin upon ligand dissociation. *Science*. 2015 Oct 23;350(6259):445–50.
  100. Levantino M, Schirò G, Lemke HT, Cottone G, Glowia JM, Zhu D, et al. Ultrafast myoglobin structural dynamics observed with an X-ray free-electron laser. *Nat Commun*. 2015 Apr 2;6:6772.
  101. Mizohata E, Nakane T, Fukuda Y, Nango E, Iwata S. Serial femtosecond crystallography at the SACLA: breakthrough to dynamic structural biology. *Biophys Rev*. 2018 Apr;10(2):209–18.
  102. Stagno JR, Liu Y, Bhandari YR, Conrad CE, Panja S, Swain M, et al. Structures of riboswitch RNA reaction states by mix-and-inject XFEL serial crystallography. *Nature*. 2017 12;541(7636):242–6.
  103. Herbst J, Heyne K, Diller R. Femtosecond infrared spectroscopy of bacteriorhodopsin chromophore isomerization. *Science*. 2002 Aug 2;297(5582):822–5.
  104. Kolano C, Helbing J, Kozinski M, Sander W, Hamm P. Watching hydrogen-bond dynamics in a  $\beta$ -turn by transient two-dimensional infrared spectroscopy. *Nature*. 2006 Nov;444(7118):469–72.
  105. Mizutani Y, Kitagawa T. Ultrafast dynamics of myoglobin probed by time-resolved resonance Raman spectroscopy. *Chem Rec N Y N*. 2001;1(3):258–75.
  106. Fang C, Frontiera RR, Tran R, Mathies RA. Mapping GFP structure evolution during proton transfer with femtosecond Raman spectroscopy. *Nature*. 2009 Nov 12;462(7270):200–4.
  107. Schneider R, Blackledge M, Jensen MR. Elucidating binding mechanisms and dynamics of intrinsically disordered protein complexes using NMR spectroscopy. *Curr Opin Struct Biol*. 2019;54:10–8.
  108. Findsen EW, Scott TW, Chance MR, Friedman JM, Ondrias MR. Picosecond time-resolved Raman studies of photodissociated carboxymyoglobin. *J Am Chem Soc*. 1985 May;107(11):3355–7.
  109. Heath GR, Scheuring S. High-speed AFM height spectroscopy reveals  $\mu$ s-dynamics of unlabeled biomolecules. *Nat Commun*. 2018 26;9(1):4983.
  110. Hellenkamp B, Wortmann P, Kandzia F, Zacharias M, Hugel T. Multidomain structure and correlated dynamics determined by self-consistent FRET networks. *Nat Methods*. 2017;14(2):174–80.

111. Soranno A, Holla A, Dingfelder F, Nettels D, Makarov DE, Schuler B. Integrated view of internal friction in unfolded proteins from single-molecule FRET, contact quenching, theory, and simulations. *Proc Natl Acad Sci U S A*. 2017 07;114(10):E1833–9.
112. Lesieur C, Schulten K. Editorial overview: Theory and simulation. *Curr Opin Struct Biol*. 2015 Apr;31:v–vi.
113. Perilla JR, Goh BC, Cassidy CK, Liu B, Bernardi RC, Rudack T, et al. Molecular dynamics simulations of large macromolecular complexes. *Curr Opin Struct Biol*. 2015 Apr;31:64–74.
114. Li C-B, Yang H, Komatsuzaki T. Multiscale complex network of protein conformational fluctuations in single-molecule time series. *Proc Natl Acad Sci U S A*. 2008 Jan 15;105(2):536–41.
115. Achoch M, Dorantes-Gilardi R, Wymant C, Feverati G, Salamatian K, Vuillon L, et al. Protein structural robustness to mutations: an in silico investigation. *Phys Chem Chem Phys PCCP*. 2016 18;18(20):13770–80.
116. Gheeraert A, Pacini L, Batista VS, Vuillon L, Lesieur C, Rivalta I. Exploring Allosteric Pathways of a V-Type Enzyme with Dynamical Perturbation Networks. *J Phys Chem B*. 2019 Apr 25;123(16):3452–61.
117. Degiacomi MT, Iacovache I, Pernot L, Chami M, Kudryashev M, Stahlberg H, et al. Molecular assembly of the aerolysin pore reveals a swirling membrane-insertion mechanism. *Nat Chem Biol*. 2013 Oct;9(10):623–9.
118. Degiacomi MT. Coupling Molecular Dynamics and Deep Learning to Mine Protein Conformational Space. *Struct Lond Engl 1993*. 2019 04;27(6):1034–1040.e3.
119. Marklund EG, Benesch JL. Weighing-up protein dynamics: the combination of native mass spectrometry and molecular dynamics simulations. *Curr Opin Struct Biol*. 2019;54:50–8.
120. Boyd CM, Bubeck D. Advances in cryoEM and its impact on  $\beta$ -pore forming proteins. *Curr Opin Struct Biol*. 2018;52:41–9.
121. Pak AJ, Voth GA. Advances in coarse-grained modeling of macromolecular complexes. *Curr Opin Struct Biol*. 2018;52:119–26.
122. Parsons ES, Stanley GJ, Pyne ALB, Hodel AW, Nievergelt AP, Menny A, et al. Single-molecule kinetics of pore assembly by the membrane attack complex. *Nat Commun*. 2019 06;10(1):2066.
123. van den Bedem H, Fraser JS. Integrative, dynamic structural biology at atomic resolution—it's about time. *Nat Methods*. 2015 Apr;12(4):307–18.
124. Bourgeat L, Serghei A, Lesieur C. Experimental Protein Molecular Dynamics: Broadband Dielectric Spectroscopy coupled with nanoconfinement. *Sci Rep*. 2019 Nov 29;9(1):17988.

125. Razvi A, Scholtz JM. Lessons in stability from thermophilic proteins. *Protein Sci Publ Protein Soc.* 2006 Jul;15(7):1569–78.
126. Karshikoff A, Nilsson L, Ladenstein R. Rigidity versus flexibility: the dilemma of understanding protein thermal stability. *FEBS J.* 2015 Oct;282(20):3899–917.
127. Dalziel AW, Lipka G, Chowdhry BZ, Sturtevant JM, Schafer DE. Effects of ganglioside GM1 on the thermotropic behavior of cholera toxin B subunit. *Mol Cell Biochem.* 1984 Aug;63(1):83–91.
128. Goins B, Freire E. Thermal stability and intersubunit interactions of cholera toxin in solution and in association with its cell-surface receptor ganglioside GM1. *Biochemistry (Mosc).* 1988 Mar 22;27(6):2046–52.
129. De Wolf MJ, Van Dessel GA, Lagrou AR, Hilderson HJ, Dierick WS. pH-induced transitions in cholera toxin conformation: a fluorescence study. *Biochemistry (Mosc).* 1987 Jun 30;26(13):3799–806.
130. Dorantes-Gilardi R, Bourgeat L, Pacini L, Vuillon L, Lesieur C. In proteins, the structural responses of a position to mutation rely on the Goldilocks principle: not too many links, not too few. *Phys Chem Chem Phys PCCP.* 2018 Oct 10;20(39):25399–410.
131. Prestrelski SJ, Tedeschi N, Arakawa T, Carpenter JF. Dehydration-induced conformational transitions in proteins and their inhibition by stabilizers. *Biophys J.* 1993 Aug;65(2):661–71.
132. Frauenfelder H, Chen G, Berendzen J, Fenimore PW, Jansson H, McMahon BH, et al. A unified model of protein dynamics. *Proc Natl Acad Sci U S A.* 2009 Mar 31;106(13):5129–34.
133. Iavarone AT, Patriksson A, van der Spoel D, Parks JH. Fluorescence probe of Trp-cage protein conformation in solution and in gas phase. *J Am Chem Soc.* 2007 May 30;129(21):6726–35.
134. Dominy BN, Minoux H, Brooks CL. An electrostatic basis for the stability of thermophilic proteins. *Proteins.* 2004 Oct 1;57(1):128–41.
135. Xia Y-L, Sun J-H, Ai S-M, Li Y, Du X, Sang P, et al. Insights into the role of electrostatics in temperature adaptation: a comparative study of psychrophilic, mesophilic, and thermophilic subtilisin-like serine proteases. *RSC Adv.* 2018;8(52):29698–713.
136. Kumar S, Ma B, Tsai CJ, Nussinov R. Electrostatic strengths of salt bridges in thermophilic and mesophilic glutamate dehydrogenase monomers. *Proteins.* 2000 Mar 1;38(4):368–83.
137. Xu G, Cheng K, Wu Q, Liu M, Li C. Confinement Alters the Structure and Function of Calmodulin. *Angew Chem Int Ed Engl.* 2017 09;56(2):530–4.

138. Rackham OJL, Madera M, Armstrong CT, Vincent TL, Woolfson DN, Gough J. The evolution and structure prediction of coiled coils across all genomes. *J Mol Biol.* 2010 Oct 29;403(3):480–93.
139. Silveira JA, Fort KL, Kim D, Servage KA, Pierson NA, Clemmer DE, et al. From solution to the gas phase: stepwise dehydration and kinetic trapping of substance P reveals the origin of peptide conformations. *J Am Chem Soc.* 2013 Dec 26;135(51):19147–53.
140. Hekstra DR, White KI, Socolich MA, Henning RW, Šrajter V, Ranganathan R. Electric-field-stimulated protein mechanics. *Nature.* 2016 15;540(7633):400–5.
141. Bhakuni V, Xie D, Freire E. Thermodynamic identification of stable folding intermediates in the B-subunit of cholera toxin. *Biochemistry (Mosc).* 1991 May 21;30(20):5055–60.
142. Hernandez G, Jenney FE, Adams MW, LeMaster DM. Millisecond time scale conformational flexibility in a hyperthermophile protein at ambient temperature. *Proc Natl Acad Sci U S A.* 2000 Mar 28;97(7):3166–70.
143. Roberts D, Keeling R, Tracka M, van der Walle CF, Uddin S, Warwicker J, et al. Specific ion and buffer effects on protein-protein interactions of a monoclonal antibody. *Mol Pharm.* 2015 Jan 5;12(1):179–93.
144. Tsai CJ, Ma B, Nussinov R. Folding and binding cascades: shifts in energy landscapes. *Proc Natl Acad Sci U S A.* 1999 Aug 31;96(18):9970–2.
145. Boehr DD, McElheny D, Dyson HJ, Wright PE. The dynamic energy landscape of dihydrofolate reductase catalysis. *Science.* 2006 Sep 15;313(5793):1638–42.
146. Ma B, Nussinov R. Enzyme dynamics point to stepwise conformational selection in catalysis. *Curr Opin Chem Biol.* 2010 Oct;14(5):652–9.
147. James LC. Antibody Multispecificity Mediated by Conformational Diversity. *Science.* 2003 Feb 28;299(5611):1362–7.
148. Yogurtcu ON, Erdemli SB, Nussinov R, Turkay M, Keskin O. Restricted mobility of conserved residues in protein-protein interfaces in molecular simulations. *Biophys J.* 2008 May 1;94(9):3475–85.
149. Ravindra R, Zhao S, Gies H, Winter R. Protein encapsulation in mesoporous silicate: the effects of confinement on protein stability, hydration, and volumetric properties. *J Am Chem Soc.* 2004 Oct 6;126(39):12224–5.
150. Yeung PS-W, Eskici G, Axelsen PH. Infrared spectroscopy of proteins in reverse micelles. *Biochim Biophys Acta.* 2013 Oct;1828(10):2314–8.
151. Minton AP. The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media. *J Biol Chem.* 2001 Apr 6;276(14):10577–80.

152. Xu G, Zhao J, Cheng K, Wu Q, Liu X, Liu M, et al. The Effects of Macromolecular Crowding on Calmodulin Structure and Function. *Chem Weinh Bergstr Ger*. 2017 May 17;23(28):6736–40.
153. Demir Ö, Baronio R, Salehi F, Wassman CD, Hall L, Hatfield GW, et al. Ensemble-based computational approach discriminates functional activity of p53 cancer and rescue mutants. *PLoS Comput Biol*. 2011 Oct;7(10):e1002238.
154. Vendruscolo M. Determination of conformationally heterogeneous states of proteins. *Curr Opin Struct Biol*. 2007 Feb;17(1):15–20.
155. Tóth-Petróczy A, Tawfik DS. The robustness and innovability of protein folds. *Curr Opin Struct Biol*. 2014 Jun;26:131–8.
156. Payne JL, Wagner A. The robustness and evolvability of transcription factor binding sites. *Science*. 2014 Feb 21;343(6173):875–7.
157. McLaughlin RN, Poelwijk FJ, Raman A, Gosal WS, Ranganathan R. The spatial architecture of protein function and adaptation. *Nature*. 2012 Nov 1;491(7422):138–42.
158. Salinas VH, Ranganathan R. Coevolution-based inference of amino acid interactions underlying protein function. *eLife*. 2018 20;7.
159. Matković-Calogović D, Loregian A, D'Acunto MR, Battistutta R, Tossi A, Palù G, et al. Crystal structure of the B subunit of Escherichia coli heat-labile enterotoxin carrying peptides with anti-herpes simplex virus type 1 activity. *J Biol Chem*. 1999 Mar 26;274(13):8764–9.
160. Merritt EA, Sarfaty S, van den Akker F, L'Hoir C, Martial JA, Hol WG. Crystal structure of cholera toxin B-pentamer bound to receptor GM1 pentasaccharide. *Protein Sci Publ Protein Soc*. 1994 Feb;3(2):166–75.
161. Mark AE, van Gunsteren WF. Simulation of the thermal denaturation of hen egg white lysozyme: trapping the molten globule state. *Biochemistry (Mosc)*. 1992 Sep 1;31(34):7745–8.



OPEN

# Experimental Protein Molecular Dynamics: Broadband Dielectric Spectroscopy coupled with nanoconfinement

Laëtitia Bourgeat<sup>1,2</sup>, Anatoli Serghei<sup>2\*</sup> & Claire Lesieur<sup>1,3\*</sup>

Protein dynamics covers multiple spatiotemporal scale processes, among which slow motions, not much understood even though they are underlying protein folding and protein functions. Protein slow motions are associated with structural heterogeneity, short-lived and poorly populated conformations, hard to detect individually. In addition, they involve collective motions of many atoms, not easily tracked by simulation and experimental devices. Here we propose a biophysical approach, coupling geometrical nanoconfinement and broadband dielectric spectroscopy (BDS), which distinguishes protein conformations by their respective molecular dynamics. In particular, protein-unfolding intermediates, usually poorly populated in macroscopic solutions are detected. The protein dynamics is observed under unusual conditions (sample nanoconfinement and dehydration) highlighting the robustness of protein structure and protein dynamics to a variety of conditions consistent with protein sustainability. The protein dielectric signals evolve with the temperature of thermal treatments indicating sensitivity to atomic and molecular interaction changes triggered by the protein thermal unfolding. As dipole fluctuations depend on both collective large-scale motions and local motions, the approach offers a prospect to track in-depth unfolding events.

Proteins produce biological activities in living organisms thanks to their 3D-structure and the associated dynamics. The 3D structure relies on chemical interactions between atoms of the amino acids that compose a protein. The molecular dynamics is based on molecular fluctuations, which produce motions from a local scale (amino acid side chain fluctuations) to a larger scale (structural relaxation) to respond to the protein folding and the protein function<sup>1-3</sup>. Accordingly, protein dynamics covers several orders of magnitude of spatiotemporal motions from femtosecond to second for molecular fluctuations from Angströms to nanometers. Basically, fast motions (femtosecond to nanosecond) concern local atomic motions (vibration, side chain motions) while slow motions concern the collective motions of many atoms within larger size areas such as secondary and tertiary structural elements (microsecond to millisecond) up to domains (millisecond to second)<sup>3</sup>.

To track protein dynamics, several experimental and theoretical approaches are required. Ultra resolution X-ray crystallography and x-ray laser have been successfully applied on few cases to monitor time-resolution of conformational motions<sup>4-7</sup>. Ultra fast NMR, femtosecond simulated Raman spectroscopy and ultra fast transient Infra Red spectroscopy are measuring atomic motions from femtosecond to nanosecond<sup>8-13</sup>. Slower motions above nanosecond are monitored by low spatial resolution techniques such as AFM and fluorescence spectroscopy (FRET) where the molecular dynamics are inaccessible<sup>14-16</sup>.

Theoretical approaches such as molecular dynamics (MD) simulations have atomic resolution and cover motions up to millisecond now thanks to ANTON super computer<sup>1</sup>. But MD simulation at high resolution is limited in terms of size and it remains hard to assess motions above microsecond range<sup>17,18</sup>. Network models are also useful to study protein dynamics<sup>19-22</sup>.

Most successful approaches are integrative, combining experimental and multiple theoretical approaches in order to cover more length and time scale motions of protein structure dynamics<sup>23-25</sup>. They are particularly successful in monitoring the slow motions involved in protein assembly and pore-formation<sup>26-32</sup>.

<sup>1</sup>AMPERE, CNRS, Univ. Lyon, 69622, Lyon, France. <sup>2</sup>IMP, CNRS, Univ. Lyon, 69622, Lyon, France. <sup>3</sup>Institut Rhônealpin des systèmes complexes, IXXI-ENS-Lyon, 69007, Lyon, France. \*email: [anatoli.serghei@univ-lyon1.fr](mailto:anatoli.serghei@univ-lyon1.fr); [claire.lesieur@ens-lyon.fr](mailto:claire.lesieur@ens-lyon.fr)

Measuring the spatiotemporal dynamics of slow collective motions in proteins is crucial because it is the range where protein folding and biological activity take place. It is also the range to assess the health state of a protein as pathological mutations are distinguished from robust mutations by different collective motions<sup>33</sup>. Protein robustness to mutations or to external perturbations such as environmental changes relies on the diversity of structures adopted by a protein in order to function<sup>34–37</sup>. But there are little measurements of such structural diversity of a protein or of its sequence variants on which to explore the set of atomic and molecular interactions that sustain robustness and from which to infer potential structural and dynamics fragility<sup>24,38,39</sup>. This is because slow collective motions associate with short-lived poorly populated conformations or conformations globally close requiring high sensitivity and probing of atomic interaction fluctuations<sup>40</sup>.

To overcome these difficulties, broadband dielectric spectroscopy (BDS) was used to monitor dipole fluctuations in the frequency range from 1 Hz to 10<sup>6</sup> Hz relevant to slow motions and combined with nanoconfinement for lowering the protein dynamics and enriching homogeneous conformations. The nanoconfinement was initially developed to analyze attograms (1 attogram = 10<sup>-18</sup> g) and zeptograms (1 zeptogram = 10<sup>-21</sup> g) of matter and investigate the dielectric behavior of matter on a length-scale becoming already comparable with molecular dimensions<sup>41,42</sup>. The characteristic volume of 1 attogram (considering a density of ~1 g/cm<sup>3</sup>) is ~10 × 10 × 10 nm, approaching thus the characteristic dimensions of one protein and offering the possibility to measure protein dynamics at the scale of the protein instead of the scale of a population of proteins.

Nanoconfined cholera toxin B subunit pentamers (CtxB<sub>5</sub>) were incubated at different temperatures from 60 °C to 180 °C and the dielectric signal measured after each thermal treatment. The toxin dielectric loss  $\epsilon''$  varies with the thermal treatment showing the technic sensitivity to atomic and molecular interaction changes induced by the toxin thermal unfolding. To the best of our knowledge, investigating protein dynamics using nanoconfinement and BDS on the scale of attograms has never been performed before.

## Material and Methods

**Materials.** Lyophilized cholera toxin B pentamer purchased from Sigma Aldrich (C9903) was diluted in phosphate buffered saline (PBS; 10 mM sodium phosphate, 150 mM sodium chloride, pH 7.4) at a final concentration of 2.6 mg/ml. This bulk toxin preparation was aliquoted and kept at -20 °C, thawed and frozen twice maximum, when used for the dielectric measurements. Ten-time concentrated solution of PBS solutions were purchased from BIOSOLVE and diluted 10-times in distilled water to prepare the PBS used to prepare the bulk toxin solution. In addition, the PBS solution was filtered in 22  $\mu$ m filter to limit impurity contamination. The Aluminium Oxide (AAO) Films on Al are purchased from InRedox. The AAO membrane dimensions are 10 mm \* 10 mm, and they contain nanopores of 40 nm-diameters and 10- $\mu$ m lengths, with 12% porosity. The AAO membrane containing the nanopores is referred to as the nanomembrane. Before use, the nanomembranes were plasma treated for twenty minutes to remove organic impurities.

**Methods.** Protein sample preparation- For the dielectric measurement, a protein sample at 0,025 mg/ml was used. This protein sample was prepared by diluting 2  $\mu$ l of the bulk toxin solution at 2,6 mg/ml in 204  $\mu$ l of deionized water (Final pH 6,9). Since the bulk protein is prepared in PBS, the sample for the dielectric measurement contains protein, deionized water and PBS. A 200  $\mu$ l drop of sample is deposited on the nanomembrane using the drop technique, then heated at 50 °C (323 K) for 15 min to evaporate the bulk water and allow proteins to enter the pores. The sample is cooled down to 30 °C (303 K) for 5 minutes with a speed of 2 K/min, placed between two Al electrodes of 7 mm and 40 mm diameter, respectively and in the cell holder for measurement. The procedure is the same for the control samples.

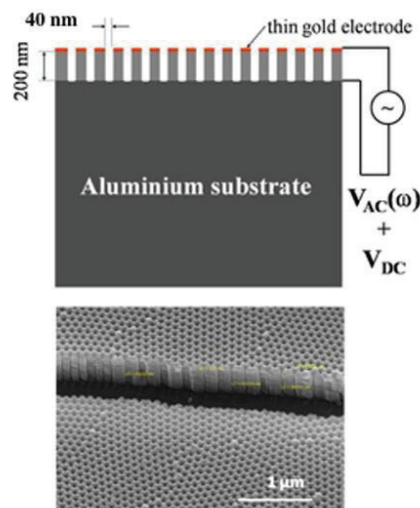
**Dielectric measurements.** The dielectric measurements were performed on a broadband dielectric spectrometer Novocontrol Alpha analyzer over a frequency range from 1 Hz to 10<sup>6</sup> Hz and over a temperature range from -80 °C (193 K) to 180 °C (453 K). The temperature ramps were carried-out with a rate of 2 K/min and a voltage of 0.2 V was applied. The dielectric loss  $\epsilon''$ , the imaginary part of the complex dielectric permittivity, was measured as a function of temperature at different frequencies. For the temperature control, a flow of pure nitrogen gas was used in a closed cryostat, providing water-free and oxygen-free experimental conditions. The sample was first heated to 60 °C (333 K) for three hours, then cooled down to -80 °C (193 K) (Cooling, C) and maintained for 30 min before being heated back to 60 °C (Heating, H). A cycle is composed of the cooling (C) and the heating (H) temperature ramps. The sample was heated again for three hours at 60 °C and a second cycle of cooling/heating, repeated to measure the reproducibility of the signal. The curves of the cooling and heating signals of the cycle 1 and of the cycle 2 are almost superimposed for the protein sample (Supplementary Fig. S1A). This demonstrates that the proteins under study are stabilized in a state where no bulk water is present anymore and where the amount of adsorbed water is stabilized as well. It also indicates that three hours incubation at 60 °C is enough for the protein to reach a stable conformational state and this incubation time can be used for thermal treatments at higher temperatures. In contrast, the cooling signal of the PBS buffer control is higher than the heating signal showing that evaporation of the bulk water still occurs in the buffer sample after the first three hours at 60 °C (Supplementary Fig. S1B). Since this is not the case for the protein sample, it suggests more adsorbed water in the protein sample. At 1 Hz, the same is observed (Supplementary Fig. S1C and D).

From 60 °C to 180 °C, little hysteresis is observed indicating that the dielectric loss  $\epsilon''$  does not detect different cooling and heating processes (Supplementary Fig. S2A). The protein remains in a steady state across the measurement and undergoes thermal unfolding mainly during the three hours incubation time.

The thermal denaturation of the toxin is studied using temperature ramps with a series of 3-hours thermal treatments at 100 °C (373 K), 140 °C (413 K) and 180 °C (453 K) followed by cooling/heating cycles after each treatment.

It is worth noting that one full dielectric experiment uses as little as ~5  $\mu$ g of toxin.





**Figure 1.** Nanopores as experimental cells to hold and measure attograms of material. The nanopores are identical, independent, and additive. Upper panel. Schematic representation of the sample cell. The lower electrode and the upper electrode are connected to the dielectric spectrometer and used to measure -in dependence on frequency and temperature- the permittivity of the sample cell. Lower panel. SEM image of the nanocontainers in cross section.

**Dielectric data.** The temperature dependence of the relaxation processes is modeled by an Arrhenius equation expressed as (Eq. 1)<sup>43,44</sup>.

$$\tau = \tau_0 \exp(E_A/k_B T) \quad (1)$$

where  $\tau$  is the relaxation time,  $\tau_0$  the relaxation time at high temperatures,  $T$  the temperature,  $E_A$  the activation energy and  $k_B$  the Boltzmann constant<sup>45</sup>.

**Control experiments.** The dielectric response of the empty nanomembrane was measured for 10 min at 60 °C and from 60 °C to –80 °C to control that no dielectric dispersions are detected. The dielectric response of nanopores filled with a 200  $\mu$ l drop of deionized water was measured for 10 min at 60 °C and from 60 °C to –80 °C to control that the deionized water used to prepare the sample show no dielectric dispersions too. Finally, the dielectric response of nanopores filled with a 200  $\mu$ l drop of PBS buffer control, is measured following a complete temperature ramp experiment as for the protein sample. The PBS buffer control is prepared with 2  $\mu$ l of PBS diluted in 204  $\mu$ l of deionized water (no protein is added). This concentration is equivalent to the concentration used to prepare and measure the protein sample. This control experiment was carried out to investigate the dielectric response of PBS mixed with deionized water in the same conditions as the protein.

To make sure the toxin is in a native state when a dielectric experiment is performed, the toxin pentameric state is checked by Trp-fluorescence beforehand<sup>46</sup>. The sample cannot be recovered from the nanopores after the BDS experiment so it is not possible to control the state of the toxin after.

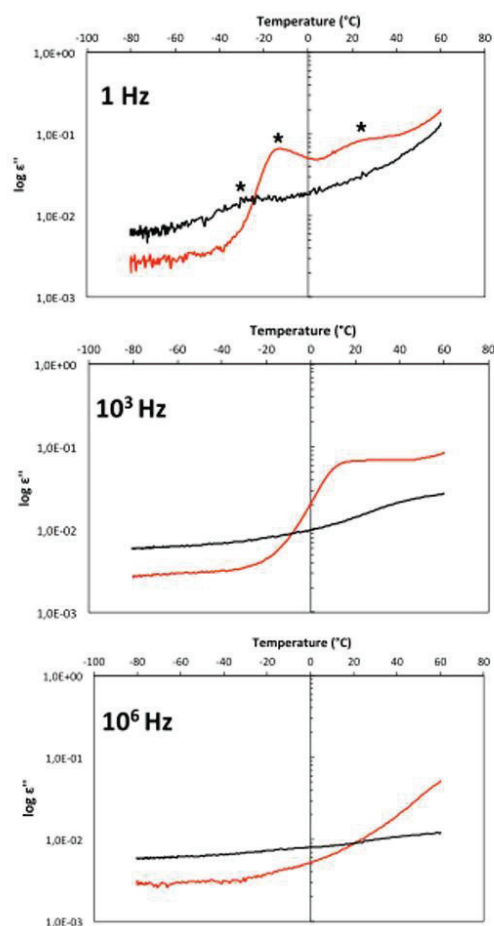
## Results

The goal of the study is to determine if combining nanoconfinement and broadband dielectric spectroscopy (BDS) allows distinguishing protein conformations by measuring different molecular dynamics. To test such a possibility, the cholera toxin B pentamer (CtxB<sub>5</sub>) is chosen because of its high stability and functional resistance to freeze-drying, which shows that the dehydration of the toxin does not lead to irreversible conformational changes. This is important since the dielectric measurements are performed under dehydrated conditions where the toxin sample contains PBS and only adsorbed water (Methods and Supplementary Fig. S1).

Toxin and PBS buffer control samples are deposited in the nanomembrane and the dielectric loss  $\epsilon''$  of the sample is measured by BDS, the whole setup is shown on Fig. 1. The samples are incubated for three hours at 60 °C and the dielectric loss  $\epsilon''$  is measured as a function of temperature (Fig. 2) (Methods). The protein and the PBS buffer control spectra can be described by several relaxation processes, which give rise to peaks in the spectra of the dielectric loss (Fig. 2, \*).

At 1 Hz (Fig. 2, top panel), the protein sample presents two dipole relaxation processes and hence two molecular dynamics (MDs). The main one, narrow, is the fastest with a maximum temperature position at –14 °C (MD<sub>1</sub>). The second, broader and less distinct, is slower with a maximum temperature position at 30 °C (MD<sub>2</sub>).

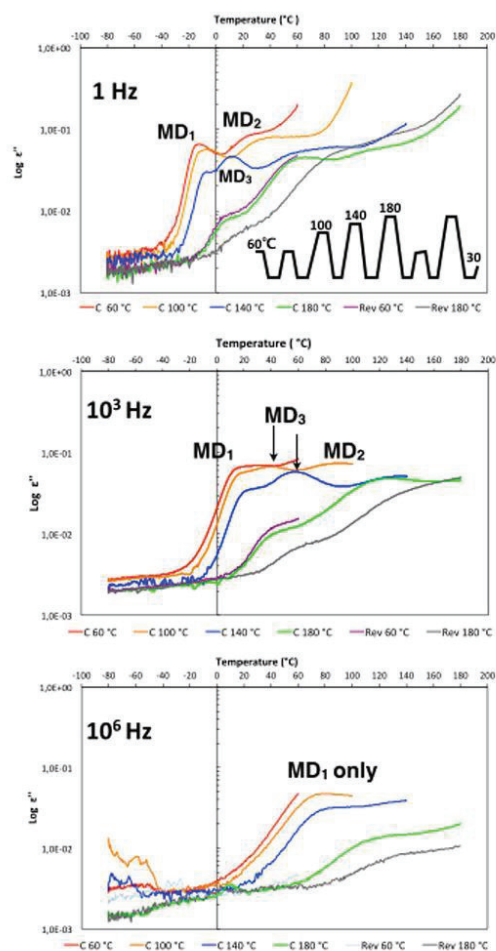




**Figure 2.** Dielectric loss as a function of temperature, at different frequencies (as indicated) for Cholera toxin B pentamer (red) and PBS buffer control (black), after a thermal treatment of 3-hours at 60 °C.

The broadness of  $MD_2$  suggests a relaxation process associated with a heterogeneous population of toxin conformations while the narrower  $MD_1$  peak suggests less heterogeneous toxin conformations. The buffer control also presents a relaxation process and a molecular dynamic ( $MD_b$ ) with a maximum dielectric loss  $\epsilon''$  at  $-27$  °C, so a fastest process than those observed for the toxin. At  $10^3$  Hz (Fig. 2, middle panel), a relaxation process is still detectable for the protein with a maximum dielectric loss  $\epsilon''$  at  $13$  °C ( $MD_1$ ). For the PBS buffer control, a relaxation process is also observed but not well enough to have an accurate maximum temperature position (Fig. 2, middle panel). The maximum temperature positions shift to higher temperatures with increasing frequencies. At  $10^6$  Hz, neither the protein nor the PBS samples have a detectable signal in the experimental window of our investigations (Fig. 2, bottom panel).

To test whether the dielectric losses probe protein conformation and the relaxation processes two different toxin conformations, the toxin is submitted to thermal denaturation (Methods). Globular proteins go from well-packed structures with numerous atomic and molecular interactions that hold their native state at low temperatures to less compact unfolded structures at high temperatures, where conformations loosing atomic and molecular interactions open to solvents and conserved only residual structures<sup>47,48</sup>. Thermal stability significantly varies for proteins from mesophilic to thermophilic organisms<sup>47,49</sup>. The CtxB<sub>3</sub> thermal stability has been studied using differential scanning calorimetry and it exhibits a single transition centered at  $74$  °C<sup>50,51</sup>. The thermal treatments disturb the toxin native atomic and molecular interactions and are hence expected to alter dipole and/or dipole environments, and consequently modify the global dielectric losses.

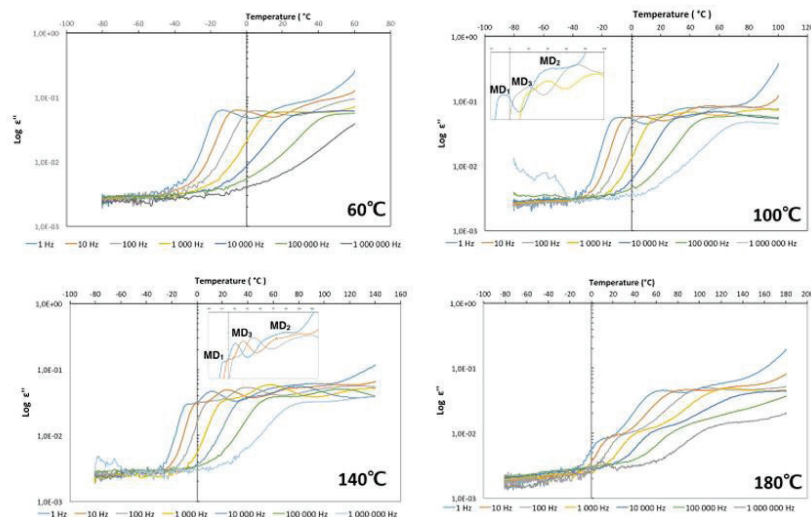


**Figure 3.** Dielectric loss as a function of temperature for Cholera toxin B pentamer, at indicated frequencies and after thermal treatments of 3-hours at 60 °C, 100 °C, 140 °C and 180 °C. Those are captioned as C 60 °C, C 100 °C, C 140 °C and C 180 °C, C for cooling, because the cooling signal is shown. After the three hours incubation at 180 °C and cooling to −80 °C, the sample is warmed back up to 60 °C for 3-hours and the subsequent cooling signal is shown, captioned Rev 60 °C for reversibility. Then the sample is warmed for 3-hours at 180 °C, and the subsequent cooling signal is shown, captioned Rev 180 °C.

Three hours thermal treatments at 100 °C, 140 °C and 180 °C are performed since it is enough time for the protein to reach a stable state at 60 °C (Methods and Supplementary Fig. S1A). After the incubation step, the dielectric loss  $\epsilon''$  is measured from the incubation temperature (100 °C, 140 °C or 180 °C) to −80 °C (Cooling) and from −80 °C back to the incubation temperature (Heating) (Fig. 3). The dielectric loss  $\epsilon''$  is shown as a function of temperature for the cooling process but both the cooling and heating signals are similar over the four temperatures of the thermal treatments (Methods and Supplementary Fig. S2A).

The thermal treatments leads to a loss of the relaxation process of the PBS buffer control and only conductivity is observed after treatment at 100 °C and 1 Hz frequency (Supplementary Fig. S2B). In contrast, the toxin maintains multiple relaxation processes evolving across the thermal treatments up to 180 °C and over the frequencies from 1 Hz to 10<sup>6</sup> Hz, confirming that the dielectric loss of the protein is conformation specific and sensitive to atomic and molecular interaction changes (Fig. 3).

At 1 Hz, the dielectric losses are similar after the thermal treatments at 60 °C and at 100 °C, suggesting that the detected molecular dynamics and protein conformations are stable up to 100 °C. The maximum temperature



**Figure 4.** Dielectric loss as a function of temperature for Cholera toxin B pentamer, at frequencies from 1 Hz to  $10^6$  Hz after three hours thermal treatments at 60 °C (Top left), 100 °C (Top right), 140 °C (Bottom left) and 180 °C (bottom right). Only few frequencies are shown in the insets for the sake of clarity.

positions of MD<sub>1</sub> shift only from  $-14$  °C to  $-7$  °C and from 30 °C to 40 °C for the MD<sub>2</sub>. In the same time, the MD<sub>2</sub> peak becomes even broader.

After incubation at 140 °C and at 180 °C, the dielectric losses change significantly more. The intensity of the dielectric loss of MD<sub>1</sub> drops at 140 °C and even more at 180 °C with, in addition to a broadening of the peak; a shift of the maximum temperature positions from  $-7$  °C to 10 °C at 180 °C. The drop of the dielectric loss intensity of MD<sub>1</sub> suggests a disappearance of the average conformations associated with MD<sub>1</sub> as the temperature rises above 100 °C.

The MD<sub>2</sub> peak broadens and the maximum temperature position shifts from 40 °C to 80 °C after the 140 °C treatment (Fig. 3, top panel, blue curve). In contrast to MD<sub>1</sub>, the intensity of the dielectric loss  $\epsilon''$  of MD<sub>2</sub> only slightly drops when the protein is heated up to 140 °C but the dielectric loss  $\epsilon''$  is undetectable after the treatment at 180 °C. This suggests as for MD<sub>1</sub> that the toxin conformations associated with MD<sub>2</sub> disappear with increasing temperature above 140 °C.

At 140 °C, simultaneously to the drop of the MD<sub>1</sub> signal, a new MD (MD<sub>3</sub>) appears with a narrow peak and a maximum temperature position at 10 °C (Fig. 3, upper panel). The MD<sub>3</sub> peak broadens after treatment at 180 °C and the maximum temperature position shifts to 70 °C.

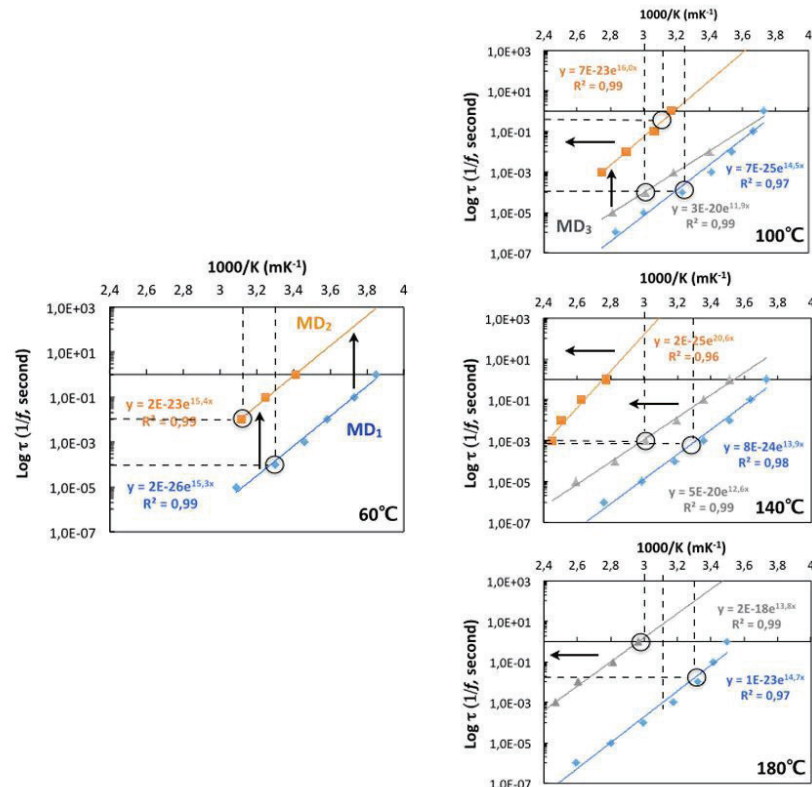
MD<sub>1</sub> signal drops with increasing temperatures but with no evident increase of the signals of MD<sub>2</sub> and MD<sub>3</sub>, compensating for the loss of the MD<sub>1</sub> signal. This indicates that the protein unfolds into conformations, others than MD<sub>2</sub> and MD<sub>3</sub>, and not detected at frequencies between 1 Hz to  $10^6$  Hz. This would suggest that unfolded states of proteins where most 'native' atomic and molecular interactions and initial dipoles are lost, might need to be investigated at lower frequencies to be detected.

To assess whether the protein initial conformation associated with MD<sub>1</sub> could be recovered after the thermal treatment at 180 °C and the cooling down to  $-80$  °C, the sample was heated back to 60 °C and kept at 60 °C for three hours (Fig. 3, purple curve). The dielectric losses measured after the treatment at 180 °C and after the subsequent three hours at 60 °C were similar, indicating that the thermal damages are irreversible and the protein does not refold into the initial conformation despite three hours at 60 °C (Fig. 3, green, purple and red curves, respectively). Finally, the sample was heated back to 180 °C and kept for three more hours and the dielectric loss measured again (Fig. 3, grey curve). The signal appears lower than after the first treatment at 180 °C and MD<sub>1</sub> is almost undetectable indicating that the associated averaged conformations is almost gone. After the second treatment at 180 °C, the MD<sub>2</sub> and MD<sub>3</sub> peaks seem merged.

At  $10^3$  Hz, MD<sub>2</sub> is barely detected while the MD<sub>1</sub> temperature-dependent features change similarly to 1 Hz, and the MD<sub>3</sub> can already be detected at 100 °C (Fig. 3, middle panel). At  $10^6$  Hz, only MD<sub>1</sub> is detectable, with a drop of the dielectric loss, a shift in the maximum temperature position with the increase of the temperature of the thermal treatments as observed for lower frequencies (Fig. 3, bottom panel). The signal is noisier at temperatures below  $-20$  °C but at  $10^6$  Hz, it becomes noisier almost over the whole temperature range.

The frequency dependency of the three relaxation processes was investigated as function of temperature (Fig. 4). MD<sub>1</sub> covers both low and high frequencies from 1 Hz to  $10^6$  Hz over the thermal treatments from 60 °C to 180 °C. The MD<sub>2</sub> is detected only at low frequencies from 1 to  $10^3$  Hz over the thermal treatments from 60 °C





**Figure 5.** Temperature dependences of the relaxation times for the three relaxation processes identified for the Cholera toxin B pentamer after thermal treatments at 60 °C (Top left), 100 °C (Top right), 140 °C (Bottom left) and 180 °C (bottom right). Dipoles changing frequencies across relaxation processes and with the increase of the temperature of the thermal treatments are highlighted with arrows and circle, respectively.

to 140 °C. MD<sub>3</sub> covers high frequencies after the thermal treatment at 100 °C, both high and low frequencies after the thermal treatment at 140 °C, and only low frequencies after the thermal treatment at 180 °C.

The temperature dependencies of the relaxation times  $\tau$  ( $\tau = 1/f$ ) of the three relaxation processes are plotted on Fig. 5. The relaxation times linearly depend on the maximum temperature positions indicating that dipoles fluctuating at different frequencies are independent of one another and that is true for the three relaxation processes. Three signal features change with the temperature of the thermal treatments and with the relaxation processes.

First, the relaxation times of dipoles with similar maximum temperature positions become significantly slower in MD<sub>3</sub> and in MD<sub>2</sub> as the temperatures of the thermal treatment increase (Fig. 5, circles). In contrast, they remain similar up to 140 °C in MD<sub>1</sub> and become slower only after the treatment at 180 °C (Fig. 5, circles). The relaxation times of dipoles with similar maximum temperature positions also become slower across the MDs (Fig. 5, vertical arrows). This means thermal unfolding leads to motions at larger spatial scale with no impact on the dipole local environment (no change in the maximum temperature positions). It could be interface motions or beta sheet motion for example. Moreover, regardless the maximum temperature positions, the relaxation times of MD<sub>3</sub> are ten times slower at 140 °C than at 100 °C (Fig. 5, circles). Likewise, the relaxation times in MD<sub>3</sub> are ten times slower than in MD<sub>1</sub> at 100 °C and hundred times slower at 140 °C, regardless the maximum temperature positions. The relaxation times of dipoles in MD<sub>2</sub> are thousand times slower than in MD<sub>1</sub>, regardless the maximum temperature positions. This indicates that dipoles fluctuating at different frequencies are independent of one another even as toxin thermal unfolding occurs. Dipole fluctuation influences as the toxin unfolds might happen during the incubation time and thus not be detected on steady state conformations.

Second, the maximum temperature positions of dipoles with similar relaxation times increase (1000/K decrease) with the increase of the temperature of the thermal treatments (Fig. 5, horizontal arrows). This is also

observed across relaxation processes as the maximum temperature positions of dipoles with similar relaxation times increase from MD<sub>1</sub> to MD<sub>3</sub> to MD<sub>2</sub>. This means thermal unfolding also leads to changes in the vicinity of the dipoles (changes in the maximum temperature positions) with no impact on the dipole relaxation time (no change in  $\tau$ ). This could be a dipole going from a buried position to a surface position but keeping same fluctuation.

Third, dipoles change in terms of both relaxation times and maximum temperature positions. For example, there are no dipoles with slow relaxations time in MD<sub>3</sub> at 100 °C but there are some at 140 °C (Fig. 5). This means the MD<sub>3</sub> dipoles with fast relaxation times at 100 °C have slowed down and changed their maximum temperature positions at 140 °C. Here the thermal unfolding leads to local changes and larger scale collective motions.

The relaxation time temperature dependence of the three relaxation processes fits well with an Arrhenius equation (Methods) over the four thermal treatments at 60 °C, 100 °C, 140 °C and 180 °C (Fig. 5). They present similar activation energies of  $\approx 27\text{--}30\text{ kcal mol}^{-1}$  across the three MD and for the different temperatures of the thermal treatments, except for the MD<sub>2</sub> at 140 °C whose activation energy is  $\approx 41\text{ kcal mol}^{-1}$  (Fig. 5, slopes). Such similar activation energies across the MD and the temperatures of the thermal treatments are consistent with having similar probes monitored in all processes but in different molecular environment as the toxin unfolds locally (around the dipole) or at larger scale.

### Discussion

The main narrow peak MD<sub>1</sub> observed at 60 °C suggests that the nanopores work as a conformational funnel, and reduce the structural heterogeneity of the toxin in solution by constraining the 3D space where the toxin can move (nanopore diameter). We are currently testing other nanopore dimensions to investigate this hypothesis. In addition to detecting MD<sub>1</sub>, the nanoconfinement combined with BDS also distinguish, via their respective molecular dynamics (MD<sub>2</sub>, and MD<sub>3</sub>) two other average toxin conformations induced by the thermal treatments.

The toxin conformations and dynamics measured by BDS combined with nanoconfinement are clearly different from the toxin dynamics measured under macroscopic and hydrated conditions as illustrated by the toxin resistance to thermal treatments up to 180 °C instead of 80 °C<sup>31</sup>. Such impact on the toxin stability is consistent with the detection of disassembled and/or unfolding species, rarely observed otherwise, made more stable under the conditions. The nanoconfinement, the dehydration or a combination of both could be responsible for the high temperature-resistance. Dehydration conditions are known to impact protein dynamics<sup>32–34</sup>. Hydrogen bonding might be reduced under dehydration while the presence of the PBS buffer could favor salt bridges and lead altogether to a more stable toxin. This would be consistent with the role of stabilizing salt bridges on temperature resistance observed in thermophilic proteins<sup>35–37</sup>. BDS experiments on myoglobin and other proteins under macroscopic hydrated conditions report one fast relaxation process corresponding to free water dynamics which disappears upon dehydration, and slower relaxation processes associated with protein-water and protein dynamics still detected upon dehydration<sup>43,45,58</sup>. Schiro and co-authors have also shown using BDS that myoglobin confined in silica gels presents only slow relaxation processes as observed in our experiments, and little dynamic changes with the dehydration level of the sample<sup>44</sup>. This would tend to suggest that with the nanoconfinement, the dehydration impacts less the protein dynamics. The nanoconfinement could increase the packing of the toxin structure and thus stabilize the protein as observed for thermophilic proteins<sup>37</sup>. Repeating the experiments using different nanoconfinements and different dehydration conditions will be one way to test the role of both parameters in the temperature-resistance of the toxin.

Assessing nanoconfined protein dynamics is of interest as in cells geometrical nanoconfinement of proteins, for example in lipid microdomains (e.g. RAFT), regulate functional protein dynamics such as pore-formation, protein sorting and protein function<sup>31,59,60</sup>.

Protein dynamical structures detected under unusual experimental conditions (nanoconfinement and dehydration) show the robustness of protein structure and protein dynamics to a large range of conditions, going from crystal to dehydrated state in agreement with the sustainability of proteins in time and environmental conditions<sup>61–63</sup>. This is thanks to structural and dynamics diversity which also fits with the function of a protein provided by an ensemble of structures and dynamics rather than unique solutions<sup>37,64</sup>.

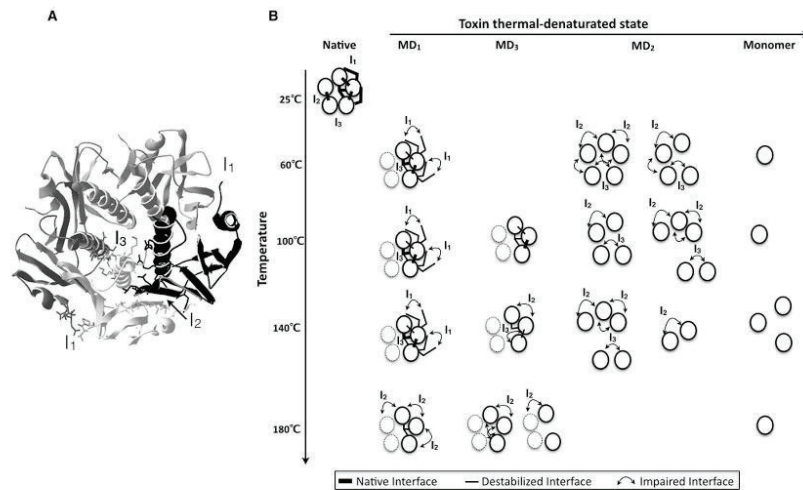
Importantly, the dielectric signal is protein conformation specific, evolves with the temperature of the thermal-treatments and is therefore sensitive to the perturbation of local atomic and molecular interactions. Thus the technique allows characterizing a protein in terms of conformation, stability (to temperature) and dynamics with multiple measures such as frequencies and maximum temperature positions. The technique could therefore be used to detect mutations by comparing the dielectric signals of a protein and its sequence-variants and diagnose the impact of mutations. Such experimental measurements would be useful to classify mutations and develop customized therapies accordingly.

Nanoconfinement allows us to investigate the dynamics of proteins at length scales comparable to the protein geometrical dimensions, which is expected to enhance the sensitivity of dielectric measurements and help detecting the impact of protein mutation. At the macroscopic scale, this ability is strongly limited by disorder-effects arising from a very broad distribution of possible protein conformations.

Presently, only the global dielectric signal can be used to distinguish molecular dynamics and conformations. It is impossible to determine the local dipoles responsible for the global signal or the scale of the collective fluctuations probed. In other words, whether the entire protein is probed or only sub-domains, and if so which one cannot be inferred using only the BDS data. This is because the sample is a mixture of three compounds, the protein, the PBS and protein bound-water. Each compound has many dipoles and the individual participation of each compound or how the individual signals combined to give the global signal is intractable.

What can be done instead is to challenge the data against known experimental and theoretical data on protein dynamics to draw hypotheses on the scale of the collective fluctuations detected by BDS. Putting the results into such perspective, we propose a model of the mechanism of the toxin thermal unfolding with the supplementary





**Figure 6.** Model of the thermal denaturation of the cholera toxin B pentamer. **(A)** Interfaces in CtxB<sub>5</sub>. Ribbon representation of the atomic structure of CtxB<sub>5</sub> with the three interface areas in sticks. Interface 1 (I<sub>1</sub>) where residues 92–93 on chain M interact with residues 1–4 on chain M – 1. Interface 2 (I<sub>2</sub>), the main toxin interface, where residues 23 to 31 on chain M interact with residues 96 to 103 on chain M + 1. Interface 3 (I<sub>3</sub>) where residues 63 to 74 of chains M – 1, M and M + 1 interact with one another. **(B)** The CtxB monomers are circles and the pentamer is composed of five circles, with two in dashed lines for the sake of clarity. The native interfaces and the monomers are not detected during the measurement. Only destabilized (fast motions) and impaired (slow motions) interfaces are detected. MD<sub>1</sub> is associated with a pentamer having both fast and slow motions at all temperatures while MD<sub>3</sub> pentamer has only fast motions at 100 °C, both at 140 °C and only slow at 180 °C. MD<sub>2</sub> is associated with CtxB assembly intermediates that contain CtxB monomers with impaired interfaces and no interface.

constraint that the model agrees with previously established mechanisms of the toxin unfolding and refolding<sup>46,51,65,66</sup> (Fig. 6).

Amino acid side-chain motions determined by other experimental approaches and molecular dynamic simulations are several decades ( $\leq$ nanoseconds) faster than the range of frequencies measured by BDS so it is reasonable to presume that such motions are not detected here.

According to molecular dynamics simulations and integrative approaches, microsecond to millisecond motions are attributed to loop, 2D and 3D structural elements whereas millisecond to second motions are folding and interdomain motions<sup>1,3</sup>. Microsecond to second dipole fluctuations on other proteins have been observed previously by BDS and attributed to protein backbone, large-scale domain or buried side chain motions<sup>44,58</sup>.

Because MD<sub>1</sub> is the main peak detected from 60 °C to 140 °C, we will assume it corresponds to the most stable state of the toxin and its most folded average conformation, a 'native-like' pentamer. Because MD<sub>3</sub> as MD<sub>1</sub> has a narrow peak, and is similar in terms of frequency ranges but with slower relaxation times, we will assume it is a partially folded pentamer where the dipole have less compact environment allowing larger scale motions. Because MD<sub>2</sub> is the slowest relaxation process and a broad peak, we will assume it to be the least folded CtxB species detected and a mixture of CtxB assembly intermediates (CtxB tetramers, trimers and dimers). Alternatively, MD<sub>1</sub> could be a folded CtxB monomer, MD<sub>3</sub> a slightly less folded monomer and MD<sub>2</sub> multiple unfolded states of CtxB monomers. But because the three relaxation processes disappear at high temperatures, it is difficult to picture what would be the undetected species if unfolded monomers are detected. The first scheme is also consistent with the CtxB (dis)assembly and (un)folding intermediates previously identified macroscopically<sup>46,65</sup>. The activation energy of Proline trans-cis isomerization is 30 kcalmol<sup>-1</sup>, and maybe the toxin relaxation processes and dipole fluctuations are under the influence of such reaction. As the toxin assembly is inhibited by the isomerization of Pro93 which slows down the formation of the main  $\beta$ -interface, it is possible although speculative<sup>46</sup>.

Under the first scheme, it is reasonable to assume that the toxin subunit interfaces are monitored during an experiment since CtxB assembly intermediates are detected (MD<sub>2</sub>). It follows that folded and unfolded CtxB monomers would be among the undetected species. The native toxin pentamer has three distinct interface areas I<sub>1</sub>, I<sub>2</sub> and I<sub>3</sub> as shown in Fig. 6A, but native interfaces are unlikely to be detected because they involve local atomic motions with fluctuations over the nanosecond scale<sup>3</sup>. In MD<sub>1</sub>, fast and slow relaxation times are detected and as the toxin unfolds with high temperature treatments only slow relaxation times are detected for MD<sub>2</sub> and MD<sub>3</sub>. The toxin interfaces I<sub>1</sub>, I<sub>2</sub> and I<sub>3</sub> would exist as native (nanosecond scale not detected), destabilized (microsecond to millisecond fluctuations: 2D/3D segment mobility) and impaired with larger and slower motions (millisecond

to second: domain motion) (Fig. 6B). These assumptions are consistent with the microsecond to millisecond time-scale motions, measured by hydrogen exchange, involved in protein thermal unfolding<sup>67</sup>. The model assumes that CtxB intermediates are detected through the slow motion interfaces, the interfaces with no adjacent chains are assumed undetected. Having  $I_1$ ,  $I_2$  and  $I_3$  interfaces with slow motions in a monomer, is assumed to trigger the monomer dissociation extremely fast.

Now, in MD<sub>1</sub> even at 180 °C fast relaxation times (microsecond to millisecond) are observed but for different maximum temperature positions than at 140 °C. This suggests that some undetected native interfaces present in MD<sub>1</sub> at 140 °C, become destabilized and detected at 180 °C but at different maximum positions than the one detected at 140 °C because they are different interfaces (Fig. 6). Alternatively, all the interfaces in MD<sub>1</sub> could be destabilized and detected with fast motions from 60 °C, and the change in maximum temperature positions with the thermal treatments could indicate local environmental changes around the dipoles with fast relaxation times due to local unfolding with no consequences at larger scales. This would have to be further investigated.

The model highlights the large combinatory of interface dynamic states within a single monomer, the CtxB assembly intermediates and the pentamers. The dielectric loss measurement is performed on hundreds of nanopores, each containing some toxin molecules so it is capturing a large statistics of structural states, referred to as average conformations and appearing within each molecular dynamics.

## Conclusion

The molecular dynamics of distinct protein conformations, among which conformations rarely detected macroscopically, become accessible experimentally by combining nanoconfinement with broadband dielectric spectroscopy. This novel biophysical approach captures the dipole fluctuations underlying the molecular dynamics of average protein conformations whose slow motions associate with protein unfolding. The level of description provided by the BDS offers a large experimental potential to study protein dynamics and is an experimental equivalent to molecular dynamics simulation over the frequency range from  $10^3$  Hz to  $10^6$  Hz.

Received: 5 August 2019; Accepted: 5 November 2019;

Published online: 29 November 2019

## References

- Lindorff-Larsen, K., Piana, S., Dror, R. O. & Shaw, D. E. How fast-folding proteins fold. *Science* **334**, 517–520 (2011).
- Vuillon, L. & Lesieur, C. From local to global changes in proteins: a network view. *Current opinion in structural biology* **31**, 1–8 (2015).
- Munoz, V. & Cerminara, M. When fast is better: protein folding fundamentals and mechanisms from ultrafast approaches. *Biochemical Journal* **473**, 2545–2559 (2016).
- Barends, T. R. *et al.* Direct observation of ultrafast collective motions in CO myoglobin upon ligand dissociation. *Science* **350**, 445–450 (2015).
- Levantino, M. *et al.* Ultrafast myoglobin structural dynamics observed with an X-ray free-electron laser. *Nature communications* **6**, 6772 (2015).
- Mizohata, E., Nakane, T., Fukuda, Y., Nango, E. & Iwata, S. Serial femtosecond crystallography at the SACLA: breakthrough to dynamic structural biology. *Biophysical reviews* **10**, 209–218 (2018).
- Stagno, J. R. *et al.* Structures of riboswitch RNA reaction states by mix-and-inject XFEL serial crystallography. *Nature* **541**, 242 (2017).
- Herbst, J., Heyne, K. & Diller, R. Femtosecond infrared spectroscopy of bacteriorhodopsin chromophore isomerization. *Science* **297**, 822–825 (2002).
- Kolano, C., Helbing, J., Kozinski, M., Sander, W. & Hamm, P. Watching hydrogen-bond dynamics in a  $\beta$ -turn by transient two-dimensional infrared spectroscopy. *Nature* **444**, 469 (2006).
- Mizutani, Y. & Kitagawa, T. Ultrafast dynamics of myoglobin probed by time-resolved resonance Raman spectroscopy. *The chemical record* **1**, 258–275 (2001).
- Fang, C., Frontiera, R. R., Tran, R. & Mathies, R. A. Mapping GFP structure evolution during proton transfer with femtosecond Raman spectroscopy. *Nature* **462**, 200 (2009).
- Schneider, R., Blackledge, M. & Jensen, M. R. Elucidating binding mechanisms and dynamics of intrinsically disordered protein complexes using NMR spectroscopy. *Current opinion in structural biology* **54**, 10–18 (2019).
- Findsen, E. W., Scott, T. W., Chance, M. R., Friedman, J. M. & Ondrias, M. R. Picosecond time-resolved Raman studies of photodissociated carboxymyoglobin. *Journal of the American Chemical Society* **107**, 3355–3357 (1985).
- Heath, G. R. & Scheuring, S. High-speed AFM height spectroscopy reveals  $\mu$ s-dynamics of unlabeled biomolecules. *Nature communications* **9**, 4983 (2018).
- Hellenkamp, B., Wortmann, P., Kandzia, F., Zacharias, M. & Hugel, T. Multidomain structure and correlated dynamics determined by self-consistent FRET networks. *Nature methods* **14**, 174 (2017).
- Soranno, A. *et al.* Integrated view of internal friction in unfolded proteins from single-molecule FRET, contact quenching, theory, and simulations. *Proceedings of the National Academy of Sciences* **114**, E1833–E1839 (2017).
- Lesieur, C. & Schulten, K. Editorial overview: Theory and simulation. *Current opinion in structural biology* **31**, v (2015).
- Perilla, J. R. *et al.* Molecular dynamics simulations of large macromolecular complexes. *Current opinion in structural biology* **31**, 64–74 (2015).
- Leitner, D. M. & Yamato, T. Mapping energy transport networks in proteins. *arXiv preprint arXiv:1805.03715* (2018).
- Liu, J. & Nussinov, R. Allostery: an overview of its history, concepts, methods, and applications. *PLoS computational biology* **12**, e1004966 (2016).
- Naganathan, A. N. Modulation of allosteric coupling by mutations: from protein dynamics and packing to altered native ensembles and function. *Current opinion in structural biology* **54**, 1–9 (2019).
- Di Paola, L. & Giuliani, A. Protein contact network topology: a natural language for allostery. *Current opinion in structural biology* **31**, 43–48 (2015).
- Li, C.-B., Yang, H. & Komatsuzaki, T. Multiscale complex network of protein conformational fluctuations in single-molecule time series. *Proceedings of the National Academy of Sciences* **105**, 536–541 (2008).
- Achoch, M. *et al.* Protein structural robustness to mutations: an in silico investigation. *Physical Chemistry Chemical Physics* **18**, 13770–13780 (2016).
- Gheeraert, A. *et al.* Exploring Allosteric Pathways of a V-Type Enzyme with Dynamical Perturbation Networks. *The Journal of Physical Chemistry B* (2019).



26. Degiacomi, M. T. *et al.* Molecular assembly of the aerolysin pore reveals a swirling membrane-insertion mechanism. *Nature chemical biology* **9**, 623–629 (2013).
27. Degiacomi, M. T. Coupling Molecular Dynamics and Deep Learning to Mine Protein Conformational Space. *Structure* (2019).
28. Marklund, E. G. & Benesch, J. L. Weighing-up protein dynamics: the combination of native mass spectrometry and molecular dynamics simulations. *Current opinion in structural biology* **54**, 50–58 (2019).
29. Boyd, C. M. & Bubeck, D. Advances in cryoEM and its impact on  $\beta$ -pore forming proteins. *Current opinion in structural biology* **52**, 41–49 (2018).
30. Pak, A. J. & Voth, G. A. Advances in coarse-grained modeling of macromolecular complexes. *Current opinion in structural biology* **52**, 119–126 (2018).
31. Parsons, E. S. *et al.* Single-molecule kinetics of pore assembly by the membrane attack complex. *Nature communications* **10**, 2066 (2019).
32. Bedem, V. D. H. & Fraser, J. S. Integrative, dynamic structural biology at atomic resolution—it's about time. *Nature methods* **12**, 307 (2015).
33. Demir, O. *et al.* Ensemble-based computational approach discriminates functional activity of p53 cancer and rescue mutants. *PLoS computational biology* **7**, e1002238 (2011).
34. Vendruscolo, M. Determination of conformationally heterogeneous states of proteins. *Current opinion in structural biology* **17**, 15–20 (2007).
35. Toth-Petroczy, A. & Tawfik, D. S. The robustness and innovability of protein folds. *Current opinion in structural biology* **26C**, 131–138 (2014).
36. Payne, J. L. & Wagner, A. The robustness and evolvability of transcription factor binding sites. *Science* **343**, 875–7 (2014).
37. Parisi, G., Zea, D. J., Monzon, A. M. & Marino-Buslje, C. Conformational diversity and the emergence of sequence signatures during evolution. *Current opinion in structural biology* **32C**, 58–65 (2015).
38. McLaughlin, R. N. Jr., Poelwijk, F. J., Raman, A., Gosal, W. S. & Ranganathan, R. The spatial architecture of protein function and adaptation. *Nature* **491**, 138–42 (2012).
39. Salinas, V. H. & Ranganathan, R. Coevolution-based inference of amino acid interactions underlying protein function. *eLife* **7**, e34300 (2018).
40. Lindorff-Larsen, K., Best, R. B., DePristo, M. A., Dobson, C. M. & Vendruscolo, M. Simultaneous determination of protein structure and dynamics. *Nature* **433**, 128 (2005).
41. Houachia, A. *et al.* Nanofluidics approach to separate between static and kinetic nanoconfinement effects on the crystallization of polymers. *Nano letters* **15**, 4311–4316 (2015).
42. Serghi, A., Zhao, W., Miranda, D. & Russell, T. P. Curie transitions for attograms of ferroelectric polymers. *Nano letters* **13**, 577–80 (2013).
43. Jansson, H., Bergman, R. & Swenson, J. Relation between solvent and protein dynamics as studied by dielectric spectroscopy. *The Journal of Physical Chemistry B* **109**, 24134–24141 (2005).
44. Schiro, G., Cupane, A., Vitrano, E. & Bruni, F. Dielectric relaxations in confined hydrated myoglobin. *The Journal of Physical Chemistry B* **113**, 9606–9613 (2009).
45. Jansson, H., Bergman, R. & Swenson, J. Role of solvent for the dynamics and the glass transition of proteins. *The Journal of Physical Chemistry B* **115**, 4099–4109 (2011).
46. Lesieur, C. *et al.* A kinetic model of intermediate formation during assembly of cholera toxin B-subunit pentamers. *J Biol Chem* **277**, 16697–704 (2002).
47. Razvi, A. & Scholtz, J. M. Lessons in stability from thermophilic proteins. *Protein Science* **15**, 1569–1578 (2006).
48. Bowler, B. E. Residual structure in unfolded proteins. *Current opinion in structural biology* **22**, 4–13 (2012).
49. Karshikoff, A., Nilsson, L. & Ladenstein, R. Rigidity versus flexibility: the dilemma of understanding protein thermal stability. *The FEBS journal* **282**, 3899–3917 (2015).
50. Dalziel, A. W., Lipka, G., Chowdhry, B. Z., Sturtevant, J. M. & Schafer, D. E. Effects of ganglioside G M1 on the thermotropic behavior of cholera toxin B subunit. *Molecular and cellular biochemistry* **63**, 83–91 (1984).
51. Goins, B. & Freire, E. Thermal stability and intersubunit interactions of cholera toxin in solution and in association with its cell-surface receptor ganglioside GM1. *Biochemistry* **27**, 2046–2052 (1988).
52. Prestrelski, S. J., Tedeschi, N., Arakawa, T. & Carpenter, J. F. Dehydration-induced conformational transitions in proteins and their inhibition by stabilizers. *Biophysical journal* **65**, 661–671 (1993).
53. Frauenfelder, H. *et al.* A unified model of protein dynamics. *Proceedings of the National Academy of Sciences* **106**, 5129–5134 (2009).
54. Iavarone, A. T., Patriksson, A., van der Spoel, D. & Parks, J. H. Fluorescence probe of Trp-cage protein conformation in solution and in gas phase. *Journal of the American Chemical Society* **129**, 6726–6735 (2007).
55. Dominy, B. N., Minoux, H. & Brooks, C. L. III. An electrostatic basis for the stability of thermophilic proteins. *Proteins: Structure, Function, and Bioinformatics* **57**, 128–141 (2004).
56. Xia, Y.-L. *et al.* Insights into the role of electrostatics in temperature adaptation: a comparative study of psychrophilic, mesophilic, and thermophilic subtilisin-like serine proteases. *RSC advances* **8**, 29698–29713 (2018).
57. Kumar, S., Ma, B., Tsai, C.-J. & Nussinov, R. Electrostatic strengths of salt bridges in thermophilic and mesophilic glutamate dehydrogenase monomers. *Proteins: Structure, Function, and Bioinformatics* **38**, 368–383 (2000).
58. Khodadadi, S., Pawlus, S. & Sokolow, A. P. Influence of hydration on protein dynamics: combining dielectric and neutron scattering spectroscopy data. *The Journal of Physical Chemistry B* **112**, 14273–14280 (2008).
59. Fivaz, M. *et al.* Differential sorting and fate of endocytosed GPI-anchored proteins. *Embo J* **21**, 3989–4000 (2002).
60. Lafourcade, C., Sobo, K., Kieffer-Jaquinod, S., Garin, J. & van der Goot, F. G. Regulation of the V-ATPase along the endocytic pathway occurs through reversible subunit association and membrane localization. *PLoS One* **3**, e2758 (2008).
61. Rackham, O. J. *et al.* The evolution and structure prediction of coiled coils across all genomes. *Journal of molecular biology* **403**, 480–93 (2010).
62. Silveira, J. A. *et al.* From solution to the gas phase: stepwise dehydration and kinetic trapping of Substance P reveals the origin of peptide conformations. *Journal of the American Chemical Society* **135**, 19147–19153 (2013).
63. Hekstra, D. R. *et al.* Electric-field-stimulated protein mechanics. *Nature* **540**, 400 (2016).
64. Kondrashov, D. A., Zhang, W., Aranda, R. IV, Stec, B. & Phillips, G. N. Jr. Sampling of the native conformational ensemble of myoglobin via structures in different crystalline environments. *Proteins: Structure, Function, and Bioinformatics* **70**, 353–362 (2008).
65. Zrimi, J., Ng Ling, A., Giri-Rachman Arifin, E., Feverati, G. & Lesieur, C. Cholera toxin B subunits assemble into pentamers - proposition of a fly-casting mechanism. *PLoS One* **5**, e15347 (2010).
66. Bhakuni, V., Xie, D. & Freire, E. Thermodynamic identification of stable folding intermediates in the B-subunit of cholera toxin. *Biochemistry* **30**, 5055–60 (1991).
67. Hernández, G., Jenney, F. E., Adams, M. W. & LeMaster, D. M. Millisecond time scale conformational flexibility in a hyperthermophile protein at ambient temperature. *Proceedings of the National Academy of Sciences* **97**, 3166–3170 (2000).



### Acknowledgements

We would like to thank David Leitner and Lorenza Pacini for critical reading of the manuscript and the Interdisciplinary Mission (MITI, Defi 'Instrumentations aux limites' ATTOGRAM) from CNRS for supporting the research.

### Author contributions

L.B. performed research and analyzed data, A.S. designed research, analyzed data and wrote the manuscript, C.L. designed research, analyzed data and wrote the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-019-54562-8>.

**Correspondence** and requests for materials should be addressed to A.S. or C.L.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019

## Analysis of Nanoconfined Protein Dielectric Signals Using Charged Amino Acid Network Models

Lorenza Pacini,<sup>A,B,D</sup> Laetitia Bourgeat,<sup>A,C,D</sup> Anatoli Serghei,<sup>C,E</sup> and Claire Lesieur<sup>A,B,E</sup>

<sup>A</sup>AMPERE, CNRS, University of Lyon, 69622, Lyon, France.

<sup>B</sup>Institut Rhônealpin des Systèmes Complexes, IXXI-ENS-Lyon, 69007, Lyon, France.

<sup>C</sup>IMP, CNRS, University of Lyon, 69622, Lyon, France.

<sup>D</sup>These authors contributed equally to this work.

<sup>E</sup>Corresponding authors. Email: [anatoli.serghei@univ-lyon1.fr](mailto:anatoli.serghei@univ-lyon1.fr); [claire.lesieur@ens-lyon.fr](mailto:claire.lesieur@ens-lyon.fr)

Protein slow motions involving collective molecular fluctuations on the timescale of microseconds to seconds are difficult to measure and not well understood despite being essential to sustain protein folding and protein function. Broadband dielectric spectroscopy (BDS) is one of the most powerful experimental techniques to monitor, over a broad frequency and temperature range, the molecular dynamics of soft matter through the orientational polarisation of permanent dipole moments that are generated by the chemical structure and morphological organisation of matter. Its typical frequency range goes from  $10^7$  Hz down to  $10^{-3}$  Hz, being thus suitable for investigations on slow motions in proteins. Moreover, BDS has the advantage of providing direct experimental access to molecular fluctuations taking place on different length-scales, from local to cooperative dipolar motions. The unfolding of the cholera toxin B pentamer (CtxB<sub>5</sub>) after thermal treatment for 3 h at 80°C is investigated by BDS under nanoconfined and dehydrated conditions. From the X-ray structure of the toxin pentamer, network-based models are used to infer the toxin dipoles present in the native state and to compute their stability and dielectric properties. Network analyses highlight three domains with distinct dielectric and stability properties that support a model where the toxin unfolds into three conformations after the treatment at 80°C. This novel integrative approach offers some perspective into the investigation of the relation between local perturbations (e.g. mutation, thermal treatment) and larger scale protein conformational changes. It might help ranking protein sequence variants according to their respective scale of dynamics perturbations.

Manuscript received: 5 October 2019.

Manuscript accepted: 27 February 2020.

Published online: 19 May 2020.

### Introduction

Protein dynamics covers multiple spatiotemporal scale processes, including slow motions (microsecond to second), which are not much understood because they are difficult to measure, even though they are underlying protein folding and protein functions.<sup>[1–3]</sup>

Fast motions (femtosecond to nanosecond) concern local atomic motions (vibration, amino acid side chain motions) and are measured by ultra-fast techniques such as ultra-fast NMR spectroscopy, femtosecond simulated Raman spectroscopy, ultra fast transient Infra Red spectroscopy, or ultra resolution X-ray crystallography and X-ray laser.<sup>[4–13]</sup> Molecular dynamics (MD) simulations have atomic resolution and can now also cover motions up to a millisecond thanks to the ANTON super computer.<sup>[1]</sup> However, MD simulations at high resolution are limited in terms of size and it remains hard to assess motions slower than the microsecond range.<sup>[14,15]</sup>

Slow motions concern the collective motions of many atoms within areas of large size such as secondary and tertiary structural elements (microsecond to millisecond) up to domains

or chain (millisecond to second).<sup>[3]</sup> The slow fluctuation dynamics is monitored by lower spatial resolution techniques such as atomic force microscopy (AFM) or fluorescence spectroscopy (e.g. Förster resonance energy transfer, FRET), where the detailed MDs are generally lost.<sup>[16–18]</sup> The slow motions are also investigated with network models and integrative approaches combining experimental and multiple theoretical approaches.<sup>[19–30]</sup> They are particularly successful in monitoring the slow motions involved in protein assembly and pore formation.<sup>[28,31–33]</sup>

Broadband dielectric spectroscopy (BDS), although not often used on proteins, can monitor protein MDs in the frequency range from 1 Hz to  $10^6$  Hz where the slow motions from microseconds to seconds involved in protein folding and protein functions take place.<sup>[34–37]</sup> The advantage of BDS is the capacity to characterise dipole features both at the local scale (dipole environment) with the temperature at maximum signal, and at a larger scale through fluctuations of collective motions exhibiting a characteristic relaxation  $\tau$  ( $\tau$ ), defined as the inverse of the frequency. However, as for other biophysics methods such as fluorescence



spectroscopy, the measurement is performed at the macroscopic scale where responses come from the bulk composed of many protein and solvent molecules containing many dipoles.

To overcome some of these limits, the dynamics of the cholera toxin B pentamer (CtxB<sub>5</sub>) was measured by BDS after different thermal treatments, but under nanoconfined and dehydrated conditions.<sup>[37]</sup> Thus, no bulk water contributed to the dielectric signal and attograms (considering a density of  $\sim 1 \text{ g cm}^{-3}$ , 1 attogram is  $\sim 10 \text{ nm} \times 10 \text{ nm} \times 10 \text{ nm}$ ) of matter were probed, approaching the characteristic dimension of one protein.<sup>[37]</sup> The nanoconfinement, initially developed to analyse attograms (1 attogram =  $10^{-18} \text{ g}$ ) and zeptograms (1 zeptogram =  $10^{-21} \text{ g}$ ) of matter, allows one to investigate the dielectric behaviour of matter on a length-scale comparable with the molecular dimensions.<sup>[38,39]</sup> In our study, the BDS signal was found to probe protein conformations that arise from the toxin thermal unfolding.<sup>[37]</sup> Based on knowledge on the scale of protein dynamics and on toxin unfolding mechanisms, a model of the toxin thermal unfolding monitored by BDS was proposed where two non-native pentamers and toxin assembly intermediates were captured.

Nevertheless, because the sample was a mixture of protein, phosphate buffered saline (PBS), and adsorbed water, the global dielectric signal was not de-convoluted in terms of distinct dipolar relaxation.

Here, we use network-based models to infer the protein dipoles present in the native toxin pentamer and network based analyses to show that, consistently with our model, the toxin interfaces are distinguishable dielectrically and thermally.

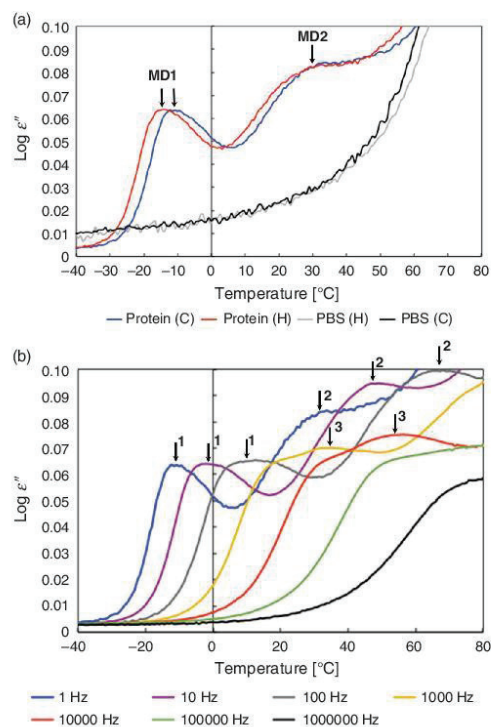
## Results and Discussion

We have recently shown that the dielectric signals of CtxB<sub>5</sub> measured by BDS in nanoconfined and dehydrated conditions evolved with the thermal treatments of the toxin at 60, 100, 140, and 180°C, probing for the toxin thermal unfolding.<sup>[37]</sup> However, because the protein sample was a mixture of three compounds – the protein, phosphate buffer saline (PBS), and protein adsorbed-water – it was impossible to analyse the global dielectric signals in terms of individual dipolar relaxations. In fact, understanding the individual contribution of each compound and how the individual signals combine together to give rise to the global signal is intractable.

Instead, we proposed a model of the thermal toxin unfolding where the global dielectric signal was interpreted based on the knowledge on protein dynamics from MD simulations and from experimental investigations of protein assembly and disassembly under macroscopic conditions where non-native pentamers and CtxB assembly intermediates were identified.<sup>[39–42]</sup> Protein assembly can be described by two mechanisms, the induced fit mechanism where the protein monomer folds before assembling and the fly-casting mechanism where the folding and assembling are concomitant.<sup>[40,43–45]</sup> Protein disassembly involves transitions from native oligomers to either folded monomers or non-native oligomers.<sup>[46–49]</sup> In the model, the three distinct MDs detected by BDS, referred to as MD<sub>1</sub>, MD<sub>2</sub> and MD<sub>3</sub>, were assigned to two different non-native pentamers (MD<sub>1</sub> and MD<sub>3</sub>) and to toxin assembly intermediates (MD<sub>2</sub>), suggesting that the dielectric signal probed the toxin interfaces.<sup>[37]</sup>

Here, we use network-based models to describe the dipoles present at the toxin interface and to test this hypothesis.

As the sample is composed of three compounds, water, salt (PBS), and protein, the first issue is to determine whether the



**Fig. 1.** Dielectric loss as a function of temperature after 3 h of thermal treatment at 80°C. (a) Dielectric loss as a function of temperature for the cholera toxin B pentamer and for the PBS buffer control at 1 Hz. The signals for the heating (H) and cooling (C) temperature ramps are shown. The cooling ramp goes down to  $-80^\circ\text{C}$  but the signal remains flat already below  $-40^\circ\text{C}$ . The relaxation processes corresponding to the molecular dynamics peaks MD<sub>1</sub> and MD<sub>2</sub> are indicated by arrows. (b) Dielectric loss as a function of temperature for the cholera toxin B pentamer at frequencies from 1 to  $10^6$  Hz. The three molecular dynamics peaks (MD<sub>1</sub>, MD<sub>2</sub>, and MD<sub>3</sub>) are observed corresponding to three relaxation processes. The relaxation processes corresponding to the molecular dynamics peaks MD<sub>1</sub>, MD<sub>2</sub>, and MD<sub>3</sub> are indicated by arrows. The MD letters have been removed for clarity.

MDs are the water, PBS, and protein respective individual dielectric signals.

### BDS Combined with Nanoconfinement

First, the dielectric behaviour of the toxin is investigated after three hours incubation at 80°C during which the toxin unfolds. Toxin samples are deposited in the nanomembrane and the dielectric loss  $\epsilon''$  is measured by BDS (see *Experimental*). After thermal treatment for 3 h at 80°C, the dielectric loss  $\epsilon''$  is measured as a function of temperature going from 80 to  $-80^\circ\text{C}$  (cooling) and from  $-80$  back to  $80^\circ\text{C}$  (heating). The dielectric loss  $\epsilon''$  of the samples during the cooling ramp is shown in Fig. 1 for frequencies ranging from 1 to  $10^6$  Hz.

The cooling and heating curves show similar dielectric signals demonstrating that the proteins under study are stabilised in a state where no bulk water is present anymore (i.e. no more water evaporation) (Fig. 1a). In the previous study, it was shown



that the protein sample contains adsorbed water.<sup>[37]</sup> As observed in our previous study, the protein spectra can be described by three relaxation processes, which give rise to three peaks in the spectra of the dielectric loss (Fig. 1b). The PBS buffer control shows only a conductivity contribution under the same conditions indicating that the signal in the protein sample is protein dependent (Fig. 1a).

The absence of MD for the PBS buffer control indicates that the individual dielectric signals of the water, the salt (PBS), and the protein are non-additive and the Havriliak–Negami functions cannot be applied to de-convolute the global signal. The three MDs are not the respective individual dielectric signal of the water, the salt, and the protein but their coupled responses.

The second issue is to determine how many sets of dipoles are associated with the three MDs.

The main relaxation process, referred to as MD<sub>1</sub> (molecular dynamics peak 1), is identified by a narrow peak and it is the fastest, with a maximum dielectric loss  $\epsilon''$  at temperature  $\theta_{\max}$  equal to  $-14^\circ\text{C}$  at 1 Hz. MD<sub>1</sub> is detected for frequencies from 1 to  $10^5$  Hz. The second relaxation process, referred to as MD<sub>2</sub>, has a broader peak and is slower with  $\theta_{\max}$  equal to  $30^\circ\text{C}$  at 1 Hz. The broadness of MD<sub>2</sub> suggests a relaxation process associated with a more heterogeneous population of average toxin conformations compared with the average toxin conformations associated with the narrower MD<sub>1</sub> peak. This leads to the assumption that MD<sub>2</sub> is associated with more thermally perturbed toxin conformations than MD<sub>1</sub> and CtxB assembly intermediates (tetramer, trimer, and dimer). Unfolding leads to loss of atomic interactions such that MD<sub>2</sub>, which is detected only at low frequencies from 1 to  $10^2$  Hz and therefore exhibits fluctuations of larger-scale atomic motions (domain or chain) than the faster frequencies of  $10^3$  Hz and above observed in MD<sub>1</sub>, is in a more unfolded state than MD<sub>1</sub>.<sup>[3]</sup> The third relaxation process, referred to as MD<sub>3</sub>, has a narrow peak as MD<sub>1</sub> but appears only at frequencies higher than 100 Hz indicating the loss of the dipoles detected at low frequencies in MD<sub>1</sub> and so a more unfolded state than the average conformations associated with MD<sub>1</sub>. Because MD<sub>3</sub> has fast motion dipoles and smaller-scale motions than the slow dipoles detected in MD<sub>2</sub>, it is in a less unfolded state than MD<sub>2</sub>. Thus as reported previously for thermal treatments at higher temperatures, three unfolding toxin intermediates can be associated with the BDS signal, two non-native pentamers, MD<sub>1</sub> and MD<sub>3</sub>, and toxin assembly intermediates with MD<sub>2</sub>.<sup>[37]</sup>

For the three MDs, the positions of the maximum dielectric signals ( $\theta_{\max}$ ) shift to higher temperatures with increasing frequencies. The temperature dependencies of the relaxation times tau ( $\tau$ ), which are the inverse of the frequency ( $\tau = 1/f$ ), can be followed by plotting  $\tau$  against the temperature at the maximum signal for the respective frequency (Fig. 2).

The relaxation times follow an Arrhenius linear temperature dependency indicating that dipoles fluctuating at different frequencies are independent from one another (see *Experimental*). This is true for the three relaxation processes. The relaxation times become longer from MD<sub>1</sub> to MD<sub>3</sub> to MD<sub>2</sub> (Fig. 2, vertical arrows).

Longer relaxation times mean dipole fluctuations involving larger collective motions consistently with more advanced stage of the toxin unfolding. The temperatures at maximum signals ( $\theta_{\max}$ ) shift to higher temperatures (lower  $1000/T$ ) for dipoles with similar relaxation times from MD<sub>1</sub> to MD<sub>3</sub> and for MD<sub>1</sub> to MD<sub>2</sub> (Fig. 2, horizontal arrows). This shift means that thermal unfolding also leads to changes in the vicinity of dipoles

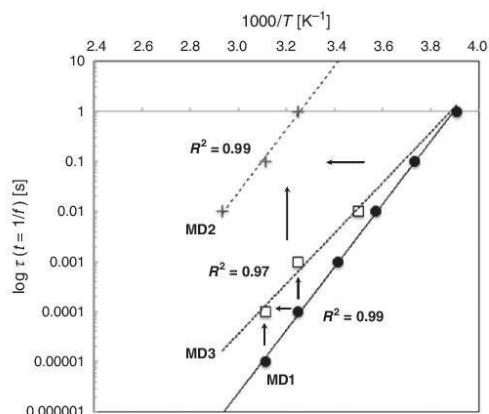


Fig. 2. Temperature dependencies of the relaxation times for the three relaxation processes MD<sub>1</sub>, MD<sub>2</sub>, and MD<sub>3</sub> identified for the cholera toxin B pentamer after thermal treatments at  $80^\circ\text{C}$ . On the x axis the temperature ( $T$ ) is  $\theta_{\max}$ . Dipoles changing frequencies across relaxation processes are highlighted with vertical arrows. Dipoles changing environment but not relaxation times are indicated with horizontal arrows. The data fit with exponential curves that are indicated by straight lines with their respective  $R^2$  values.

(changes in the temperature at maximum signals) with no impact on the dipole relaxation time (no change in tau) or that different sets of dipoles (changes in the temperature at maximum signals) with similar fluctuation scales (no change in tau) are probed.

MD<sub>1</sub> and MD<sub>2</sub> have dipoles with slow motions but not with the same temperatures at maximum signals, indicating that two different sets of dipoles are detected in MD<sub>1</sub> and MD<sub>2</sub>. This is confirmed by the loss of the MD<sub>1</sub> slow fluctuations dipoles in MD<sub>3</sub>, interpreted as an intermediate unfolding stage between MD<sub>1</sub> and MD<sub>2</sub>. The set of dipoles with fast motions in MD<sub>1</sub> is detected as well in MD<sub>3</sub> but with slightly slower motions indicating larger collective motions probably due to further unfolding (Fig. 2, vertical arrows). This set of dipoles is no more detected in the MD<sub>2</sub>, probably because the area unfolded further and the resulting motions fall out of the measured frequency range. Native dipole fluctuations are unlikely to be detected as they fall within time scales equal or below nanoseconds.<sup>[1,3]</sup> According to other experiments and MD simulations, slow motions (millisecond to second) are large domains and chains, while fast motions (microsecond to millisecond) involve smaller areas such as secondary and tertiary structural elements.<sup>[1,3]</sup>

In summary the MD<sub>1</sub> is associated with two sets of dipoles, one with slow fluctuations and the other one with fast fluctuations. The set with fast motions is also detected in MD<sub>3</sub> and a third set of dipoles with slow motions is detected in MD<sub>2</sub>. The dipoles with fast motions in MD<sub>1</sub> and MD<sub>3</sub> could be detected in MD<sub>2</sub> but with slower motions due to a more advanced toxin unfolding. However, because the motions in MD<sub>2</sub> are a thousand times slower it is unlikely that such larger unfolding leads to no change in the local environment of the dipoles (no change in the temperature at maximum signals). It is more reasonable to consider that MD<sub>2</sub> is associated with a third set of dipoles which has similar temperatures at maximum signals than the set of dipoles with fast fluctuations in MD<sub>1</sub> and MD<sub>3</sub>. Thus, two sets of dipoles with slow motions are detected and one set with fast motions.



### Network Based Models

The BDS data indicate that three sets of dipolar relaxations are detected and the model of the toxin thermal unfolding assumes that MD<sub>1</sub> and MD<sub>3</sub> are non-native pentamers while MD<sub>2</sub> is associated with assembly intermediates, implying that the toxin interfaces are probed by the BDS signal.

Hence, for the model to be valid, the toxin interfaces need to be dielectrically detectable, thermally sensitive, and be composed of three different sets of dipolar relaxations with distinct thermal sensitivity.

To investigate such a possibility, network-based models are used to study the toxin interfaces in terms of stability and dielectric features. The steps to build the different networks are schematised in Fig. S1 (Supplementary Material).

### Modelling Interactions in the Toxin Interfaces and Ranking Interface Stability

An intermolecular interaction network, referred to as a 4D-network, was built based on the X-ray structure of the toxin native pentamer to describe the toxin interface (see *Experimental*). The nodes of the network are amino acids which belong to different chains and which can be connected because they have at least one pair of atoms within a threshold distance of 5 Å. The links are weighted according to the exact number of pairs of atoms, one per chain that are within a threshold distance of 5 Å. The 4D-network models intermolecular atomic and amino acid interactions, namely the interactions involved in the toxin interface. As an approximation of the interface stability, the number of amino acid and atomic interactions involved in the 4D-network are computed.

The toxin interface is rather complex and made of three physically distinct domains, I<sub>1</sub>, I<sub>2</sub>, and I<sub>3</sub> (Fig. 3).

For simplicity, in Fig. 3, I<sub>1</sub>, I<sub>2</sub>, and I<sub>3</sub> sub-domain interfaces are shown on a trimer of the toxin instead of a pentamer involving the chains D, E, and F. I<sub>1</sub> is the N-terminal sub-domain interface domain of the toxin, it involves residues 1 to 12 of chain E and interacts with two sub-domains on the adjacent chain F. The residues 1 to 3 of chain E interact with residues 92 and 93 on chain F and constitute a weak patch (I<sub>1a</sub>) of the sub-domain interface I<sub>1</sub> with only 5 amino acid pairs and 46 atomic interactions (Table 1). The sub-domain interface I<sub>1b</sub> is made of residues 1 to 12 of chain E and residues 28 to 39 of chain F and constitutes a moderate interface composed of 19 amino acid pairs and 137 atomic interactions (Table 1). I<sub>2</sub> constitutes the main sub-domain β-interface of the toxin and is composed of the β-strand number 3 (residues 25 to 39) and the β-strand number 6 (residues 96 to 103). The residues 25 to 32 of chain E interact with the residues 88 and 96 to 103 on chain D (I<sub>2a</sub>) and constitute a strong interface (I<sub>2a</sub>) with 26 amino acid pairs and 198 atomic interactions (Table 1). The residues 28 to 39 of chain E interact with some of the residues 58 to 68 of the central helix (residues 58 to 81) of chain D, constituting a moderate interface (I<sub>2b</sub>) with 19 amino acid pairs and 136 atomic interactions (Table 1). Finally, there is a weak interface patch (I<sub>2c</sub>) involving interactions between residues 101 to 103 of chain E and residues 73, 76–77 on the adjacent chain F with only 5 amino acid pairs and 22 atomic interactions (Table 1). I<sub>3</sub> is the central α-helix sub-domain interface, composed of residues 58 to 81 of chain E interacting with residues 58 to 81 on the adjacent chains D and F. This is a tri-partite interface as it involves interactions between chain E and F as well as between chain E and D (Fig. 3), but it still constitutes a single interface with two ‘sticky’ moieties. The

central α-helix sub-domain interface I<sub>3</sub> is the strongest interface with 24 amino acid pairs and 214 atomic interactions (Table 1).

Based on the number of atomic and amino acid interactions, the toxin stability of the sub-domain interfaces can be ranked as follows: I<sub>1a</sub> and I<sub>2c</sub> are the weakest, I<sub>1b</sub> and I<sub>2b</sub> are moderate, and I<sub>2a</sub> and I<sub>3</sub> are the strongest (Table 1).

### Dielectric Properties of the Toxin Interface and Ranking

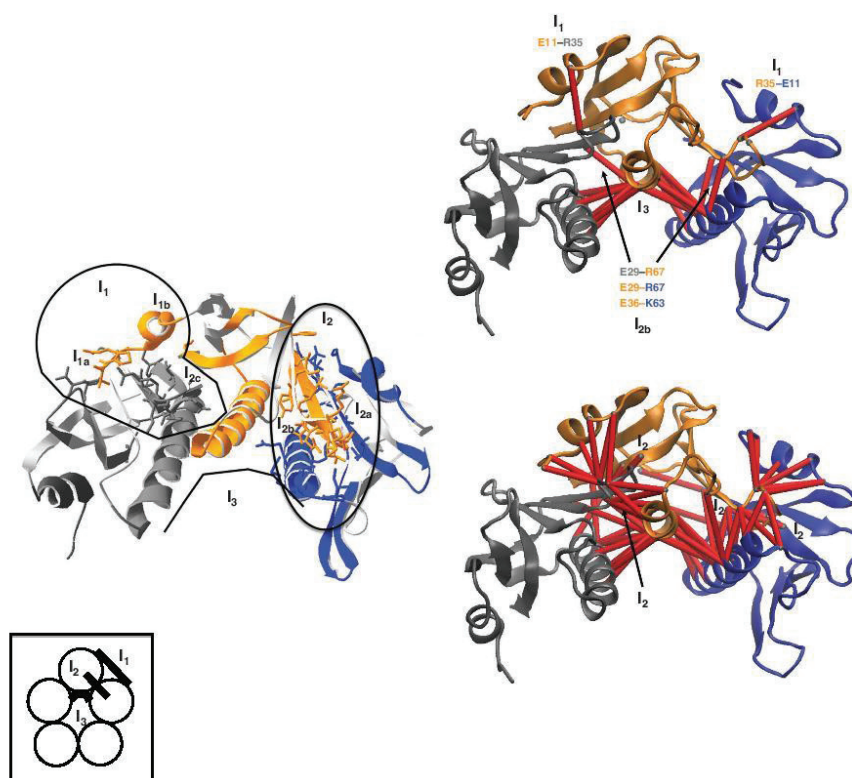
In order to assess the dielectric properties of the toxin interface, an intermolecular charged amino acid network and an intermolecular induced charged amino acid network were built (see *Experimental*). The former is referred to as the 4D-charged network and the latter as the 4D induced charged network. The 4D-charged network is a sub-network of the 4D-network where only charged amino acids are retained (see *Experimental*). The pairs of charged residues are called ionic dipoles. The 4D induced charged network is initially built as the 4D-charged network but the non-charged amino acid neighbours in the 4D-network of the nodes are added. The pairs of non-charged residues are called induced dipoles. The construction of the different networks is schematised in Fig. S1 (Supplementary Material).

The choice of these two types of networks to model the dipoles in the native state of the toxin pentamer and infer dipoles which are detected by BDS is supported by the following reasons. First, protein structures and protein dynamics rely on non-specific hydrophobic interactions (van der Waals) and specific electrostatic interactions (e.g. salt bridges, charged-dipoles), the latter are contributing more to the dielectric signal as they involve Coulomb forces. Moreover, the electrostatic interactions are key to the folding and the stability of a protein and are believed to play a major role in the protein thermal stability.<sup>[43]</sup> Second, charged residues and their proximal neighbours are the major constituents of the electrostatic interactions in proteins.<sup>[50]</sup> Third, the electrostatic interactions in protein structures depend on the pH and the salt conditions, such that considering interactions between charged residues as an approximation of the coupling between protein, water, and salt (PBS) dipoles is relevant.<sup>[50]</sup> Because the pH is 6.9 and based on disassembly and reassembly pH-dependency studies, the N- and C-terminals, lysine and arginine residues are considered protonated while glutamic and aspartic acids are considered deprotonated and hence participating in the 4D-charged network.<sup>[40,46,51–53]</sup>

Based on the 4D-charged network and the 4D-induced charged network, the dielectric properties of the different interface sub-domains are deduced (Table 1) (Fig. 3). I<sub>1a</sub> and I<sub>2c</sub> are dielectrically invisible and weakly visible, respectively, as they have no charged residues but I<sub>2c</sub> has two induced dipoles. I<sub>1b</sub> and I<sub>2b</sub> have two ionic dipoles (pair of charged residues) and some induced dipoles (charged–non charged pairs). I<sub>2a</sub> is invisible dielectrically as it has no charged residues but it can be detected through the induced dipoles in the environment of the charged residue Glu29 (Table 1). I<sub>3</sub> is the strongest dielectric interface domain of the toxin with 13 ionic dipoles and 5 induced dipoles.

### Dielectric and Stability Features of the Toxin Interfaces and BDS Signal Interpretation

Based on the dielectric and stability features of the toxin interface, we can draw the following conclusions on the dipoles that are participating to the BDS signal. The I<sub>1a</sub> and I<sub>2c</sub> domains are the weakest and ranked as the most temperature-sensitive domains but they cannot be detected by BDS or only weakly



**Fig. 3.** The toxin interface. The cholera toxin B pentamer has a complex chain-interface made of mainly three sub-interfaces:  $I_1$ ,  $I_2$ , and  $I_3$ , that are shown on a trimer of the X-ray structure of the toxin for simplicity (left). The chain F is in dark grey, the chain E in orange and the chain D in blue.  $I_1$  is sub-divided into two interfaces:  $I_{1a}$  made of residues 1 to 3 of chain E interacting with residues 92 and 93 on the adjacent chain F, and  $I_{1b}$  made of residues 1 to 12 of chain E interacting with residues 28 to 39 on chain F.  $I_2$  is also sub-divided into three interfaces:  $I_{2a}$  made of residues 25 to 32 of chain E interacting with residues 88 and 96 to 103 on chain D,  $I_{2b}$  made of residues 28 to 39 of chain E interacting with residues 58 to 68 on the adjacent chain D, and  $I_{2c}$  made of residues 101 to 103 of chain E interacting with residues 73 to 77 on the adjacent chain F.  $I_3$  is a tripartite interface composed of residues 58 to 81 of chain M (e.g. E) interacting with residues 58 to 81 on chain M - 1 (D) and M + 1 (F). The electrostatic networks, made using VMD 3D-representations are on the left, with the 4D-charged network at the top and the 4D-induced charged network at the bottom (see *Experimental*). The dipoles are represented by red sticks. A schematic of the three interfaces in the native pentamer are shown in the left side bottom corner.

through induced dipoles (Table 1). The thermal perturbation of  $I_{1a}$  will be detected when  $I_{1b}$  is thermally perturbed as well through the ionic dipoles T1-R35 and E11-R35 (Table 1) (Fig. 4). Likewise, the thermal perturbation of the  $I_{2c}$  domain will be detected via the perturbation of the induced dipoles and the ionic dipoles involving Glu29 when  $I_{2a}$  and  $I_{2b}$  are thermally disturbed (Table 1) (Fig. 4). Because  $I_{1b}$  is a moderate interface in terms of amino acid and atomic interactions compared with  $I_{2a}$ ,  $I_{1b}$  unfolding is likely to occur before the unfolding of  $I_{2a}$  and so  $I_{1b}$  is associated with the slow motion dipoles detected in MD<sub>1</sub> but lost in MD<sub>3</sub>. It follows that  $I_{2a}$  and  $I_{2b}$  are associated with the fast motion dipoles observed in MD<sub>1</sub> and MD<sub>3</sub>. The slow dipoles of MD<sub>2</sub> are associated with the  $I_3$  interface sub-domain as this domain has the strongest stability and the highest number of ionic dipoles and would be detected in the most unfolded stage of the toxin. In MD<sub>2</sub>, the toxin conformations would have lost the  $I_1$  and  $I_2$  sub-domain interfaces, their dipoles being no more

detected by BDS such that the chain dissociation would occur as only an  $I_3$  destabilised interface remains (slow motion = large fluctuation).

The thermal unfolding perturbation and the dipoles detected in MD<sub>1</sub> are schematized in Fig. 4.

The weak interfaces unfold, leading to the perturbation of the  $I_1$  N-terminal interface detected by dipoles with slow motions and to the perturbation of the  $I_2$   $\beta$ -interface detected by dipoles with fast motions. The dipoles detected in MD<sub>3</sub> and MD<sub>2</sub> are shown on the 4D-charged network and the 4D-induced charged network, respectively on Fig. 5. The interface  $I_2$  further unfolds in MD<sub>3</sub> where the dipoles involving the residues Glu29 and Glu36 and the residues Lys63 and Arg67 on the adjacent chains are detected. In MD<sub>2</sub>, the N-terminal  $I_1$  interface and the  $I_2$  interface are unfolded and not detected anymore, while only dipoles from the central helix  $I_3$  interface are still detected.



**Table 1. Interface interaction and dipole features inferred from amino acid network based models**

Thermal strength	Interfaces	Chain E	Chain F	Interface parameter <sup>A</sup>	Chain D	Interface parameter <sup>A</sup>	Ionic dipoles	Induced dipoles
Weak interfaces	I <sub>1a</sub>	1–3	92–93	5, 46, 0, 0			—	—
	I <sub>2c</sub>	101–103	73–77	5, 22, 0, 2			—	M101–R73 N103–Y76
Moderate interfaces	I <sub>1b</sub>	1–12	28–39	19, 137, 2, 3			T1–R35 E11–R35	P2–R35 L8–R35 Y12–R35
	I <sub>2b</sub>	28–39			58–68	19, 136, 2, 9	E29–R67 E36–K63	E29–A64 E29–M68 K34–I58 K34–S60 E36–I58 E36–S60 E36–Q61 E36–A64
Strong interfaces	I <sub>2a</sub>	25–32			88, 96–103	26, 198, 0, 4	—	A97–E29 A98–E29 I99–E29 S100–E29
	I <sub>3</sub> Central $\alpha$ -helix: 63 to 81	63–77			63–81	12, 97, 7, 2	K63–E66	I65–R67
		67–81	66–77	12, 117, 6, 3			E66–K63 E66–R67 R67–R67 K69–K63 K69–R67 D70–R67 R67–R67 R67–E66 R67–K69 R67–D70 R67–R73 D70–R73	Y76–K81 T71–R73 I74–R73 K81–Y76

<sup>A</sup>Parameters characterising interfaces: number of amino acid pairs, number of atomic interactions, number of ionic dipoles (pairs of charged residues), and number of induced dipoles (pairs of charged and non-charged residues).

## Conclusion

This novel network approach to analyse dielectric signals, using network-based models and infer toxin interface dielectric and stability properties from the X-ray structure, supports the hypothesis that BDS probes the thermal perturbation of the toxin interfaces. The toxin interface is composed of three sub-domains that have distinct dielectric and stability features as expected from the model of the toxin thermal unfolding. Moreover, the scale of the motions detected by the BDS is also in good agreement with the unfolding stage of the toxin: the I<sub>1</sub> N-terminal unfolding likely leads to chain motion and the detected slow fluctuations (millisecond to second), in MD<sub>2</sub> only the I<sub>3</sub> sub domain interface is connecting two chains which is likely to lead to chain dissociation and the detected slow fluctuations, the fast motions involving the I<sub>2</sub> interface perturbation is a smaller scale fluctuation involving secondary structure element motions.

Our study is not an exhaustive investigation of the dipoles involved in the BDS signal, and intramolecular contributions are probably also taking place, a possibility that can be explored in the future using charged intramolecular networks instead. Nevertheless, the results allow the pinpointing of some residues potentially important for the toxin unfolding and detected dielectrically, such as the charged pair R35–E11 in the I<sub>1</sub> interface, the charged dipoles E29–(A97, A98, I99, S100) and R73–M101 in the I<sub>2</sub> interface and several charged pairs

K63–E66, R67–D70, and D70–R73 in the central  $\alpha$ -helix. The results also suggest that the I<sub>3</sub> interface is the most stable and most dielectrically detectable interface in the toxin. Thus, based on the results, toxin mutants can be proposed to test the roles of these amino acids both in the dielectric signals and in the toxin thermal unfolding. This integrative approach could also be used to investigate protein sequence variants and rank them according to dynamics perturbations. The network analysis combined with the experimental molecular dynamics measured by BDS in nanoconfined conditions offers a powerful approach to investigate the relation between local perturbations and global protein conformational changes.

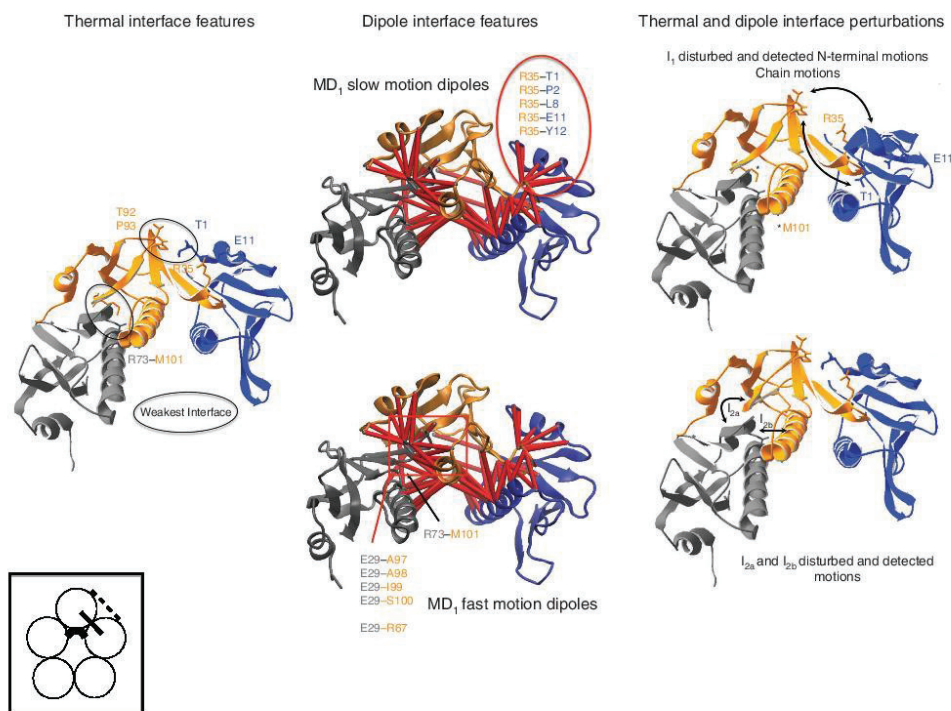
## Experimental

### BDS and Nanoconfinement

#### Materials

Lyophilized cholera toxin B pentamer purchased from Sigma Aldrich (C9903) was diluted in phosphate buffered saline (PBS; 10 mM sodium phosphate, 150 mM sodium chloride, pH 7.4) at a final concentration of 2.6 mg mL<sup>-1</sup>. This bulk toxin preparation was aliquoted and kept at –20°C, and thawed and frozen two times at most, when used for the dielectric measurements. Concentrated PBS solutions were purchased from BIOSOLVE and diluted 10 times in distilled water to prepare the PBS used in the bulk toxin solution. In addition, the PBS solution was filtered





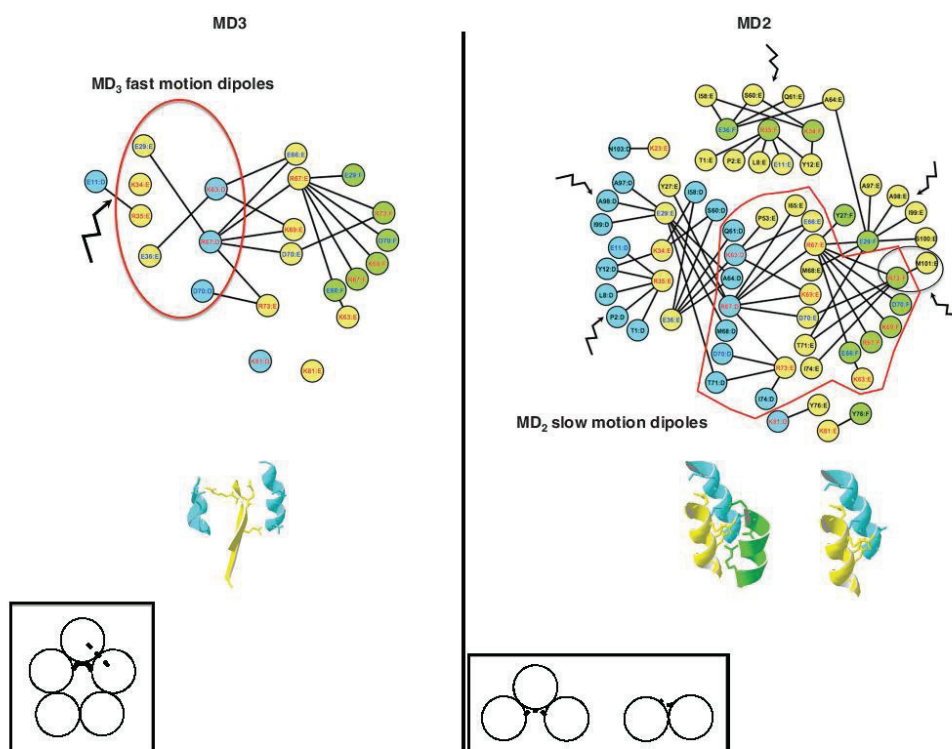
**Fig. 4.** Toxin interface state and dipole detected in MD<sub>1</sub>. MD<sub>1</sub> is associated with the interface state closest to the native pentamer. The weakest interfaces I<sub>2a</sub> and I<sub>2b</sub> are thermally disturbed leading to I<sub>2a</sub> and I<sub>1</sub> interface fluctuations (left and middle structures). I<sub>1b</sub> is weaker than I<sub>2a</sub> so it is associated with the slowest relaxation times of MD<sub>1</sub> (ten milliseconds to second) while I<sub>2a</sub> and I<sub>2b</sub>'s fluctuations are associated with the fastest relaxation times of MD<sub>1</sub> (microsecond to millisecond). The charged dipoles participating in the MD<sub>1</sub> signals are shown on the right structures. The box in the left bottom corner is a schematic of the toxin interface state in MD<sub>1</sub>.

by 22  $\mu\text{m}$  filters to limit impurity contamination. The aluminium oxide (AAO) films on Al were purchased from InRedox. The AAO membrane dimensions were 10 mm  $\times$  10 mm, and they contained nanopores of 40 nm diameter and 10  $\mu\text{m}$  length, with 12 % porosity. The AAO membrane containing the nanopores is referred to as the nanomembrane. Before use, the nanomembranes were plasma-treated for 20 min to remove organic impurities.

#### Methods

**Protein sample preparation.** For the dielectric measurement, a protein sample at 0.025 mg mL<sup>-1</sup> was used. This protein sample was prepared by diluting 2  $\mu\text{L}$  of the bulk toxin solution at 2.6 mg mL<sup>-1</sup> in 204  $\mu\text{L}$  of deionised water (final pH 6.9). Since the bulk protein is prepared in PBS, the sample for the dielectric measurement contains protein, deionised water, and PBS. A 200  $\mu\text{L}$  drop of sample was deposited on the nanomembrane using the drop technique, and then heated at 50°C (323 K) for 15 min to evaporate the bulk water and allow proteins to enter the pores. The sample was cooled to 30°C (303 K) for 5 min with a speed of 2 K min<sup>-1</sup>, placed between two Al electrodes of 7 and 40 mm diameter, respectively, and placed in the cell holder for measurement. The procedure was the same for the control samples.

**Dielectric measurements.** The dielectric measurements were performed on a broadband dielectric spectrometer Novocontrol Alpha analyser over a frequency range from 1 to 10<sup>6</sup> Hz and a temperature range from -80 (193) to 80°C (353 K). The temperature ramps were carried out at a rate of 2 K min<sup>-1</sup> and a voltage of 0.2 V was applied. The dielectric loss  $\epsilon''$ , the imaginary part of the complex dielectric permittivity, was measured as a function of temperature at different frequencies. For the temperature control, a flow of pure nitrogen gas was used in a closed cryostat, providing water-free and oxygen-free experimental conditions. The sample was first heated to 80°C (353 K) for 3 h, and to -80°C (193 K) (cooling, C) and maintained for 30 min before being heated back to 80°C (heating, H). A cycle was composed of the cooling (C) and the heating (H) temperature ramp. The curves of the cooling and heating signals were very similar for the protein sample (Fig. 1a). This demonstrates that the proteins under study were stabilised in a state where no bulk water is present anymore and where the amount of adsorbed water is stabilised as well. It also shows that the protein remains in a steady-state across the measurement and undergoes thermal unfolding mainly during the 3 h incubation time. The PBS buffer control showed no molecular dynamics and only had conductivity after treatment at 80°C (Fig. 1a). The dielectric signals are therefore from the protein.



**Fig. 5.** Toxin interface state and dipole detected in MD<sub>2</sub> and MD<sub>3</sub>. The MD<sub>3</sub> interface state is shown on the left panel. The charged network of the whole interface is shown for a toxin trimer with the nodes (charged residues) of chain D in cyan, of chain E in yellow, and of chain F in green. The residue name and sequence positions are indicated within the nodes, with non-charged residues written in black, positively charged residues in red, and negatively charged residues in blue. MD<sub>3</sub> has lost the slow motion dipoles associated with the N-terminal interface, which is considered damaged and not detected (jagged arrows). Consequently, the I<sub>2<sub>3</sub></sub> interface becomes destabilised and detected (red box) with dipoles E29–R67 and E36–K63 shown as linked on the charged network. The structure of the I<sub>2<sub>3</sub></sub> and I<sub>2<sub>1</sub></sub> interfaces detected in MD<sub>3</sub> are shown below the network. The box in the left bottom corner is a schematic of the toxin interface state in MD<sub>3</sub>. The MD<sub>2</sub> interface state is shown on the right panel. MD<sub>2</sub> is associated with the most destabilised interface state of the toxin. Only the central  $\alpha$ -helix interface I<sub>3</sub> is assumed to be still present. The induced-charged network of the whole interface is shown for a toxin trimer with the nodes (charged residues) of chain D in cyan, of chain E in yellow, and of chain F in green. The damage interfaces are indicated with jagged arrows while the red box indicates the destabilised I<sub>3</sub> interface and the maximum dipoles detected in MD<sub>2</sub>. The structure of the I<sub>3</sub> interfaces detected in MD<sub>2</sub> is shown below the network. The box in the left bottom corner is a schematic of the toxin interface state in MD<sub>2</sub>.

**Dielectric data.** The temperature dependence of the relaxation processes was modelled by an Arrhenius equation expressed as Eqn 1:<sup>[34,35]</sup>

$$\tau = \tau_0 \exp(E_A/k_B T) \quad (1)$$

where  $\tau$  is the relaxation time,  $\tau_0$  the relaxation time at high temperatures,  $T$  the temperature,  $E_A$  the activation energy, and  $k_B$  the Boltzmann constant.<sup>[45]</sup>

**Control experiments.** The dielectric response of the empty nanomembrane was measured for 10 min at 60°C and from 60 to –80°C to control that no dielectric dispersions were detected. The dielectric response of nanopores filled with a 200  $\mu$ L drop of deionised water was measured for 10 min at 60°C and from 60 to –80°C to control that the deionised water used to prepare the sample show no dielectric dispersions too. Finally, the dielectric response of nanopores filled with a 200  $\mu$ L drop of PBS buffer control was

measured following a complete temperature ramp experiment as for the protein sample. The PBS buffer control was prepared with 2  $\mu$ L of PBS diluted in 204  $\mu$ L of deionised water (no protein was added). This concentration was equivalent to the concentration used to prepare and measure the protein sample. This control experiment was carried out to investigate the dielectric response of PBS mixed with deionised water in the same conditions as the protein.

To make sure the toxin was in a native state when a dielectric experiment is performed, the toxin pentameric state was checked by Trp-fluorescence beforehand.<sup>[45]</sup> The sample could not be recovered from the nanopores after the BDS experiment so it was not possible to control the state of the toxin after.

#### Network Based Models

##### Intermolecular Amino Acid Network

The toxin interfaces were modelled by a network of the amino acids that belong to different chains and had at least



one pair of atoms within a 5 Å distance, so-called hotspots.<sup>[24]</sup> The amino acid network (AAN) of the whole pentamer was built first and the intermolecular amino acid network was generated after by selecting only intermolecular interactions. The networks were created using the Python Biograph module, available at <https://github.com/rodigi/biographs>, or directly with the script available at [https://github.com/lorpac/network\\_vmd](https://github.com/lorpac/network_vmd). Briefly this was done using the atomic coordinates of the toxin structure (PDB: 1EE1) freely available from the RCSB website (<https://www.rcsb.org/>) and computing intermolecular atomic distances between amino acids that belong to different chains (nodes). Hotspots were linked when their respective atoms were close enough for a chemical interaction using 5 Å distance as a threshold. Such spatial networks based on Euclidian distances between nodes are classically used in proteins to infer atomic and amino acid interactions intractable otherwise due to the size of the system.<sup>[2,22,54]</sup> From the intermolecular amino acid network of the toxin, simple network measures such as number of hotspot pairs, number of atomic interactions, number of charged amino acid pairs, and number of charged dipole pairs within the interfaces were computed. A charged dipole pair was a charged amino acid interacting with a polar or hydrophobic amino acid. The intermolecular amino acid networks were referred to as a 4D-network for simplicity, 4D corresponds to the quaternary structure of proteins, namely the number of chains present in a protein oligomer.

#### Intermolecular Charged Amino Acid Network

The intermolecular charged amino acid networks were sub-networks of the intermolecular amino acid networks where only charged amino acids are kept as nodes of the network. The charged amino acid networks model the Coulomb interactions with charged-charged interactions within the toxin interface and were used to analyse the protein dielectric signals which result from such interactions. The intermolecular charged amino acid networks were referred to as the 4D-charged network for simplicity. The orientation of the dipoles was not taken into account in the model although it could cancel out the contribution to the dielectric signal of some of the dipoles considered in the model when dipoles of equal strength had opposite orientation. Nevertheless, the amino acid neighbourhood was sufficiently anisotropic for this to be unlikely, making the approximation locally reasonable. The isoelectric point (pI) of the toxin pentamer was around pH 5.6 such that at pH 6.9 the global charge was negative and a global dipolar moment was expected, consistent with having a dielectric signal. It also supported the approximation at the global scale of the entire protein.

#### Intermolecular Induced Charged Amino Acid Network

The intermolecular induced charged amino acid network was as the intermolecular charged amino acid network but with the addition of the non-charged amino acids located in the chemical vicinity of the charged hotspots. This network models the Coulomb interactions as well as the ion-dipole interactions with the charged non-charged amino acid pairs, present at the toxin interface. The intermolecular induced charged amino acid networks were referred to as 4D-induced charged network for simplicity.

Mathematically, if the network of all amino acid interactions is represented by a graph  $G(V, E)$  with set of nodes  $V$  (Vertices)

and set of edges ( $E$ ), then the 4D-network is represented by a graph  $G_{4D}(V_{4D}, E_{4D})$  with nodes

$$V_{4D} = \{i | i \in V \text{ and } \exists (i, j) \in E \text{ with } \text{chain}(i) \neq \text{chain}(j)\}$$

and edges

$$E_{4D} = \{(i, j) | (i, j) \in E \text{ and } \text{chain}(i) \neq \text{chain}(j)\}$$

The 4D-charged network is then represented by a graph  $G_{4D}^D(V_{4D}^D, E_{4D}^D)$  with nodes

$$V_{4D}^D = \{i | i \in V_{4D} \text{ and } i \text{ is charged}\}$$

and edges

$$E_{4D}^D = \{(i, j) | (i, j) \in E_{4D} \text{ and } i, j \in V_{4D}^D\}$$

and the 4D-induced charged network is represented by a graph  $G_{4D}^{D+}(V_{4D}^{D+}, E_{4D}^{D+})$  with nodes

$$V_{4D}^{D+} = V_{4D}^D \cup \{j | \exists (i, j) \in E \text{ for some } i \in V_{4D}^D\}$$

and edges

$$E_{4D}^{D+} = \{(i, j) | (i, j) \in E_{4D} \text{ and } i \in V_{4D}^D\}.$$

The construction of the networks is shown in Fig. S1 (Supplementary Material).

#### Network Generation and Representation

All networks were generated by manipulating the amino acid network of the toxin using the *Python* package NetworkX.<sup>[46]</sup> Two-dimensional network representations were produced using the *Gephi* software<sup>[48]</sup> and three-dimensional network representations were produced using the *Python* scripts available at [https://github.com/lorpac/network\\_vmd](https://github.com/lorpac/network_vmd) and visualised with the VMD software.<sup>[49]</sup> For simplicity, the toxin pentamer having a central symmetry, only links between nodes involved in the interfaces between chain E and D and between chain E and F are represented.

#### Supplementary Material

A schematic of the network construction flow is available on the Journal's website.

#### Author contributions

LB and LP performed research and analysed data; AS designed research, analysed data, and wrote the manuscript; CL designed research, analysed data, and wrote the manuscript.

#### Conflicts of Interest

The authors declare no conflicts of interest.

#### Acknowledgements

The authors thank the Interdisciplinary Mission of the CNRS for funding (MITI, Defi 'Instrumentations aux limites' ATTOGRAM).

#### References

- [1] K. Lindorff-Larsen, S. Piana, R. O. Dror, D. E. Shaw, *Science* **2011**, 334, 517. doi:10.1126/SCIENCE.1208351

- [2] L. Vuillon, C. Lesieur, *Curr. Opin. Struct. Biol.* **2015**, *31*, 1. doi:10.1016/J.SBI.2015.02.015
- [3] V. Munoz, M. Cerminara, *Biochem. J.* **2016**, *473*, 2545. doi:10.1042/BCJ20160107
- [4] T. R. Barends, L. Foucar, A. Ardevol, K. Nass, A. Aquila, S. Botha, R. B. Doak, K. Falahati, E. Hartmann, M. Hilpert, M. Heinz, M. C. Hoffmann, J. Köfinger, J. E. Koglin, G. Kovacsova, M. Liang, D. Milathianaki, H. T. Lemke, J. Reinstein, C. M. Roome, R. L. Shoeman, G. J. Williams, I. Burghardt, G. Hummer, S. Boutet, I. Schlichting, *Science* **2015**, *350*, 445. doi:10.1126/SCIENCE.AAC5492
- [5] M. Levantino, G. Schiro, H. T. Lemke, G. Cottone, J. M. Glownia, D. Zhu, M. Chollet, H. Ihee, A. Cupane, M. Cammarata, *Nat. Commun.* **2015**, *6*, 6772. doi:10.1038/NCOMMS7772
- [6] E. Mizohata, T. Nakane, Y. Fukuda, E. Nango, S. Iwata, *Biophys. Rev.* **2018**, *10*, 209. doi:10.1007/s12551-017-0344-9
- [7] J. R. Stagno, Y. Liu, Y. R. Bhandari, C. E. Conrad, S. Panja, M. Swain, L. Fan, G. Nelson, C. Li, D. R. Wendel, T. A. White, J. D. Coe, M. O. Wiedorn, J. Knoska, D. Oberthuer, R. A. Tuckey, P. Yu, M. Dyba, S. G. Tarasov, U. Weierstall, T. D. Grant, C. D. Schwieters, J. Zhang, A. R. Ferré-D'Amaré, P. Fromme, D. E. Draper, M. Liang, M. S. Hunter, S. Boutet, K. Tan, X. Zuo, X. Ji, A. Barty, N. A. Zatsepin, H. N. Chapman, J. C. H. Spence, S. A. Woodson, Y.-X. Wang, *Nature* **2017**, *541*, 242. doi:10.1038/NATURE20599
- [8] J. Herbst, K. Heyne, R. Diller, *Science* **2002**, *297*, 822. doi:10.1126/SCIENCE.1072144
- [9] C. Kolano, J. Helbing, M. Kozinski, W. Sander, P. Hamm, *Nature* **2006**, *444*, 469. doi:10.1038/NATURE05352
- [10] Y. Mizutani, T. Kitagawa, *Chem. Rec.* **2001**, *1*, 258. doi:10.1002/TCR.1012
- [11] C. Fang, R. R. Frontiera, R. Tran, R. A. Mathies, *Nature* **2009**, *462*, 200. doi:10.1038/NATURE08527
- [12] R. Schneider, M. Blackledge, M. R. Jensen, *Curr. Opin. Struct. Biol.* **2019**, *54*, 10. doi:10.1016/J.SBI.2018.09.007
- [13] E. W. Finsen, T. W. Scott, M. R. Chance, J. M. Friedman, M. R. Ondrias, *J. Am. Chem. Soc.* **1985**, *107*, 3355. doi:10.1021/JA00297A056
- [14] C. Lesieur, K. Schulten, *Curr. Opin. Struct. Biol.* **2015**, *31*, v. doi:10.1016/J.SBI.2015.05.008
- [15] J. R. Perilla, B. C. Goh, C. K. Cassidy, B. Liu, R. C. Bernardi, T. Rudack, H. Yu, Z. Wu, K. Schulten, *Curr. Opin. Struct. Biol.* **2015**, *31*, 64. doi:10.1016/J.SBI.2015.03.007
- [16] G. R. Heath, S. Scheuring, *Nat. Commun.* **2018**, *9*, 4983. doi:10.1038/S41467-018-07512-3
- [17] B. Hellenkamp, P. Wortmann, F. Kandzia, M. Zacharias, T. Hugel, *Nat. Methods* **2017**, *14*, 174. doi:10.1038/NMETH.4081
- [18] A. Soranno, A. Holla, F. Dingfelder, D. Nettels, D. E. Makarov, B. Schuler, *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E1833. doi:10.1073/PNAS.1616672114
- [19] D. M. Leitner, T. Yamato, *Mapping Energy Transport Networks in Proteins*, Vol. 31 **2018** (Wiley Online Library: Hoboken, NJ).
- [20] J. Liu, R. Nussinov, *PLoS Comput. Biol.* **2016**, *12*, e1004966. doi:10.1371/JOURNAL.PCBL1004966
- [21] A. N. Naganathan, *Curr. Opin. Struct. Biol.* **2019**, *54*, 1. doi:10.1016/J.SBI.2018.09.004
- [22] L. Di Paola, A. Giuliani, *Curr. Opin. Struct. Biol.* **2015**, *31*, 43. doi:10.1016/J.SBI.2015.03.001
- [23] C.-B. Li, H. Yang, T. Komatsuzaki, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 536. doi:10.1073/PNAS.0707378105
- [24] M. Achoch, R. Dorantes-Gilardi, C. Wymant, G. Feverati, K. Salamatián, L. Vuillon, C. Lesieur, *Phys. Chem. Chem. Phys.* **2016**, *18*, 13770. doi:10.1039/C5CP06091E
- [25] A. Gheeraert, L. Pacini, V. S. Batista, L. Vuillon, C. Lesieur, I. Rivalta, *J. Phys. Chem. B* **2019**, *123*, 3452. doi:10.1021/ACS.JPCB.9B01294
- [26] E. G. Marklund, J. L. Benesch, *Curr. Opin. Struct. Biol.* **2019**, *54*, 50. doi:10.1016/J.SBI.2018.12.011
- [27] M. T. Degiacomi, *Structure* **2019**, *27*, 1034. doi:10.1016/J.STR.2019.03.018
- [28] A. J. Pak, G. A. Voth, *Curr. Opin. Struct. Biol.* **2018**, *52*, 119. doi:10.1016/J.SBI.2018.11.005
- [29] H. Van den Bedem, J. S. Fraser, *Nat. Methods* **2015**, *12*, 307. doi:10.1038/NMETH.3324
- [30] M. T. Degiacomi, I. Iacovache, L. Pernet, M. Chami, M. Kudryashev, H. Stahlberg, F. G. van der Goot, M. Dal Peraro, *Nat. Chem. Biol.* **2013**, *9*, 623. doi:10.1038/NCHEMBIO.1312
- [31] C. M. Boyd, D. Bubeck, *Curr. Opin. Struct. Biol.* **2018**, *52*, 41. doi:10.1016/J.SBI.2018.07.010
- [32] E. S. Parsons, G. J. Stanley, A. L. Pyne, A. W. Hodel, A. P. Nievergelt, A. Menny, A. R. Yon, A. Rowley, R. P. Richter, G. E. Fantner, D. Bubeck, B. W. Hoogenboom, *Nat. Commun.* **2019**, *10*, 2066. doi:10.1038/S41467-019-10058-7
- [33] G. Schiro, A. Cupane, E. Vitrano, F. Bruni, *J. Phys. Chem. B* **2009**, *113*, 9606. doi:10.1021/JP901420R
- [34] H. Jansson, R. Bergman, J. Swenson, *J. Phys. Chem. B* **2005**, *109*, 24134. doi:10.1021/JP052418V
- [35] H. Frauenfelder, G. Chen, J. Berendzen, P. W. Fenimore, H. Jansson, B. H. McMahon, I. R. Stroe, J. Swenson, R. D. Young, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 5129. doi:10.1073/PNAS.0900336106
- [36] S. Khodadadi, S. Pawlus, A. P. Sokolov, *J. Phys. Chem. B* **2008**, *112*, 14273. doi:10.1021/JP8059807
- [37] L. Bourgeat, A. Serghei, C. Lesieur, *Sci. Rep.* **2019**, *9*, 17988. doi:10.1038/S41598-019-54562-8
- [38] A. Serghei, W. Zhao, D. Miranda, T. P. Russell, *Nano Lett.* **2013**, *13*, 577. doi:10.1021/NL304103Y
- [39] C. Lesieur, M. J. Clifff, R. Carter, R. F. James, A. R. Clarke, T. R. Hirst, *J. Biol. Chem.* **2002**, *277*, 16697. doi:10.1074/JBC.M110561200
- [40] J. Zrimi, A. Ng Ling, E. Giri-Rachman Arifin, G. Feverati, C. Lesieur, *PLoS One* **2010**, *5*, e15347. doi:10.1371/JOURNAL.PONE.0015347
- [41] B. Goins, E. Freire, *Biochemistry* **1988**, *27*, 2046. doi:10.1021/B100406A035
- [42] V. Bhakuni, D. Xie, E. Freire, *Biochemistry* **1991**, *30*, 5055. doi:10.1021/B100234A031
- [43] B. A. Shoemaker, J. J. Portman, P. G. Wolynes, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 8868. doi:10.1073/PNAS.160259697
- [44] K. Luke, M. Perham, P. Wittung-Stafshede, *J. Mol. Biol.* **2006**, *363*, 729. doi:10.1016/J.JMB.2006.08.058
- [45] L. W. Ruddock, J. J. Coen, C. Cheesman, R. B. Freedman, T. R. Hirst, *J. Biol. Chem.* **1996**, *271*, 19118. doi:10.1074/JBC.271.32.19118
- [46] L. W. Ruddock, S. P. Ruston, S. M. Kelly, N. C. Price, R. B. Freedman, T. R. Hirst, *J. Biol. Chem.* **1995**, *270*, 29953. doi:10.1074/JBC.270.50.29953
- [47] M. N. Blackburn, E. A. Noltmann, *Arch. Biochem. Biophys.* **1981**, *212*, 162. doi:10.1016/0003-9861(81)90355-6
- [48] K. E. Neet, D. E. Timm, *Protein Sci.* **1994**, *3*, 2167. doi:10.1002/PRO.5560031202
- [49] M. Perham, M. Chen, J. Ma, P. Wittung-Stafshede, *J. Am. Chem. Soc.* **2005**, *127*, 16402. doi:10.1021/JA055574O
- [50] D. Roberts, R. Keeling, M. Tracka, C. F. Van Der Walle, S. Uddin, J. Warwicker, R. Curtis, *Mol. Pharm.* **2015**, *12*, 179. doi:10.1021/MP500533C
- [51] M. De Wolf, G. Van Dessel, A. Lagrou, H. J. Hilderson, W. Dierick, *Biochim. Biophys. Acta* **1985**, *832*, 165. doi:10.1016/0167-4838(85)90328-0
- [52] M. J. De Wolf, G. A. Van Dessel, A. R. Lagrou, H. J. Hilderson, W. S. Dierick, *Biochemistry* **1987**, *26*, 3799. doi:10.1021/B100387A010
- [53] R. Gilardi-Dorantes, L. Bourgeat, L. Pacini, L. Vuillon, C. Lesieur, *Phys. Chem. Chem. Phys.* **2018**, *20*, 25399.
- [54] J. S. Vilorio, M. F. Allega, M. Lambrugh, E. Papaleo, *Sci. Rep.* **2017**, *7*, 1.