Analyse de l'assemblage de peptides amyloïdes bactériens

David Partouche

▶ **To cite this version:**

David Partouche. Analyse de l'assemblage de peptides amyloïdes bactériens. Biochimie [q-bio.BM]. Université Paris Saclay (COmUE), 2018. Français. NNT : 2018SACLX084 . tel-01959336

HAL Id: tel-01959336
https://pastel.hal.science/tel-01959336

Submitted on 18 Dec 2018

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.

NNT : 2018SACLX084

Analyse de l'assemblage de peptides amyloïdes bactériens

Thèse de doctorat de l'Université Paris-Saclay
Préparée à l'École Polytechnique

École doctorale n°573 Interfaces : Approches interdisciplinaires, fondements, applications et innovation

Spécialité de doctorat : Chimie

Thèse présentée et soutenue à Saint-Aubin, le 6 novembre 2018, par

M. David Partouche

Composition du Jury :

Mme Brigitte Hartmann Directrice de recherche, CNRS UMR 8113, Laboratoire de Biologie et Pharmacologie Appliquée	Présidente
M. Codjo Hountondji Professeur, Université Pierre et Marie Curie EA 7336, Enzymologie de l'ARN	Rapporteur
Mme Andrea Gomez Zavaglia Directrice de recherche, Conicet Center for Research and Development in Food Cryotechnology	Rapporteur
M. Bertrand Castaing Directeur de recherche, CNRS UPR 4301, Centre de Biophysique Moléculaire	Examinateur
Mme Véronique Arluison Maitre de conférences, Université Paris Diderot UMR 12, Laboratoire Léon Brillouin	Co-Directrice de thèse
M. Matthieu Réfrégiers Responsable de ligne de lumière DISCO Synchrotron SOLEIL	Co-Directeur de thèse

Thèse de doctorat

A mes parents,

A mes frères,

Remerciements

Le travail de thèse présenté dans ce manuscrit a été effectué entre différents laboratoires : principalement au centre de rayonnement Synchrotron SOLEIL à Saint-Aubin, mais également aux Laboratoire Léon Brillouin (LLB) se trouvant sur le site Saclay du centre du Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA) à Gif sur Yvette et le laboratoire Chimie-Physique (LCP) de l'université Paris-Sud à Orsay.

Je tiens à remercier tout d'abord le Directeur Général de SOLEIL, Jean Daillant et le Directeur Scientifique Andrew Thompson de m'avoir accueilli au sein de SOLEIL ainsi que Ferenc Borondics responsable de la ligne SMIS.

Mes premiers remerciements vont également à ma directrice et mon directeur de thèse. Je remercie chaleureusement Véronique Arluison, ma directrice de thèse. Son soutien infaillible a été essentiel à l'avancement de mes recherches. Je remercie Matthieu Reffregier, mon directeur de thèse pour sa gentillesse, ses conseils et sa disponibilité.

J'adresse mes sincères et respectueux remerciements aux membres du Jury : à M. Codjo Hountondji et Mme Andrea Gomez Zavaglia pour avoir accepté d'être rapporteurs et à Mme Brigitte Hartmann et M. Bertrand Castaing pour avoir accepté d'être examinateurs à la soutenance de ma thèse.

Merci à mes encadrant et encadrante de thèse. Je remercie Christophe Sandt pour son accompagnement pédagogique et qui m'a fait découvrir le domaine infrarouge du synchrotron. Il a été très agréable et enrichissant de discuter avec lui de sujets scientifiques et non-scientifiques. Et merci également à Ariane Deniset qui assuré ma formation à l'AFM-IR, pour ses conseils et sa patience.

Je remercie Sergio Marco et Sylvain Trépout de l'institut Curie d'Orsay dont l'aide en MET et Cryo-TEM a été primordiale dans le cadre de ma thèse.

J'exprime également une pensée reconnaissante à Thierry Bizebard du laboratoire IBPC qui a bien voulu être mon tuteur de thèse.

J'aimerais remercier l'ensemble des équipes, collègues et amis que j'ai pu rencontrer durant la thèse et avec qui j'ai apprécié travailler ou échanger :

- J’adresse un grand merci à tous les chercheurs, ingénieurs et techniciens du LLB. Un merci particulier à Marianne Bombled avec qui j’ai occupé le même bureau pendant ces trois années. Un petit clin d’œil reconnaissant à Marie-Claire Bellissent-Funel et Jose Teixeira qui ont participé à mes répétitions orales et ont su me donner de bons conseils. Merci à Nancy Brodie-Linder d’avoir partagé ses connaissances sur les matériaux.
- Merci à tous les stagiaires, doctorants et post-doctorants, que j’ai eu le plaisir de rencontrer : Aline, Summit, Marc-Antoine, Erigène, Marguerite et Rodrigo, Florian, N’Daye et Camille. J’ai vécu avec eux trois années de bonne humeur et une excellente ambiance de travail. Merci à Oriana Osta pour sa gentillesse, son soutien et qui communique tant sa joie de vivre. Merci aussi à Antoine Malabirade, pour tous ces moments passés à discuter de nos thèses respectives, pour les échanges scientifiques (et culturels) et aussi sa précieuse aide au laboratoire.
- À SOLEIL, sur les lignes DISCO et SMIS, je remercie : Frédéric, Alexandre, Francesco, Stéphane et Valérie pour leur aide mais surtout Frank Wien qui m’a initié à la SRDC et pour son implication déterminante dans mon projet de thèse. J’aimerais aussi remercier les autres scientifiques : Thomas Bizien, sur la ligne SWING qui m’a formé au traitement de données de diffusion des rayons X. Merci également à Omar El Hamoui avec qui j’ai passé tant d’heures à faire des acquisitions de spectres CD. Merci à Manon Gallard pour son amitié, l’entraide psychologique autour d’une pause-café, l’occasion d’échanger sur nos travaux de thèse ou bien débattre sur le monde.
- Au laboratoire Physique-chimie (Université Paris-Sud) : je remercie aussi Alexandre Dazzi et Jérémie Mathurin de l’équipe AFM-IR.

Je remercie également Frédéric Geinguenaud pour son aide en spectroscopie infrarouge d’eau lourde du laboratoire SMBH de l’université Paris-13 Nord.

Merci à Jean-Francois Berret du laboratoire MSC de l’université Paris Diderot qui m’a formé à la technique de microrhéologie. De la même université, je remercie également Florent Busi qui a bien voulu écouter une de mes répétitions et m’a apporté ses conseils.

Et, je remercie aussi l’équipe de mon stage de Master du laboratoire ITODYS, notamment Catherine Combellas, Jean-Marc Noel et Jérôme Médard qui m’ont soutenu et ont été présents à ma soutenance de thèse malgré leur domaine différent.

Et puis aussi : Marisela Velez (ICP de Madrid), Frédécic Gaubeaux (LIONS/CEA), Danièle Joseleau-Petit, Grzegorz Cech et Krzysiek Kubiak (Université de Gnansk/Pologne) : merci !

La thèse est aussi trois années de vie, de partage, de joie et de peine durant lesquelles de nouvelles amitiés se forment et les anciennes se renforcent. Merci à tous mes amis.

Et tout ceci n'aurait pu être possible sans la présence et l'affection de ma famille. Je tiens à remercier profondément mes parents, ainsi que mes frères Jonathan et Mickaël qui m'ont toujours encouragé sans relâche et m'ont aidé à surmonter les moments de doute et les obstacles, jusqu'au bout.

Résumé

Hfq est une protéine bactérienne qui a un rôle pléïotropique. La principale fonction de la protéine Hfq bactérienne consiste à répondre aux stress que peut rencontrer la bactérie lors d'un changement environnemental, en utilisant essentiellement un contrôle post-transcriptionnel. La protéine, par sa capacité à interagir avec les ARN et notamment les petits ARN non codant, permet ainsi une régulation rapide de l'expression génétique. En outre, la protéine interagit aussi avec l'ADN qu'elle aide à se structurer. Les mutations dans le gène qui code pour Hfq ont des effets pléïotropes c'est-à-dire déterminant plusieurs caractères phénotypiques.

D'un point de vue structural, la protéine adopte un repliement de type Sm caractérisé par un oligomère toroïdal reposant sur la formation d'un feuillet β continu à 30 brins. Cependant, outre cette région Sm N-terminale (NTR), Hfq possède également une région C-terminale (CTR) de taille et de séquence variables selon les bactéries. Mon travail de thèse a porté sur l'analyse de cette région CTR chez la bactérie *Escherichia coli*. Cette région a en effet la capacité de former une structure de type amyloïde : structures auto-assemblées *in vivo*, à proximité de la membrane interne et dans le nucléoïde.

Par l'utilisation de diverses techniques physico-chimiques (microscopie moléculaire, spectroscopie et microscopie infrarouge, dichroïsme circulaire et diffusion aux petits angles), mon travail a consisté à caractériser l'assemblage de cette région de Hfq ainsi que les facteurs l'influençant, en particulier la présence d'acide nucléique. Une partie de mon travail de thèse a aussi consisté à mettre en place une méthode d'imagerie corrélative innovante permettant d'analyser la signature chimique et morphologique d'une fibre amyloïde unique. Mon travail a enfin porté sur l'analyse de l'effet de composés inhibant l'agrégation de la structure amyloïde, ce qui pourrait constituer une piste pour développer une nouvelle classe d'antibiotiques.

Abstract

Hfq is a pleiotropic bacterial protein that determines several phenotypic characteristics. Its main function is to facilitate responses to stresses that bacteria may encounter during environmental changes, mainly by using post-transcriptional genetic control. The protein, by its capacity to interact with RNA, in particular small non-coding RNA, enables a rapid regulation of gene expression.

In addition, the protein also interacts with DNA and compacts it. From a structural point of view, the protein adopts an Sm-like fold, characterized by a toroidal oligomer formed by a continuous 30-stranded β-sheet. Besides its conserved N-terminal Sm domain, Hfq also possesses a C-terminal region (CTR) that can vary in size and sequence between bacteria. My PhD work focused on the analysis of this CTR region in *Escherichia coli* bacteria. Indeed, this region has the capacity to form an amyloid structure. This structural dynamic is related to the formation of self-assembled structures *in vivo*, in the proximity of the inner membrane and in the nucleoid.

Using various physicochemical techniques (molecular microscopy, spectroscopy and infrared microscopy, circular dichroism and small angle X-ray scattering), my work consisted in characterizing the assembly of this region of Hfq, as well as the factors influencing its assembly (in particular, the presence of nucleic acids). A part of my work consisted in setting up an innovative correlative–imaging method to analyze the chemical and morphological signature of a single amyloid fibre. Finally, my work focused on the analysis of the effect of compounds that inhibit the aggregation of the amyloid structure, which could constitute a new way to develop a novel class of antibiotics.

Liste des illustrations

Figure 1 – Schématisation de structure secondaire de type L-Sm montrant l'hélice α en position N-terminale et les cinq brins β antiparallèles. 28
Figure 2 – Structure tridimensionnelle de la protéine Hfq (code PDB 1HK9), mettant en évidence les 6 sous unités du feuillet β à 5 brins. 29
Figure 3 – Calcul des caractéristiques physico-chimiques de la séquence Hfq chez *E. coli*. 30
Figure 4 – Différentes actions de régulation effectuées par Hfq. 33
Figure 5 – Schématisation des classes d'interaction d'ARN avec Hfq. 34
Figure 6 – Quelques exemples de NAPs chez *E. coli*. 38
Figure 7 – Localisation de la protéine Hfq marquée par des clusters d'or en microscopie électronique à transmission. 39
Figure 8 – Auto-assemblage de Hfq. 41
Figure 9 – Schématisation des différentes voies d'agrégation de protéines. 45
Figure 10 – Modèle structural de la fibre amyloïde et des feuillets β croisés, formés par le peptide $A\beta_{40}$. 47
Figure 11 – Structures tridimensionnelles de fibres amyloïdes provenant de différentes sources. 48
Figure 12 – Formation progressive et processus secondaire de fibres amyloïdes. 49
Figure 13 – Thermodynamique et cinétique de formation de fibres amyloïdes. 51
Figure 14 – Domaines d'énergie et spectroscopies associées au spectre électromagnétique. 61
Figure 15 – Diagramme des différents niveaux d'énergie dans le cas d'une molécule diatomique. 62
Figure 16 – Régions des groupes fonctionnels en spectroscopie IR. 66
Figure 17 – Schéma d'un interféromètre de Michelson : principe du fonctionnement de la spectroscopie IRTF. 67
Figure 18 – Schémas des montages infrarouges les plus fréquents. 70
Figure 19 – Schéma d'un spectromicroscope confocal IRTF. 73
Figure 20 – Représentation des champs lointain et proche respectivement par des ondes propagatives et évanescentes d'un objet diffractant un rayonnement monochromatique. 74
Figure 21 – Principe et fonctionnement d'un microscope à force atomique. 75
Figure 22 – Schéma de principe de l'AFM-IR première génération : éclairage par-dessous. 78

Figure 23 – Schéma et principe de l'AFM-IR deuxième génération : éclairage par-dessus. 81
Figure 24 – Angles dièdres et structures secondaires caractéristiques. 83
Figure 25 – Déconvolution par FSD et dérivées secondes (à gauche) et ajustement de courbes (à droite) d'un spectre générique de protéine. 85
Figure 26 – Illustration des différentes polarisations d'un champ électromagnétique. 91
Figure 27 – Illustration de l'effet du dichroïsme circulaire. 92
Figure 28 – Présentation des énantiomères (L) et (D) des acides aminés, ainsi que le diagramme des transitions électroniques de la liaison peptidique dans l'UV. 94
Figure 29 – Réponses des différentes structures secondaires "pures" obtenues en CD. 95
Figure 30 – Interaction électrons/matière. 98
Figure 31– Motif typique de diffraction des fibres amyloïdes par les rayons X. 100
Figure 32 – Caractérisation de la structure amyloïde du domaine CTR-Hfq par quatre techniques complémentaires. 105
Figure 33 – Spectres de fluorescence d'une solution contenant des fibres du domaine CTR-Hfq en présence de ThT. 107
Figure 34 – Spectres IRTF et SRDC de solutions de fibres du peptide 38aa CTR-Hfq à différentes dilutions. 108
Figure 35 – Courbes SAXS du peptide CTR-Hfq en présence ou absence de ThT. . 109
Figure 36 – Suivi par IRTF de la cinétique de formation des fibres amyloïdes du domaine CTR-Hfq durant 35 jours. 110
Figure 37 – Moyennes, variances et ACP des spectres infrarouges pour les différentes souches BL21(DE3). 114
Figure 38 – ACP des spectres infrarouges pour les différentes souches MG1655. ... 116
Figure 39 – Modèles et hypothèses de l'auto-assemblage Hfq par le domaine CTR. 118
Figure 40 – Schéma du support de nitrure de silicium. 120
Figure 41 – Quelques exemples d'images des fibres CTR-Hfq obtenues par la méthode corrélative MET/AFM sur les fenêtres Si_3N_4. Les images MET sont reconstruite à l'aide du plugins MosaicJ sous images J. 120
Figure 42 – Imagerie corrélative d'une solution séchée de peptide CTR-Hfq obtenue par AFM-IR sur support de Si_3N_4 (en dehors des fenêtres). 121

Figure 43 – Images MET de fibres amyloïdes en présence et absence d'inhibiteurs.. 126
Figure 44 – Courbes SAXS de fibres amyloïdes en présence et absence d'inhibiteurs. 127
Figure 45 – Cinétiques montrant l'effet des inhibiteurs en SRCD sur la structure cross-β amyloides...... 128
Figure 46 – Visualisation en MET de l'effet d'une gamme d'ECGC sur les fibres CTR11. 129
Figure 47 – Courbes cinétiques SRDC de fibres en présence de 5mM ide EGCG 5mM incubé durant 10h...... 129
Figure 48- Détermination des températures de fusion par SRCD...... 130
Figure 49 – Analyse en Cryo-EM du complexe CTR38/ADN et comparaison avec les fibres de CTR38 seul...... 133
Figure 50 – Représentation d'une micro-aiguille en rotation dans un repère...... 135
Figure 51 – Types de réponse possibles à la rotation d'une aiguille dans un fluide viscoélastique (modèle de Maxwell)...... 136
Figure 52 – Type de réponse au mouvement d'une aiguille dans un gel viscoélastique (modèle de Kelvin Voigt). 137
Figure 53 – Schéma du dispositif de bobines utilisé pour la microrhéologie...... 138
Figure 54 – Analyses rhéologiques de l'échantillon d'ADN $(dAdT)_{60}$...... 140
Figure 55 – Analyse de la fréquence critique en fonction de la longueur des aiguilles réduites L*. 141
Figure 56 – Analyses rhéologiques de l'échantillon CTR38...... 142
Figure 57 – Analyses rhéologiques de l'échantillon complexe CTR-Hfq + $(dAdT)_{60}$. 143

Liste des tableaux

Tableau 1 – Quelques exemples d'amyloses et de leurs protéines associées. 44
Tableau 2 – Résumé des caractéristiques des bandes Amides à symétrie dans le plan dans une liaison polypeptidique. 82
Tableau 3 – Affectations empiriques pour les bandes infrarouges (IR) d'Amide I caractéristiques des structures secondaires des protéines. 85
Tableau 4 – Bandes d'absorption des groupes fonctionnels bactériens dans le moyen infrarouge. 89
Tableau 5 – Présence ou absence des fibres amyloïdes dans différentes conditions de température et concentrations en sels. 111

Liste des abréviations

ACP : Analyse en Composantes Principales

ADN : Acide Désoxyribo Nucléique

AFM : *Atomic Force Microscopy*, Microscopie à force atomique

AMP : Ampicilline

APOM : Apomorphine

ARN : Acide RiboNucléique

ARNp : Acide RiboNucléique polymerase

ATR: A*ttenuated Total Reflection*, réflexion totale atténuée

CC : *Creative Commons*

Cryo-EM : Cryo-Microscopie Électronique

CTR : *C-Terminus Region*, domaine Carboxyl-terminal

EGCG : Gallate d'épigallocatéchine

FSD *: Fourier self-deconvolution*

IDP : Intrinsically Disordered Protein, protéine intrinsèquement désordonnée

IR : InfraRouge

IRTF : Infra-Rouge à Transformée de Fourrier

MET : Microscopie Électronique par Transmission

NAP : *Nucleoid Associated Protein*, protéine intrinsèquement désordonnée

NTR : *N-Terminus Region*, domaine amino-terminal

OPO : O*ptical Parametric Ocsillator*

PDB : Protein Data Bank, banque de données des protéines

QCL : Q*uantum Cascade Laser*

RBS : Ribosome Binding Site, site de fixation du ribosome

RMN : Résonance Magnétique Nucléaire

sARN : *small ARN*, petit acide ribonucléique

SRDC : S*ynchrotron Radiation Circular Dichroism*, dichroïsme circulaire avec le rayonnement synchrotron

Sommaire

Résumé 5
Abstract 6
Liste des illustrations 7
Liste des tableaux 11
Liste des abréviations 13
Sommaire 16
Introduction 22
État de l'art 26
1. La protéine Hfq 27
1.1. Structure de la protéine Hfq 27
1.1.1. Hfq, une protéine de type Sm procaryote 27
1.1.2. Les protéines Hfq eubactériennes : domaines C- et N-terminaux 28
1.1.3. La protéine Hfq de *Escherichia coli* 29
1.2. Le rôle de Hfq *in vivo* 30
1.3. Modes d'action de Hfq 31
1.3.1. Effets de l'hybridation des ARN avec Hfq 31
1.3.2. Interaction des différentes régions structurales de Hfq avec les acides nucléiques 33
1.3.3. Autres interactions : protéines et ADN 35
1.3.4. Hfq, une protéine associée au nucléoïde (NAP) 35
1.4. Localisation cellulaire et super-structure secondaire de Hfq 38
1.5. Le domaine C-terminal de Hfq 41
2. Les amyloïdes 43
2.1. Description et caractéristiques 43
2.1.1. L'amylose 43
2.1.2. Les amyloïdes : une voie d'agrégation des protéines 44
2.1.3. Les structures amyloïdes : feuillets β croisés (cross-β) 46

2.1.4. Dynamique et mécanisme de formation des amyloïdes 48
2.1.5. Caractéristiques physico-chimiques des fibres amyloïdes 52
2.1.6. Cas particulier : les prions 52
2.2. Concept d'amyloïdes fonctionnelles 53
2.2.1. Concernant la fonction des IDPs 53
2.2.2. Les amyloïdes fonctionnelles 53
2.2.3. Interaction amyloïde-membrane et amyloïde-ADN 54
2.2.4. Exemples d'amyloïdes bactériens 55
2.3. Inhibiteurs d'amyloïdes 56
2.3.1. Introduction : inhibiteurs d'amyloïdes 56
2.3.2. Les différents types d'inhibiteurs 57
Matériel et méthode : Panorama des techniques utilisées pour l'étude des amyloïdes. 59
1. Introduction - Interaction lumière/matière 60
2. La spectroscopie infrarouge 63
2.1. Introduction à la spectroscopie infrarouge 63
2.1.1. L'absorption en infrarouge 64
2.1.2. La spectroscopie infrarouge à Transformée de Fourier 66
2.2. Mode de mesure et échantillonnage 68
2.3. Imagerie chimique et résolution spatiale 71
2.3.1. Spectromicroscopie infrarouge en champ lointain 71
2.3.2. Microscopes en champ proche 74
2.3.3. Microscope à force atomique 75
2.3.4. Spectromicroscope infrarouge en champ proche 77
2.3.4.1. Introduction 77
2.3.4.2. Techniques en champ proche thermique 77
2.3.4.3. L'AFM-IR : présentation et fonctionnement 78
2.3.4.4. Évolution et amélioration de AFM-IR 79
2.4. Étude de la structure secondaire des protéines par IRTF 81

2.4.1. Les bandes d'absorption polypeptidique 81
2.4.2. Détermination de la structure secondaire par la bande amide I 82
2.4.3. Analyses structurales avec les autres bandes amides 86
2.4.4. Détermination de la structure *in vivo* 87
2.5. Conclusion sur la spectroscopie infrarouge 89
3. Dichroïsme circulaire 90
3.1. Introduction 90
3.2. Le principe de la spectroscopie CD 91
3.3. Détermination des structures secondaires par dichroïsme circulaire 93
3.3.1. Échantillons de protéines ou de peptides 93
3.3.2. Échantillons d'acides nucléiques 95
3.4. Conclusion 96
4. Techniques complémentaires pour sonder les amyloïdes 97
4.1. La microscopie électronique en transmission 97
4.2. La diffusion de rayons X aux petits angles (SAXS) 99
4.3. Spectroscopie de fluorescence : Thioflavine T colorant des amyloïdes 101
Études : résultats et discussion 102
Chapitre I : caractérisation de la structure amyloïde du domaine CTR de Hfq 103
1. Identification de l'amyloïde du domaine C-terminal d'Hfq 103
1.1. Identification par un ensemble de techniques physico-chimiques 103
1.2. Mesures cinétiques 106
1.2.1. Test en présence de Thioflavine T 106
1.2.2. Cinétiques en spectroscopie IRTF 109
1.2.3. Analyses *in vivo* 111
1.2.3.1. Souche BL21(DE3) 111
1.2.3.2. Souche MG1655 115
2. Discussion 117

Chapitre II : mise en place d'une technique corrélative pour l'étude de nano-objets : application sur les fibres amyloïdes 119

1. Présentation de l'article 119

2. Discussion 120

Chapitre III : rôle de la formation amyloïde d'Hfq sur la topologie de l'ADN 122

1. Présentation de l'article 122

2. Discussion 123

Chapitre IV : interaction entre le domaine C-terminal de la protéine Hfq et des inhibiteurs d'amyloïdes 125

1. Études de APOM et EGCG sur le peptide CTR38 126

2. Études de l'effet EGCG sur le peptide 11aa 128

3. Études de l'effet EGCG *in vivo* 131

Chapitre V : caractérisation structurale des fibres amyloïdes/ADN et analyse des propriétés rhéologiques du système. 132

1. Introduction 132

2. Analyse du complexe CTR38/ADN 133

3. Microrhéologie : la spectroscopie magnétique rotationnelle 134

4. Montage de l'expérience 137

5. Résultats et discussion 138

5.1. Double brin d'ADN (dAdT)60 seul 139

5.2. Peptide CTR38 seul 141

5.3. Complexe CTR38 + $(dAdT)_{60}$ 143

Conclusion 145

Bibliographie 148

Annexe A : Article 1 Techniques to analyze sRNA protein cofactor self-assembly in vitro 165

Annexe B : Article 2 Correlative infrared nanospectroscopy and transmission electron microscopy to investigate nanometric amyloid fibrils: Prospects and challenges 186

Annexe C : Article 3 Revised role for Hfq bacterial regulator on DNA topology 207

Annexe D : Article 4 Tailoring the Hydrophobicity of Mesoporous Organosilica for Protein Trapping and Supported Catalysis 241

Introduction

De nombreuses pathologies chez l'Homme comme les maladies d'Alzheimer, de Parkinson, ou le diabète de type II sont associées à la présence de structures dites amyloïdes dans l'organisme. Ce type de structure consiste en des dépôts riches en feuillets β croisés formés par des peptides et protéines. Les protéines ayant la faculté d'adopter ce repliement amyloïde n'ont pas *à priori* de relation structurale ni fonctionnelle communes. Cependant, lors de leur agrégation en fibres amyloïdes, elles forment une nanostructure analogue considérée comme cytotoxique. Récemment, des formes amyloïdes fonctionnelles et utiles à la cellule ont été découvertes dans de divers organismes [1]. Des mammifères aux bactéries, la fonction de ces amyloïdes fonctionnelles est variée : formation de biofilms, développement de structures cellulaires, transfert de l'information, ou bien encore support du mécanisme de défenses immunitaires ou du processus de mort cellulaire programmé [2].

Bien que longtemps ignorée, ce type de structure existe aussi chez les bactéries. La cellule bactérienne présente en effet de multiples assemblages protéiques, jouant par exemple un rôle important lors de la division cellulaire. Mes travaux de thèse ont consisté à analyser le processus d'assemblage en structure amyloïde d'une protéine bactérienne appelée Hfq. La protéine Hfq de *Escherichia coli* (*E. coli*) a été identifiée comme faisant partie de la famille des amyloïdes il y a environ cinq ans. Connue depuis les années 1970 comme un facteur important pour la réplication d'un virus, le bactériophage *Qβ* [3], Hfq est aujourd'hui connue pour avoir un rôle pléïotropique dans la cellule bactérienne. Ainsi, elle est principalement connue pour son action riborégulatrice. La protéine est présente chez un peu plus de la moitié des bactéries, mais son rôle reste peu connu chez les bactéries Gram*(+)*. Il est en revanche beaucoup mieux connu chez les bactéries Gram*(-)*. L'exemple le plus étudié est probablement la protéine Hfq de *E. coli*. C'est cette protéine particulière qui est l'objet de la présente étude.

Des travaux récents ont montré que Hfq s'auto-assemble et forme des superstructures en forme d'hélices sous-membranaires [4]. Ces structures, que l'on pense être fibrillaire, ont été observées i*n vivo* par imagerie immunofluorescence conventionnelle. Cependant, la résolution des images limitée par la diffraction, n'a pas permis d'identifier la nature précise de ces structures à ce jour (il est par exemple possible que les fibres ne soient pas continues mais organisées en « *patchs* »). En parallèle, des études *in vitro* ont montré que Hfq est capable de s'auto-assembler de façon spontanée sous forme de longues fibres amyloïdes. Ces fibres ont pu être caractérisées par microscopie électronique et par spectroscopie infrarouge [5]. Il a aussi pu être montré que c'est la région Carboxyl-Terminal de Hfq (CTR-

Hfq) qui est à la base de cette structure et de cet assemblage [6]. Mais plusieurs points d'ombre subsistent, notamment en termes de cinétiques d'assemblage qui semblent différentes *in vivo* et *in vitro*, ce qui rend difficile la compréhension de la fonction mais aussi la formation *in vivo* des nanostructures. Il est donc essentiel, afin de mieux comprendre la fonction et plus généralement le rôle de Hfq, d'analyser précisément ces structures.

Le domaine C-terminal de la protéine Hfq a été jusque-là peu étudié. Mes travaux ont consisté à caractériser la formation des structures i*n vivo* et *in vitro*, et à voir en quoi l'interaction avec des acides nucléiques influence la structure. Enfin, nous nous sommes intéressés aux effets de petites molécules inhibitrices sur la formation des structures amyloïdes.

Ce manuscrit se divisera en cinq parties :

- ⇨ **État de l'art :** dans cette première partie est présentée la protéine Hfq, ses fonctions, et sa structure, principalement chez *E. coli*. Une description des fibres amyloïdes est aussi exposée en seconde partie.

- ⇨ **Matériel et méthode : panorama des techniques utilisées pour l'étude des amyloïdes.** Les principales techniques pour analyser la formation et caractériser les structures amyloïdes y sont présentées d'un point de vue théorique.

- ⇨ **Études : résultats et discussion**

  - **Chapitre I :** « Caractérisation de la structure amyloïde du domaine C-terminal de Hfq ». Dans ce chapitre, nous étudierons la structure amyloïde du domaine C-terminal de Hfq de *E. coli*. Une partie détaille l'aspect technique des mesures en se basant sur un chapitre de livre rédigé dans le cadre de cette thèse. Par la suite, nous présenterons l'analyse par infrarouge de la présence de structure amyloïde *in vivo* dans les différentes souches d'*E. coli*. Les travaux de cette deuxième partie n'ont pas été publiés à ce jour.

  - **Chapitre II :** « Mise en place d'une technique corrélative pour l'étude de nano-objets : application sur les fibres amyloïdes ». Dans ce chapitre, nous exposerons une technique d'imagerie corrélative nanométrique présentée dans une publication soumise dans le cadre de cette thèse.

- **Chapitre III :** « Rôle de la formation amyloïde d'Hfq sur la topologie de l'ADN ». Dans ce chapitre, nous présenterons le rôle de Hfq sur la topologie de l'ADN, ainsi que l'effet de l'ADN sur la structure de Hfq. Ce chapitre repose sur une publication soumise.

- **Chapitre IV :** « Interaction entre le domaine C-terminal de la protéine Hfq et des molécules inhibitrices ». Dans ce chapitre, nous présenterons les résultats obtenus, notamment le Gallate d'épigallocatéchine (EGCG). Les résultats de ce travail n'ont pas encore été publiés mais font l'objet d'une publication en cours de rédaction.

- **Chapitre V :** « Caractérisation structurale des fibres ADN/amyloïde et analyse des propriétés rhéologiques du système ». Dans cette dernière partie, nous présenterons les premiers éléments d'une étude en cours montrant l'influence qu'a la formation du complexe amyloïde/ADN sur la dynamique du système. Les conséquences de cette dynamique *in vivo* seront discutées.

⇨ **Conclusion**

⇨ **Annexes :** articles publiés ou en cours de publication**.**

État de l’art

1. La protéine Hfq

La protéine Hfq a été identifiée dans les années 1970 comme un facteur indispensable à la réplication d'un virus infectant les bactéries, le bactériophage Qβ, d'où son acronyme Hfq pour ***H**ost **F**actor **Q**β*. Ce n'est que des années plus tard qu'a été mis à jour le rôle de Hfq chez les bactéries, principalement lié au métabolisme des ARN. Dans cette partie, nous nous intéresserons au rôle de Hfq dans la régulation de l'expression génétique, et plus particulièrement à la protéine de *E. coli*, étudiée dans le cadre le cadre de cette thèse. Les protéines Hfq des autres bactéries, notamment des bactéries Gram positives (+), sont en effet très différentes et ne possèdent pas en général de région C-terminale (CTR), sur laquelle a porté cette thèse.

1.1. Structure de la protéine Hfq

1.1.1. Hfq, une protéine de type Sm procaryote

Hfq fait partie de la grande famille des protéines L-Sm (pour Sm-Like). Celles-ci sont en général retrouvées chez les eucaryotes [7][8]. Cette famille de protéines est caractérisée par une structure conservée (voir ci-dessous), bien que leurs séquences soient peu conservées. La conservation de séquence se retrouve essentiellement dans deux courts motifs nommés Sm1 et Sm2. Les fonctions de ces protéines peuvent être diverses, mais sont en général reliées au métabolisme des ARN: régulation post-transcriptionnelle (épissage, dégradation…), maintenance des télomères, traduction… [9][10]

La caractéristique de la famille des L-Sm est leur structure tridimensionnelle, en forme d'anneau (ou tore) hexamèrique ou heptamérique. Les protéines L-Sm retrouvées chez les eubactéries présentent une forme homo-hexamèrique. Celle des archées sont homo-heptamériques, tandis que leurs homologues eucaryotes sont hétéro-heptamèrique [11]. Le tore se forme grâce à la structure secondaire particulière des monomères, consistant en une hélice-α N-terminale suivie d'un feuillet β antiparallèle à cinq brins. L'interaction des feuillets entre différents monomères permet l'assemblage en tore. Les monomères s'assemblent entre eux par interaction entre le feuillet β5 d'une sous-unité et le feuillet β4* d'une autre sous-unité. Cette interaction se fait de manière orientée, conduisant à l'oligomérisation, où toutes les hélices α sont situées sur la même face de l'anneau [12]. Cette structure impliquant un feuillet β continu à trente brins chez Hfq (et de trente-cinq chez les Sm d'archae et eucaryotes) permet aux protéines Sm d'être thermostables. Hfq n'est par

exemple pas totalement dissociée sur un gel de type SDS-PAGE et résiste à un chauffage à 80°C de plusieurs minutes. Ces caractéristiques sont communes aux protéines Sm.

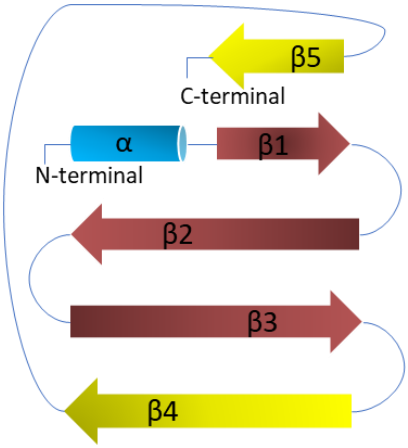


Figure 1 – Schématisation de structure secondaire de type L-Sm montrant l'hélice α en position N-terminale et les cinq brins β antiparallèles. Le motif Sm1 regroupe les trois premiers brins β en rouge. Le motif Sm2, quant à lui, rassemble les deux derniers brins, en jaune.

Il a été établi que les trois premiers feuillets β forment le motif appelé Sm1, dont la séquence est assez bien conservée entre les protéines L-Sm. Les deux derniers feuillets β forment le motif Sm2, qui est plus variable entre les différentes protéines L-Sm.

Il convient de noter que, chez les eucaryotes, il est possible de dénombrer au moins une quinzaine de protéines L-Sm différentes. Par exemple, les protéines L-Sm impliquées dans l'épissage sont constituées par sept sous-unités codées par sept gènes distincts, tandis que chez les bactéries, il n'en existe généralement qu'une (parfois deux), appelée Hfq.

1.1.2. Les protéines Hfq eubactériennes : domaines C- et N-terminaux

Généralement d'une longueur d'une centaine d'acides aminés, la séquence des protéines Hfq bactériennes peut être séparée en deux domaines. Le domaine N-terminal correspondant à la partie Sm et formant le tore est plutôt bien conservé entre les espèces bactériennes, quand Hfq est présente car toutes les bactéries n'ont pas un gène *hfq*. La structure de cette région est bien connue car les protéines Hfq de *Staphylocoque aureus*, *E. coli* et *Pseudomonas aeruginosa*, entre autres, ont été radio-cristallographiées [13][14][15]. Cependant, la structure du domaine C-terminal n'a pu être obtenue dans aucune de ces études. Celui-ci peut d'ailleurs être très différent d'une souche à une autre, voire même être absent (c'est le

cas pour *S. aureus*). Il peut aussi à l'inverse être très long comme par exemple chez *Sphingomonas assacharolytica* (160 acides aminés) ou *Acinetobacter baylyi* (172 acides aminés). Bien que de structure inconnue, le domaine C-terminal semble cependant avoir une fonction *in vivo* [16][17].

Cette thèse a eu pour objet d'étudier la structure et la fonction du domaine C-terminal (CTR) de la protéine Hfq de *E. coli*.

1.1.3. La protéine Hfq de *Escherichia coli*

Hfq chez *E. coli* est une protéine de 102 acides aminés d'une masse molaire de 11kDa environ. Plus précisément, le gène *hfq* code pour 102 acides aminés, mais après maturation la méthionine N-terminale est enlevée et la protéine mature ne fait que 101 acides aminés. Hfq existe *in vivo* sous forme hexamérique, et sa présence dans la cellule sous forme monomérique est très controversée. Son domaine N-terminal (NTR) est composé d'une soixantaine d'acides aminés et se replie sous forme Sm (voir ci-dessus et Figure 2) : son domaine C-terminal est non structuré et fait une quarantaine d'acides aminés.

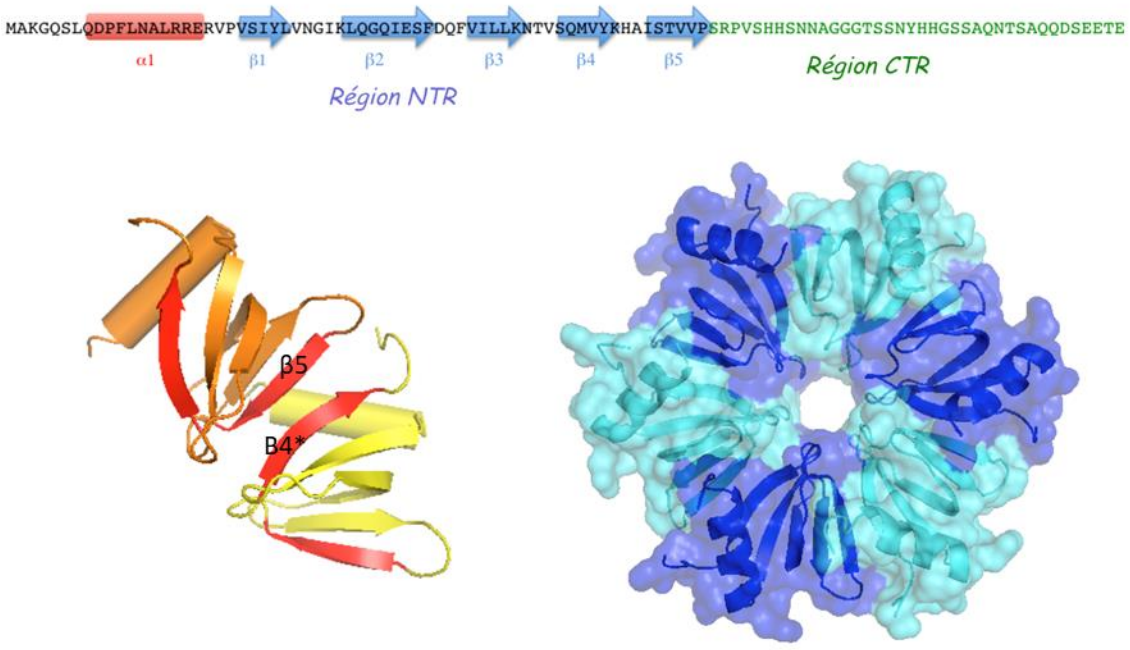


Figure 2 – Structure tridimensionnelle de la protéine Hfq (code PDB 1HK9), mettant en évidence les 6 sous unités du feuillet β à 5 brins. Haut : séquence de la protéine Hfq chez *E. coli* indiquant la position des éléments de structure secondaire dans le NTR (hélice α suivie de cinq brins β) et le CTR (en vert). Bas gauche :

structure 3D du NTR (72 acides aminés). On peut voir l'interaction des monomères avec une sous unité voisine par le biais d'une interaction entre les brins β5 et β4* de deux sous-unités différentes. Bas droit : structure toroïdale de l'hexamère (régions CTR absentes).

Quant au domaine C-terminal, il semble être principalement désordonné (Figure 3), ce qui explique l'impossibilité de déterminer sa structure par les techniques de radiocristallographie. En fait, la présence même de cette région rend la cristallisation de la protéine entière impossible, ce qui est probablement dû à son auto-assemblage (voir ci-dessous). On le supprime donc en général génétiquement pour favoriser la cristallisation [14]. Cependant, il a été avancé que la flexibilité et la charge plutôt positive de cette région CTR permettent à Hfq d'interagir avec les ADN et avec elle-même [18] [19].

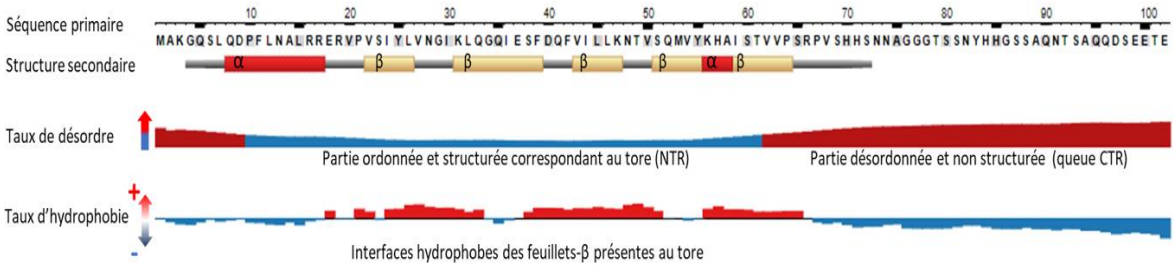


Figure 3 – Calcul des caractéristiques physico-chimiques de la séquence Hfq chez *E. coli*. Structure secondaire tirée de 1HK9 [14] et de UniProtKB. Taux de probabilité de désordre de la séquence primaire d'Hfq calculé selon [20]. En rouge : région potentiellement désordonnée ; En bleu : région probablement ordonnée. Le taux d'hydrophobie a été calculé en utilisant une fenêtre glissante de 15 résidus et en additionnant les scores des tables d'hydrophobicité standard [21]. En rouge : hydrophobe et en bleu : hydrophile.

1.2. Le rôle de Hfq *in vivo*

Hfq est un acteur important dans l'expression génétique des bactéries. Un mutant délété du gène *hfq* chez *E. coli* (mutant Δ*hfq*) montre de nombreux phénotypes (effet pléiotropie d'une mutation) tels qu'une sensibilité accrue aux stress (entre autres sensibilité accrue à la pression osmotique, aux rayonnements UV et aux changements d'environnement en général) mais aussi une virulence affaiblie (diminution de la capacité de coloniser et survivre dans l'hôte infecté) [22]. On parle ici d'un rôle pléïotropique.

Les cellules Δ*hfq* présentent aussi une différence au niveau de leur morphologie, avec un allongement des cellules [23], montrant un rôle dans le contrôle de la division cellulaire (régulation des gènes *ftsZ* et *mreB*) [24] [25].

Le gène *hfq* codant pour la protéine Hfq comporte plusieurs promoteurs (séquence d'ADN située en amont de la séquence codant la protéine) permettant la transcription de l'ARN messager (ARNm) *hfq*. Mais la synthèse de Hfq est autorégulée également au niveau traductionnel. La protéine limite sa traduction en se liant à son propre ARNm et en bloquant le site de fixation du ribosome (RBS) [26]. Hfq est abondante dans la cellule, on compte de 6 000 à 10 000 hexamères par bactérie (environ 60 000 monomères), avec une concentration maximum mesurée à 15µM (ce qui est comparable à la concentration de ribosomes). Mais les mesures sont différentes selon les souches, la phase de croissance et les conditions de culture [27]. La protéine est par exemple plus abondante en phase stationnaire.

1.3. Modes d'action de Hfq

1.3.1. Effets de l'hybridation des ARN avec Hfq

Hfq fait partie de la famille des protéines dites chaperon à ARN, permettant l'appariement entre deux séquences d'ARN [28]. Dans le cas de Hfq, ce rôle est important pour permettre l'hybridation d'un ARN régulateur non codant (ARNnc) avec un ARN messager (ARNm). Il faut noter que, chez les bactéries, ces ARNnc sont en général de petite taille et sont appelés sRNA pour « *small RNA* ». Le terme ARNnc est plutôt réservé aux eucaryotes.

Cette hybridation se fait en trois étapes. La première est l'interaction des ARN avec Hfq entraînant l'ouverture et le dépliement de la structure secondaire de l'ARN (surtout au niveau de l'ARNnc qui est en général très structuré). Cela permet dans un second temps un appariement sARN:ARNm:Hfq. Hfq permet aussi la formation d'un duplex en rapprochant les deux ARN. Ce processus est suivi par l'appariement ARNm:sARN. Enfin Hfq libère le duplex. Ce mécanisme est rapide, permettant à la bactérie une riborégulation efficace [29] [30] [31] [32]. La formation du complexe sARN-Hfq-ARNm a plusieurs conséquences (Figure 4) :

(A) L'inhibition de la traduction de l'ARNm : la fixation d'un sARN par l'intermédiaire de Hfq sur ARNm cache le site de fixation du ribosome (RBS), empêchant la traduction. On peut retrouver cette inhibition par exemple pour les

couples sARN/ARNm suivants : SgrS/*ptsG*, permettant la régulation et le transport du glucose-phosphate [33], RprA/*csgD* permettant la formation de biofilm pendant la phase stationnaire et répondant au stress [34], MicA/*ompA* permettant la synthèse de protéine de membrane externe [35] et RyhB/*sodB* permettant la régulation et la disponibilité en fer chez *E. coli* [36].

(B) Hfq permet d'accélérer la dégradation des ARNm qui ne sont plus traduits, étant devenus inutiles. L'arrêt de la traduction (voir (A)) entraîne la dégradation de l'ARNm sous l'action d'une RNase, principalement la RNAse E, endonucléase majeure chez *E. coli* [37]. En fait, cette dégradation peut avoir lieu de manière passive résultant uniquement de l'arrêt de la traduction, mais aussi de manière active, Hfq interagissant avec la RNAse E et attirant avec elle la machinerie de dégradation des ARN [38].

(C) Moins fréquent, Hfq peut au contraire libérer un RBS permettant la traduction de l'ARN. Cette libération est retrouvée pour les couples sARN/ARNm concernant le stress par choc thermique ou osmotique par la traduction du facteur sigma de la phase stationnaire, sigmaS. La traduction du gène *rpoS* (codant pour sigmaS) est stimulé par au moins deux sARN, DsrA pour le choc thermique « basse température » [39] et RprA pour un choc osmotique [40].

(D) Enfin Hfq permet également de protéger la dégradation par des ribonucléases des sARN [41].

Les effets décrits ci-dessus ont donc un rôle dans la régulation de la traduction et de la dégradation des ARN et sont dits post-transcriptionnels. Mais Hfq peut aussi avoir un effet de régulation sur la transcription, bien que celui-ci soit encore méconnu [42] [43].

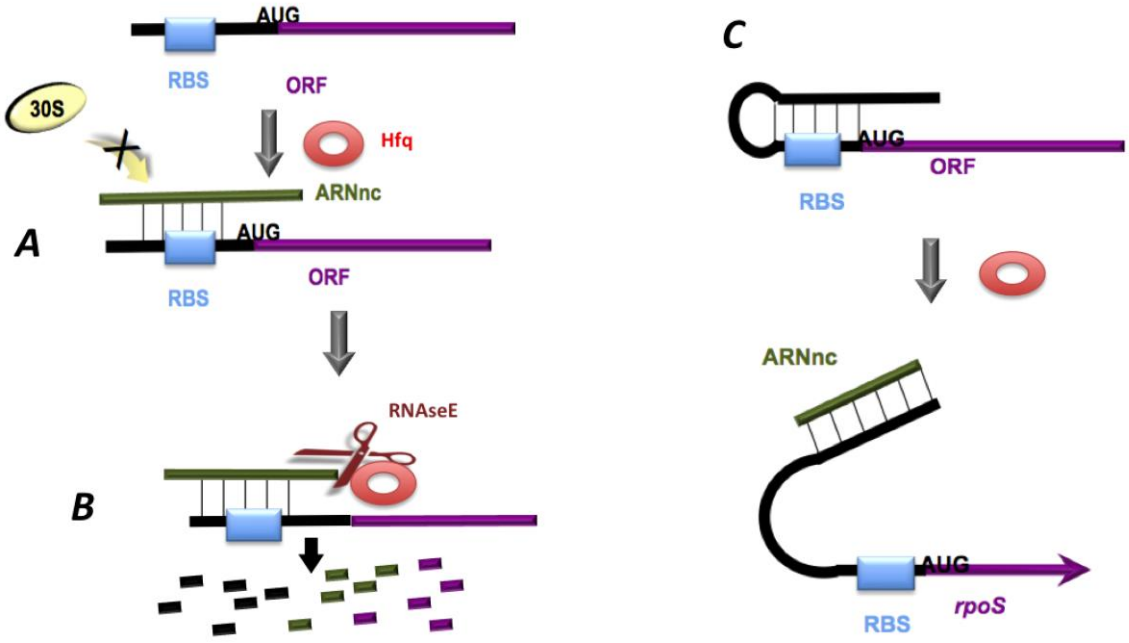


Figure 4 – Différentes actions de régulation effectuées par Hfq. (A) Inhibition de la traduction de l'ARNm par le blocage du RBS par un sARN grâce à Hfq. (B) Dégradation de sARN et ARNm par la RNase E. (C) Activation de la traduction par la libération du RBS à l'aide d'un sARN.

Il a pu être montré que Hfq est capable d'interagir avec plus de 500 ARNm et plus de 100 sARN différents. Les affinités des liaisons ARN-Hfq sont très variables, et peuvent descendre jusqu'à quelques pM [44]. Ce nombre important de partenaires rend difficile la compréhension de discrimination et donc de régulation par Hfq car, bien que la protéine soit abondante, il pourrait y avoir compétition entre les ARN. Des études récentes cherchent à montrer de nouveaux mécanismes de régulation avec des protéines tierces permettant, par exemple, la liaison des ARN dont l'affinité pour Hfq est faible [45] [46].

1.3.2. Interaction des différentes régions structurales de Hfq avec les acides nucléiques

Comme nous l'avons vu précédemment, on peut distinguer deux régions : le domaine N-terminal et le domaine C-terminal. Le domaine N-terminal est le domaine possédant un repliement de type Sm permettant de former un tore. On distingue donc trois régions dans Hfq : la face proximale (où se trouvent les hélices α), la face distale (face opposée du tore) et le domaine C-terminal périphérique [47].

La face proximale a tendance à interagir avec les régions d'ARN riches en uraciles avec une stœchiométrie d'un nucléotide par monomère. Les sARN possèdent en général une séquence riche en uracile qui sera donc liée par cette face [28]. La face distale interagit quant à elle plutôt avec des régions riches en adénine, et en particulier avec le motif Adénine (A), Purine (R) et un nucléotide quelconque (N) (noté « ARN » sur la figure 5). Du fait de la présence de séquence riches en A dans la région 5'UTR des ARNm, cette face liera plutôt les ARNm [48][49].

Deux classes d'interaction ont été déterminées entre Hfq et sARN. Concernant la première classe (classe I), le sARN se fixe sur la face proximale, tandis que l'ARNm se fixe sur la face distale. L'appariement sARN/ARNm se fait avec le détachement de Hfq. La seconde classe (classe II) regroupe les sARN se fixant sur les faces proximales et distales, tandis que l'ARNm s'enroule autour de l'anneau. Néanmoins, certaines interactions ne peuvent être classifiées dans aucune de ces deux catégories [50].

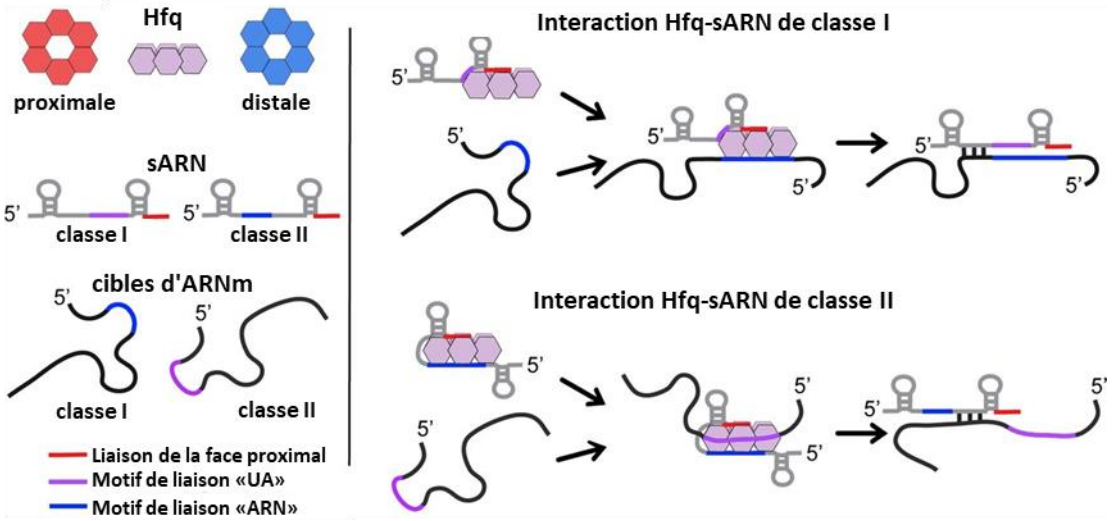


Figure 5 – Schématisation des classes d'interaction d'ARN avec Hfq. Les séquences des sARN et ARNm présentent différents éléments : en rouge, la séquence qui se lie à la face proximale riche en uracile ; en violet, la séquence « UA » riche en uridine et adénine qui se lient au bord/anneau latéral contenant trois arginines consécutives (hélice α). Les éléments en bleu représentent des motifs de liaison ARN à la face distale riche en adénine. La classe I : sARN avec son extrémité 3' riche en uracile se lie à la face distale tandis que l'ARNm se lie avec le motif « ARN » riche en nucléotides adénine et purine. La classe II : sARN se lie à la fois aux faces proximale et distale, et ARNm se lie au bord (modifiée de [50]).

1.3.3. Autres interactions : protéines et ADN

Comme nous l'avons vu, Hfq est acteur dans plusieurs voies de régulation par son action liée à l'ARN. Cependant, des études ont démontré qu'elle peut également se lier à d'autres protéines. Des interactions entre Hfq et la RNAseE ainsi que l'ARN polymérase ont été démontrées [51][38]. Il a également été montré que Hfq interagit avec la poly(A) polymérase bactérienne, ce qui influencera l'efficacité de dégradation des ARNm [52].

1.3.4. Hfq, une protéine associée au nucléoïde (NAP)

Hfq interagit aussi avec l'ADN et pourrait donc remplir des fonctions de régulation liées à l'ADN. Hfq affecte en effet de manière significative la structure de l'ADN [53], ce qui pourrait avoir un impact sur des processus tels que la réplication ou la transcription [54]. Des études ont montré que Hfq fait partie de la famille des protéines associées au nucléoïde (NAP, de l'anglais *Nucleoid Associated Protein*) [55][56][57][58].

Les NAPs ont une fonction essentielle dans la compaction et l'organisation de l'ADN chez la bactérie. En effet, l'ADN chromosomique, contrairement aux cellules eucaryotes, n'est pas contraint par une membrane. Cependant, il forme une structure compacte, due à l'implication d'une dizaine de protéines : le nucléoïde. Ces protéines permettent de replier et compacter cette macromolécule mille fois plus grande que la cellule bactérienne la contenant. Au moins une dizaine de protéines NAPs ont été identifiées, parmi lesquelles on pourra citer les protéines CbpA (« *curved DNA-binding protein A* »), CbpB (« *curved DNA-binding protein B* »), DnaA (« *DNA-binding protein A* »), Dps (« *DNA-binding protein from starved cells* »), Fis (« *factor for inversion stimulation* »), Hfq, H-NS (« *histone-like nucleoid structuring protein* »), HU (« *Histone-like protein of U93* »), IciA (« *inhibitor of chromosome initiation A* »), IHF (« *integration host factor* »), Lrp (« *leucine-responsive regulatory protein* ») et StpA (« *suppressor of td(-) phenotype A* »). Les affinités de ces NAPs peuvent aller de 25 nM (pour HU) à 250 nM (pour Hfq). Les NAPs peuvent avoir des mécanismes variés. Elles ont pour propriété commune de se lier et de modifier la conformation de l'ADN et parfois son superenroulement. Elles peuvent agir comme agent pontant, comme agent pliant et/ou rigidifiant. Ainsi, l'on donne souvent comme exemple de protéine pontant l'ADN la protéine H-NS, alors que HU est le modèle le plus couramment cité comme protéine pliant et rigidifiant l'ADN.

Exemple, la protéine HU : la protéine HU a été mise en évidence pour la première fois chez la souche *E. coli* U93 en 1975 et sa fonction était définie alors comme semblable à celle d'une histone, d'où son nom H pour *Histone-like protein* et U pour la souche U93. HU est un hétérodimère chez *E. coli* car elle est composée d'un monomère α et d'un monomère β d'environ 9 kDa chacun. Lorsque la protéine est libre, (non fixée) sa conformation est plutôt désordonnée. Mais lors d'une fixation, un feuillet β se forme et chaque monomère entoure l'hélice d'ADN, ce qui la courbe et forme un tore (rotation d'un volume circulaire autour d'un axe). HU est la protéine la plus abondante dans le nucléoïde. Lorsque de nombreux dimères se fixent dans une même zone, ils peuvent former des filaments rigides d'ADN. Cette compaction extrême de l'ADN permet à la bactérie de contraindre son génome, mais également de pouvoir répondre à divers stress. L'angle de rotation de l'ADN par HU varie selon la séquence fixée, contrairement aux autres protéines du nucléoïde, ce qui facilite et multiplie les interactions possibles [59].

Exemple, la protéine H-NS : la protéine H-NS, pour *Histone like Nucleoid Structuring*, est une protéine également très abondante dans le nucléoïde de *E. coli* (environ 20 000 copies par cellule) [60]. Elle forme un homodimère de 2 x 15,6 kDa capable de se fixer sur des séquences d'ADN spécifiques riches en A/T (adénine et tyrosine) [61]. La structure complète de la protéine est encore méconnue. La structure de sa partie N-terminale a pu être obtenue par radiocristallographie tandis que celle de la partie C-terminale a été observée par RMN. La partie N-terminale permet la dimérisation de la protéine en utilisant un motif de type « *leucine zipper* » ou « *coiled-coil* » [61]. Suite à la liaison à l'ADN, d'autres dimères H-NS se fixent sur la séquence, formant des polymères de la protéine le long de la chaine d'acide nucléique. Ces interactions entre protéines peuvent avoir lieu sur de grandes distances, ce qui permet de compacter des régions d'ADN éloignées et ainsi de structurer complètement le nucléoïde par des phénomènes de pontage ou « *bridging* » [62]. H-NS inhibe alors la transcription dans les zones où elle se fixe et permet donc de réguler l'expression génétique de manière pléïotropique. On pense, mais ce point sera discuté dans le chapitre III, que Hfq aurait une fonction proche de celle de H-NS, mais en utilisant une région amyloïde pour ponter l'ADN alors que H-NS utilise une région de type *leuzine zipper*. De plus les deux protéines ont des spécificités de séquence assez proches [63] et pourraient interagir physiquement et fonctionnellement [64][65].

En plus de compacter le génome, toutes ces interactions ont un rôle de modulation des processus liés à l'ADN (transcription, réplication, réparation…) et à sa réorganisation

dynamique. Elles peuvent ainsi par exemple influencer positivement ou négativement la transcription en ouvrant ou contraignant des boucles d'ADN. Il faut aussi savoir que les NAPs ne sont pas exprimées de manière constante lors du cycle cellulaire, et leur expression dépend de la phase de croissance et du changement environnemental. Les différentes NAPs ont une abondance propre, ainsi que des fonctions et des affinités différentes pour l'ADN. *In vivo*, les NAPs sont généralement des régulateurs pléïotropiques, ce qui rend difficile la compréhension de leur mission. Mais on peut déterminer plus facilement le type de contraintes exercées par les NAPs sur l'ADN par des analyses *in vitro*.

Il a été par différentes techniques montré que Hfq, de manière comparable à d'autres NAPs telles que HU et H-NS, condense fortement l'ADN [19]. En fait, à ce jour, Hfq est la protéine qui a le plus fort effet compactant *in vitro*, même si elle est moins abondante que HU et H-NS et avec une affinité moindre.

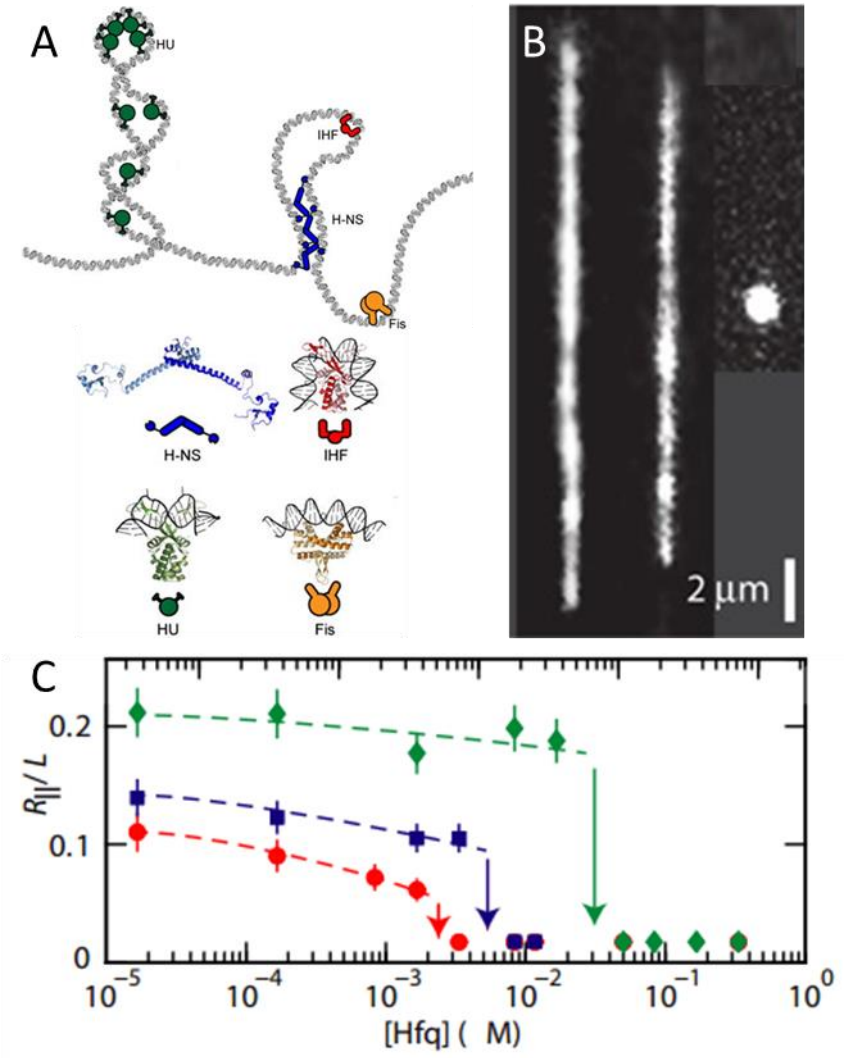


Figure 6 – Quelques exemples de NAPs chez *E. coli*. (A) Schéma présentant les effets des NAPs sur l'ADN et leurs structures PDB (1P78 pour HU, 1IHF pour IHF, 3JRA pour Fis et une combinaison de 3NR7 et 1NHR pour H-NS). HU, IHF et Fis plient l'ADN tandis que H-NS recouvre et ponte les brins d'ADN. Schéma reproduit de [66]. (B) Compaction croissante avec la concentration de Hfq de gauche à droite de l'ADN de phage T4 (marqué par une sonde fluorescente) dans un nano-canal. (C) Courbes d'extension relative de l'ADN phage T4 à l'intérieur de canaux en fonction de la concentration de Hfq et aux concentrations de NaCl suivantes : 0mM de NaCl pour la courbe verte, 3mM pour la courbe bleue et 30mM pour la courbe rouge. La présence de sel favorise la compaction de l'ADN en présence de Hfq [19].

1.4. Localisation cellulaire et super-structure secondaire de Hfq

Comme nous l'avons précédemment souligné, Hfq est une protéine fortement présente dans la cellule bactérienne. Elle se retrouve majoritairement dans le cytoplasme, en particulier

proche de la membrane (environ 50 % de la protéine), alors que 10 à 20 % de la protéine se situent dans le nucléoïde. Le reste de la protéine est retrouvé dans le cytoplasme associé à des ARNm, sARN ou des ribosomes. Ceci a pu être établi grâce à des expériences en immunoprécipitation et par microscopie optique et électronique en effectuant un marquage de la protéine [67][58]. Comme Hfq interagit principalement avec des sARN s'hybridant avec des ARNm qui codent pour des protéines membranaires, il est probable que cette localisation à proximité de la membrane favorise ce type de régulation [68]. Il faut aussi noter que l'observation de Hfq dans le nucléoïde n'est pas aisée compte-tenu de sa forte présence à la périphérie de la cellule, masquant la localisation intra-cytoplasmique [69].

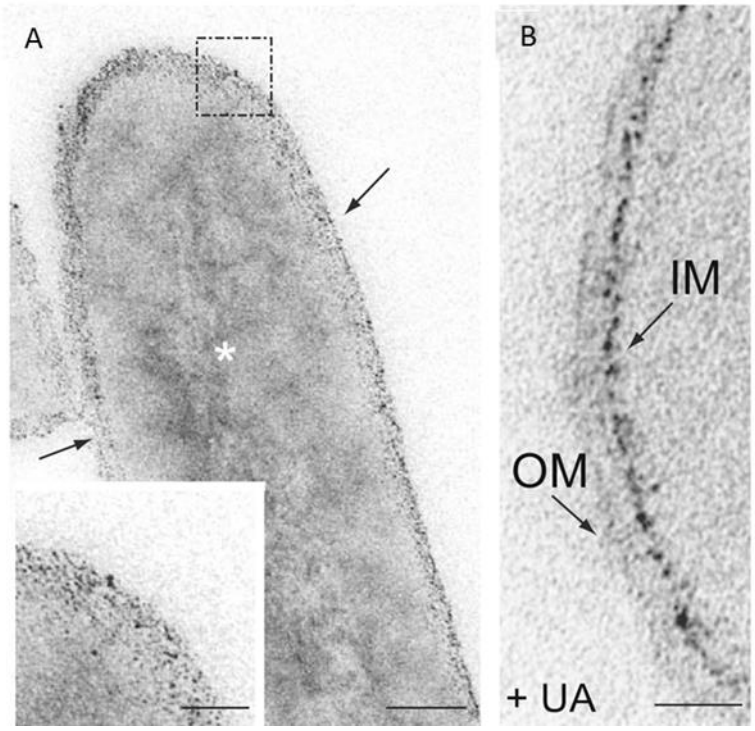


Figure 7 – Localisation de la protéine Hfq marquée par des clusters d'or en microscopie électronique à transmission. Barres d'échelle 100nm. (A) Hfq est majoritairement présente au niveau de la membrane de la bactérie. On en retrouve également une portion dans le nucléoïde (*). (B) Observation de la membrane externe (OM) et interne (IM) possible par l'ajout d'agent contrastant (acétate d'uranyle, +UA), montrant que les protéines se retrouvent au niveau de la membrane interne. Tiré de [67].

Comme indiqué plus haut, il est difficile d'observer la fraction de Hfq dans le nucléoïde en immunofluorescence. En revanche, on observe bien en microscopie optique sa répartition à la périphérie de la cellule, où elle forme de longues structures en forme d'hélices [6]. Il

convient de préciser que ces analyses sont effectuées par immunofluorescence et non en utilisant une étiquette fluorescente, puisqu'à ce jour tous les marquages réalisés en ce sens n'ont pas permis de donner une protéine fonctionnelle. La plupart des marquages aboutissent en outre à la formation de corps d'inclusion. Malgré le manque de résolution en microscopie optique conventionnelle, on soupçonne que ces structures sont sous-membranaires, et ceci a pu être confirmé par une analyse en microscopie de fluorescence par réflexion totale interne (TIRF). Cependant, les résultats restent difficilement interprétables en raison du manque de résolution de la technique. Des études en super-résolution de type PALM/STORM *(photo-activated localization microscopy and stochastic optical reconstruction microscopy*) sont envisagées, mais posent problème car, comme indiqué ci-dessus, le marquage de Hfq par une protéine fluorescente photo-activable est difficile (dû à la formation de l'hexamère). À ce jour, ces études n'ont pas encore pu aboutir.

En parallèle, une étude *in vitro* a montré que Hfq est capable de s'auto-assembler et de former des structures fibrillaires de nature amyloïde [5][6]. Les superstructures observées *in vivo* par immunofluorescence ont été rapprochées des structures observées *in vitro,* en analysant un mutant de Hfq délété de la région CTR [6]. La figure 8 présente la comparaison en immunofluorescence entre trois souches de *E. coli* cultivées dans les mêmes conditions :

(A) Une souche délétée de *hfq* dans laquelle un plasmide contenant le gène *hfq* dont la partie CTR a été tronquée (séquence de 1-72 premiers acides aminés d'Hfq présente notée NTR-Hfq) a été ajouté.

(B) Une souche WT contenant le gène *hfq* entier.

(C) Une souche délétée de *hfq* dans laquelle un plasmide a été ajouté contenant le gène *hfq* entier.

Lorsque Hfq entière est exprimée, des structures sous-membranaires sont observées. En retirant la partie CTR, ces structures sont absentes. La présence du CTR est donc nécessaire à la formation des structures *in vivo*.

Pour aller plus loin dans la compréhension de l'assemblage, ces structures ont été analysées au cours de cette thèse à l'aide de techniques de biophysique telles que la spectroscopie infrarouge IRTF, la diffusion des rayons X (SAXS) ou le dichroïsme circulaire sur synchrotron SRCD [70].

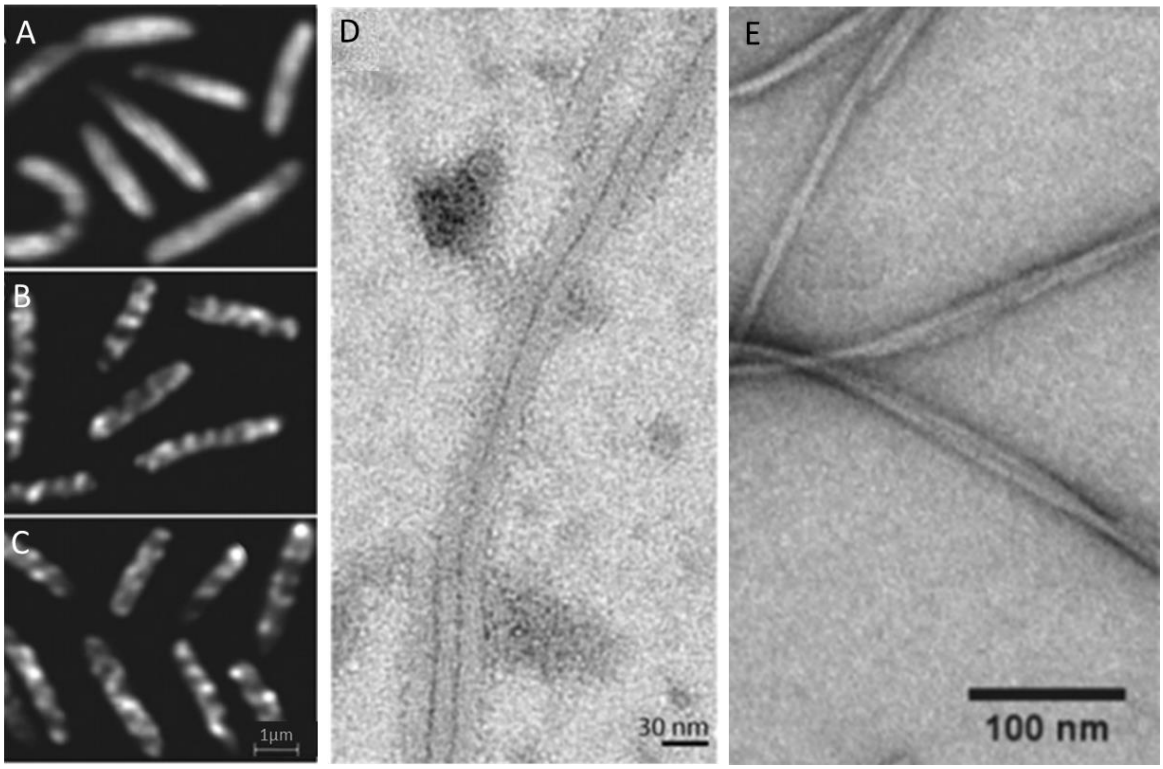


Figure 8 – Auto-assemblage de Hfq. (A), (B) et (C) indiquent la distribution de Hfq dans les cellules de *E. coli* observées par immunofluorescence. (A) Absence de structures sous-membranaires en l'absence du domaine CTR. (B) et (C) Observation d'une répartition hétérogène avec une organisation structurée hélicoïdale sous-membranaire en présence de Hfq entière. Le contrôle en l'absence de Hfq ne montre aucun marquage (non présenté ici). (D) Image de microscopie électronique de Hfq entière formant une structure fibrillaire *in vitro*. (E) Image de microscopie électronique du domaine CTR de Hfq formant une structure en forme de fibre. La nature amyloïde des fibres a pu être établie par d'autres techniques telle que la spectroscopie IRTF. Reproduit d'après [5][6].

Enfin, une analyse récente a montré que le domaine CTR permet l'interaction de Hfq avec la membrane bactérienne, ce qui explique en partie la localisation membranaire de Hfq vue en imagerie. Le domaine CTR a aussi pu être montré comme déstabilisant les membranes. L'ensemble de ces résultats permet d'attribuer de potentielles fonctions à ce domaine, comme par exemple l'export d'ARN à travers les membranes [71].

1.5. Le domaine C-terminal de Hfq

Le domaine C-terminal de Hfq n'est pas, à priori, primordial aux fonctions de riborégulation de Hfq, bien que cela soit parfois controversé. En effet, il a été montré que des versions

tronquées de Hfq possèdent encore la majorité de leurs fonctions de riborégulation qui résident donc principalement dans son tore [72]. D'autres fonctions non analysées pourraient en revanche être affectées.

Ainsi, les précédents résultats (figure 8) ont montré que le domaine CTR joue un rôle important pour son auto-assemblage. Il a aussi été mis en évidence par cinétique rapide (« stopped-flow ») que le CTR joue un rôle d'écarteur en permettant de libérer le complexe sARN-ARNm formé, rendant ainsi la protéine disponible à d'autres ARN [73].

Selon les récentes études effectuées au sein de notre laboratoire, le domaine C-terminal est aussi essentiel à la condensation et à la compaction de l'ADN. Ceci serait lié au fait que cette région s'auto-assemble et forme des structures amyloïde, en plus d'interagir avec l'ADN [6], ce qui induit un effet pontant. En effet, il a été montré que la partie NTR seule ne compacte pas l'ADN, tandis que le domaine C-terminal (seul ou au sein de la protéine) le compacte fortement [53]. Une partie de mon travail a consisté à caractériser cette dernière propriété.

2. Les amyloïdes

Comme indiqué précédemment, le domaine C-terminal de Hfq est capable de se replier sous forme de structure amyloïde. Mais ce domaine a également été décrit comme faisant partie des protéines ou sous-domaines protéiques intrinsèquement désordonnés (IDPs, *Intrinsically Disordered Proteins* en anglais). Ces protéines (ou fragments de protéines), caractérisées par un manque de structure secondaire ou tertiaire en absence de ligand ou de partenaire cellulaire, représentent environ 5% des protéines de *E. coli*. Ce pourcentage augmente chez les eucaryotes allant parfois jusqu'à 25%. Au niveau structural, elles sont dynamiques, flexibles et peuvent adopter une variété de conformations [74][75]. Elles sont quasiment impossibles à cristalliser. En termes de séquence, elles sont particulièrement riches en alanine, asparagine, glutamine, aspartate, glutamate, glycine, lysine, histidine et serine [76]. Ces acides aminés sont d'ailleurs retrouvés en grand nombre dans le domaine CTR de Hfq. De plus, il est fréquent de retrouver des répétitions dans ces régions. On remarque également cette particularité dans le cas de Hfq avec des répétions d'histidines et un motif quasiment répété (SAQNTSAQQDS). Enfin, les IDPs ont deux caractéristiques là aussi retrouvées chez Hfq : leur migration à une masse apparemment plus élevée dans un gel dénaturant de type SDS-PAGE et leur résistance à la chaleur. À noter, les IDPs ont une tendance à se convertir en structure amyloïde, comme c'est le cas pour la région CTR de Hfq.

Dans ce chapitre, nous détaillerons les propriétés structurales de ces protéines, ainsi que les inhibiteurs permettant d'affecter la formation des structures amyloïdes.

2.1. Description et caractéristiques

2.1.1. L'amylose

Les structures amyloïdes ont été découvertes car elles sont associées à des pathologies neurodégénératives, telles que la maladie d'Alzheimer ou la « maladie de la vache folle » (*eucéphalopathie spongiforme*). Le terme amylose, créé par Rudolf Virchow au début du XIXe siècle, désigne ces pathologies associées aux dépôts amyloïdes. Ces dépôts consistent en des plaques formées par l'agrégation de protéines ou de peptides qui ne nouent pas forcément de relations fonctionnelles ou structurales entre elles et sont caractérisées par un repliement anormal [77][78]. Une trentaine de protéines impliquées ont été identifiées. Quelques exemples de pathologies associées sont présentés dans le tableau 1. De manière générale, ces dépôts sont dus à un changement de conformation d'une protéine native

fonctionnelle en une protéine capable de s'auto-assembler ou de s'agréger de façon hautement organisée, ce qui aboutit à la formation de fibres. Les dépôts amyloïdes sont présents soit dans un milieu extracellulaire, soit sous forme d'inclusion intracellulaire. Les personnes atteintes de maladies neurodégénératives en sont porteuses. Lorsque les amyloses sont présentes dans un seul tissu autre que le cerveau, elles sont dites non neuropathiques localisées. Quand elles sont présentes dans de multiples tissus, elles sont dites non neuropathiques systématiques. Ces maladies peuvent être sporadiques, héréditaires ou transmissibles [79].

Pathologies	Peptides ou protéines	Structures à l'état natif
Amyloses présentes dans les pathologies neurodégénératives		
Alzheimer	Peptides β amyloïdes (Aβ)	Non structurée
Encéphalopathie spongiforme	Protéine Prion (PrP)	Partiellement en hélice α
Maladie de Parkinson	α-synucléine	Non structurée
Démence à corps de Lewy		
Maladie de Huntington	Huntingtin Polyglutamique	Non structurée
Amyloses non neuropathiques systématiques		
Amylose AL	IDP de l'immunoglobuline	Feuillet β
Amylose AA	Protéine sérum amyloïde	Hélice α
Amylose systématique sénile	Transthyrétine	Feuillet β
Amylose de l'hémodialyse	β2-microglobuline	Feuillet β
Amylose du lysozyme	Mutant du lysozyme	Feuillet β et hélice α
Amyloses non neuropathiques localisées		
Diabète de type II	Amyline/polypeptide amyloïde islet	Non structurée
Amylose au point d'injection	Insuline	Hélice-α

Tableau 1 – Quelques exemples d'amyloses et de leurs protéines associées. Tiré de [80].

2.1.2. Les amyloïdes : une voie d'agrégation des protéines

Les protéines peuvent avoir plusieurs voies et états d'agrégation possibles. Elles peuvent former plusieurs types d'agrégats dans un système vivant depuis l'étape de sa biosynthèse jusqu'à sa dégradation finale. Elles peuvent former des dépôts de « protéines natives », des dépôts amyloïdes ou bien encore des dépôts amorphes. Deux types de structures peuvent amener à l'agrégation amyloïde : les structures globulaires dans leur état natif et celles qui

sont intrinsèquement désordonnées. Selon les types de dépôts, il existe différentes séries d'états d'équilibres présentées en figure 9.

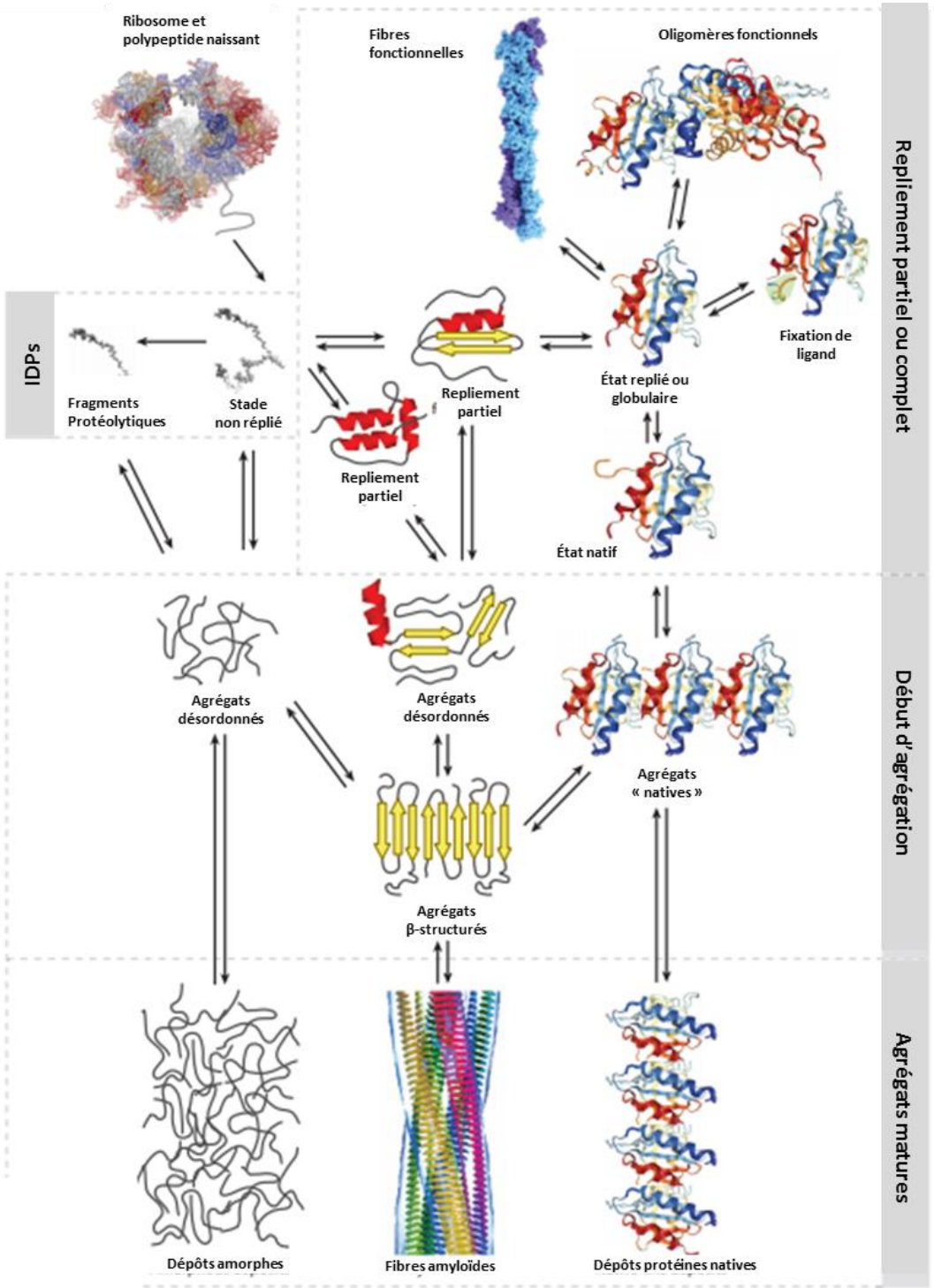


Figure 9 – Schématisation des différentes voies d'agrégation de protéines. Les agrégats peuvent former des dépôts amorphes (bas gauche), des dépôts de fibres amyloïdes (bas milieu) ou des dépôts de protéines natives. D'après [79].

2.1.3. Les structures amyloïdes : feuillets β croisés (cross-β)

Dans ce type de structure, les peptides ou protéines se replient et s'auto-assemblent en fibres formées de feuillets β croisés, partiellement ordonnés, le plus souvent insolubles, et se déposant sur les tissus. Les conditions de ce repliement sont encore mal connues. En effet, il n'existe pas de signature commune entre les structures natives de ces protéines que ce soit au niveau de la taille et de la composition de la séquence ou de la structure tridimensionnelle. *A contrario*, les fibres amyloïdes présentent de nombreuses similitudes. Elles partagent une morphologie commune. L'agglomération se fait en couches successives de protéines linéaires, non ramifiées et riches en feuillets β, selon un axe perpendiculaire à la direction des feuillets où les liaisons de type hydrogène sont perpendiculaires.

Des études structurales à haute résolution par microscopie électronique à transmission (MET), par spectroscopie à résonance magnétique nucléaire (RMN) du solide, ainsi que des modélisations moléculaires ont permis d'élaborer des modèles possibles d'auto-assemblage comme par exemple celui du peptide $A\beta_{40,}$ le peptide amyloïde le plus étudié (présenté en figure 10). L'auto-assemblage formerait des paires de brins ainsi qu'une boucle perpendiculaire à l'axe de la fibre par l'intermédiaire de liaisons de type hydrogène pour constituer des feuillets β croisés, parallèles et colinéaires à l'axe des fibres. Le diamètre de la fibre peut aussi augmenter par empilement des brins via leur face hydrophobe. Il peut prendre deux symétries probables dans le cas d'$A\beta_{40}$ pour constituer le protofilament (modèle proposé par Tycko, figure 10). D'autres conformations de ce repliement sont possibles comme la présence de trois molécules/peptides par couche.

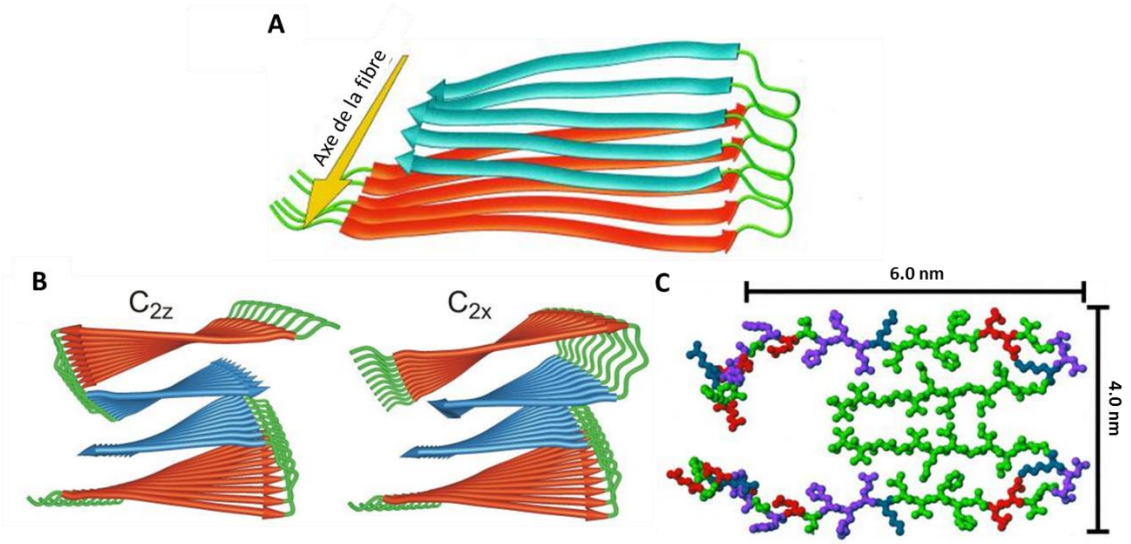


Figure 10 – Modèle structural de la fibre amyloïde et des feuillets β croisés, formés par le peptide $A\beta_{40}$ (proposé par Tycko). (A) Modèle des feuillets β en forme de rubans (en haut). (B) Modèles des arrangements des feuillets β entre deux molécules dans une symétrie C_{2z} et C_{2x}. (C) Modèle atomique de la structure C_{2x} (en vert : hydrophobe ; en magenta : polaire ; en rouge : chargé négativement ; en bleu : chargé positivement). Copyright (2002) *National Academy of Sciences*, U.S.A pour A et C [81][82].

Des structures atomiques d'amyloïdes d'autres peptides et de protéines ont été proposées, comme la structure de la partie C-terminale du prion (*proteinaceous infectious particles*) ou de la protéine Amyline. La conformation est commune à l'ensemble des amyloïdes mais avec des arrangements variables (figure 11). Les similitudes dans la formation des fibres peuvent être attribuées à la chaîne carbonée commune à tous les polypeptides. Les différents arrangements dépendent des différences de longueurs et de composés de la chaîne polypeptidique et des conditions physico-chimiques lors de d'agrégation [83].

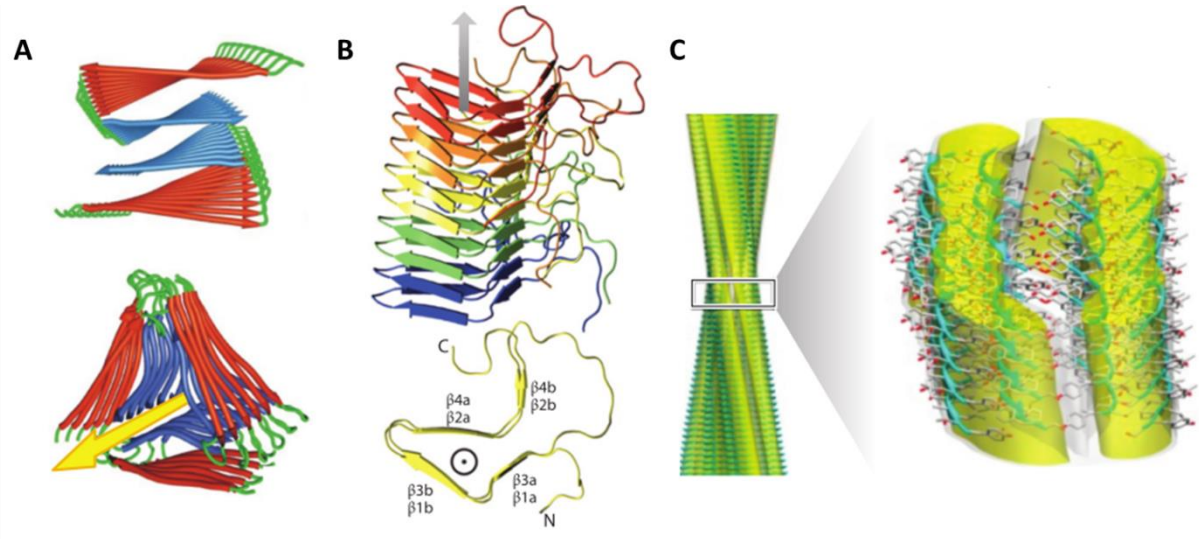


Figure 11 – Structures tridimensionnelles de fibres amyloïdes provenant de différentes sources. (A) Deux modèles structuraux polymorphes de protofilaments $A\beta_{40}$ observés sur l'axe longitudinal de la fibrille (représentés par des modèles de rubans). (B) Fibre formée par le domaine prion du HET de Podospora Anserina. (C) La structure de fibre à six protofilaments d'un peptide correspondant aux résidus 105-115 de la transthyrétine. Adapté de [79].

2.1.4. Dynamique et mécanisme de formation des amyloïdes

Les réactions d'auto-assemblage nucléées, y compris la formation de fibres amyloïdes, présentent typiquement une cinétique de croissance selon une courbe sigmoïde (figure 13) [84]. En effet, il est généralement admis que la formation de fibres amyloïdes adopte un mécanisme de type nucléation/propagation, commençant par une phase de latence correspondant à la nucléation, suivie d'une propagation qui est exponentielle, conduisant à la formation de gros agrégats/fibres pour arriver à un plateau, ou équilibre [85]. Le début et la fin de la transition peuvent être plus ou moins nets selon le mécanisme. Cependant, il est important de noter qu'aucune de ces trois phases ne peut être attribuée à un seul mécanisme ou processus moléculaire simple. Le groupement qui se forme au début de la phase d'agrégation garde souvent sa conformation de départ et s'agrège de ce fait entre protéines ou peptides non structurés ou à partir de la structure native. La formation d'oligomères est peu stable et est en équilibre avec l'état fonctionnel car elle implique des interactions de faible énergie. Cet équilibre « métastable » rend difficile l'étude expérimentale des oligomères qui sont aujourd'hui reconnus comme en partie responsables de la cytotoxicité, en raison de leurs fortes interactions et à leur capacité de pénétrer et déstabiliser les

membranes. Les premiers oligomères formés se réorganisent en des espèces plus stables. La compaction ainsi que le nombre et la taille des feuillets β augmentent au cours du processus de formation des fibres amyloïdes. Ensuite, les fibres peuvent incorporer de nouveaux monomères ou de nouvelles fibres pour s'allonger ou former des super-fibres. Une fois les fibres formées, un second processus peut se mettre en place afin de former et catalyser de nouveaux agrégats ou fibres. Ces processus sont la fragmentation ou la nucléation secondaire (figure 12) [86].

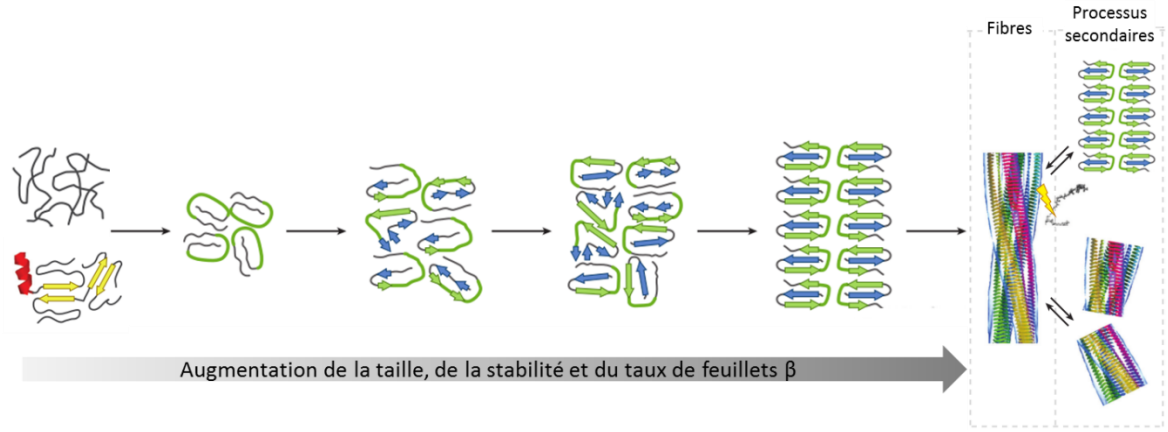


Figure 12 – Formation progressive et processus secondaire de fibres amyloïdes. Schématisation de l'oligomérisation, des réarrangements des segments amyloïdogéniques initiaux (en vert) et de l'augmentation de paramètres structuraux (en bleu) prenant place durant le processus. D'après [79].

D'un point de vue thermodynamique, la plupart des interactions secondaires entre protéines sont intramoléculaires, alors que pour l'état amyloïde les interactions sont intermoléculaires. L'état natif est thermodynamiquement plus stable que l'état amyloïde lorsque l'énergie libre (G) du peptide ou de la protéine est plus faible à l'état natif qu'à l'état amyloïde. La stabilité de l'état amyloïde (ΔG) dépend de la concentration de protéine en solution, tandis que celle de l'état natif est indépendante de la concentration (sauf s'il existe dans un complexe fonctionnel). À la concentration dite « critique », la stabilité de l'état amyloïde est égale à celle de l'état natif. À des concentrations dépassant cette valeur critique, une protéine est plus stable à l'état amyloïde qu'à son état natif. Dans de telles situations, l'état natif ne peut persister que s'il existe de hautes barrières d'énergie libre qui entravent la transition vers l'état amyloïde plus stable. La transition vers l'état amyloïde n'aura lieu que si son énergie libre est inférieure à celle de l'état natif (figure 13).

Alors que la thermodynamique décrit si une transition d'un état à un autre est spontanée ou peut se produire sans énergie d'activation externe, la cinétique aborde la question de la rapidité de cette transformation. Les mesures cinétiques ont toujours été l'un des principaux outils permettant d'élucider les mécanismes de réaction. Les mesures cinétiques expérimentales permettent de tester la vitesse relative des différents processus [80].

Concernant les forces d'interaction mises en jeu dans l'agrégation des feuillets β amyloïdes, les brins des feuillets se forment grâce à des liaisons de type hydrogène, et se trouvent tout le long de l'axe de la fibre, tandis que l'interaction entre les feuillets β dans la direction perpendiculaire à l'axe des fibres dépend fortement de la nature des chaînes latérales et de leurs caractéristiques (taille, charges...) [80].

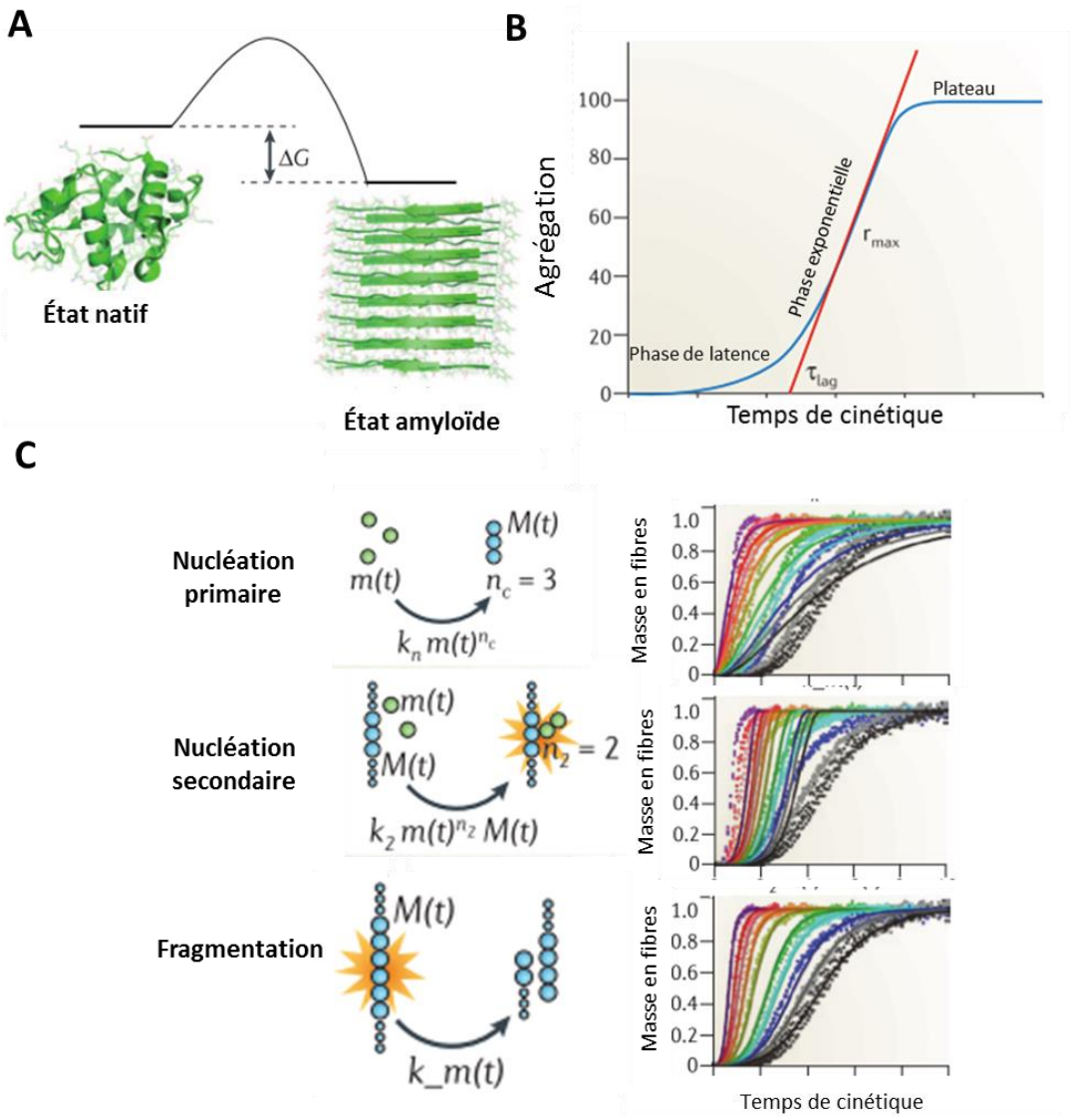


Figure 13 – Thermodynamique et cinétique de formation de fibres amyloïdes. (A) Diagramme d'énergie libre de la formation de fibres amyloïdes. Les fibres et les monomères peuvent avoir une énergie libre similaire, et la concentration totale de monomères gouverne ΔG. (B) Courbe d'agrégation cinétique caractéristique de formation de fibres amyloïdes : concentration de masse agrégée en fonction du temps. La courbe présente typiquement trois parties : une phase de latence, une phase de croissance et un plateau final. (C) Les cinétiques d'agrégation du peptide Aβ42 à des concentrations croissantes sont comparées à des lois de vitesse qui conviennent aux différents processus de nucléation primaire, de fragmentation et de nucléation secondaire. Adapté et traduit avec autorisation de Spinger Nature [80][87].

2.1.5. Caractéristiques physico-chimiques des fibres amyloïdes

Les fibres amyloïdes sont caractérisées par plusieurs propriétés morphologiques et physico-chimiques communes et ce, malgré leur séquence et structure natives différentes. Elles font typiquement une dizaine de nanomètres de diamètre et peuvent mesurer plusieurs micromètres de long. En outre, elles peuvent avoir des diamètres bien plus importants [88], grâce à la formation de superenroulements de filaments ou de fibres tournant les uns sur autres en forme de torsades ou associés en forme de rubans superposés.

Du point de vue de leur résistance mécanique, il est à noter que les fibres amyloïdes ont une élasticité et une force de rupture comparables à celles de l'acier [89]. Elles sont par ailleurs résistantes aux protéases ainsi qu'aux détergents.

Ces structures peuvent être observées en microscopie électronique à transmission (ou MET) ou en microscopie de force atomique (AFM) [90]. La distance répétée entre les feuillets et entre les brins ainsi qu'un même axe d'orientation des fibres permettent d'observer un motif typique par diffraction aux rayons X [91]. En spectroscopie infrarouge, les fibres amyloïdes ont une réponse spécifique caractéristique de l'angle de torsion de la liaison peptidique imposé par l'agrégation [92]. Elles ont également des propriétés tinctoriales, interagissant avec des sondes fluorescentes. L'interaction avec le rouge Congo déplace son maximum d'absorption de 500nm à 540nm, avec une apparition d'une biréfringence verte en microscope à lumière polarisée [93]. L'interaction avec la Thioflavine T permet d'observer une émission de fluorescence à 480nm après excitation à 440nm [94].

Ces méthodes seront présentées plus en détail dans la partie « Matériel et méthode ».

Ces trois caractéristiques que sont la morphologie, la structure et l'interaction avec des sondes/molécules permettent de spécifier l'agrégation de type amyloïde [95].

2.1.6. Cas particulier : les prions

Le terme « prion » a été créé en 1982 par Stanley Prusiner pour *PROteinaceous INfectious particle* (particule protéique infectieuse). Il fait partie de la classe de protéines formant des structures amyloïdes. Comme la plupart des protéines amyloïdogéniques, les séquences sont intrinsèquement désordonnées mais la séquence des prions présente une proportion particulièrement élevée de glutamine et d'asparagine [96]. Les prions ont pour caractéristique de favoriser la polymérisation des protéines natives en fibres amyloïdes. Elles peuvent aussi

parfois alterner entre une conformation soluble et amyloïde et vice versa [97]. Les maladies à prions sont généralement transmissibles, contrairement aux maladies associées à la présence d'amyloïdes. On les retrouve dans les maladies neurodégénératives mortelles chez les mammifères, telles que « la tremblante » du mouton, l'encéphalopathie spongiforme bovine ou encore la maladie de Creutzfeldt-Jacob chez l'homme. Cependant, il a été montré que des protéines semblables aux prions ont plutôt une fonction biologique, comme par exemple chez les levures [98].

2.2. Concept d'amyloïdes fonctionnelles

2.2.1. Concernant la fonction des IDPs

Les IDPs ne sont pas dépourvues de fonctions : régulation de la transcription et de la traduction mais aussi l'auto-assemblage de complexes protéiques, transmission de signaux, stockage de petites molécules, maintien de l'interaction entre partenaires protéiques [99] [100]. Ces fonctions sont généralement possibles grâce à la capacité des IDPs à effectuer des transitions vers des états de repliement plus ordonnés et stables lorsqu'ils interagissent, par un processus de couple dit de « repliement et interaction ». Ce processus est une compensation entropie-enthalpie, avec une énergie entropique générée par le repliement et compensée par l'énergie enthalpique de la liaison. Les IDPs sont susceptibles de former des amyloïdes présentant des capacités fonctionnelles (Figure 9).

2.2.2. Les amyloïdes fonctionnelles

Bien que l'amyloïde ait généralement été considérée comme une structure pathologique, des recherches ont montré que les amyloïdes peuvent également être fonctionnelles, utiles et contribuer à la physiologie normale des cellules. Les amyloïdes fonctionnelles ont été découvertes dans de nombreux organismes, allant des bactéries aux mammifères, y compris chez l'homme [101]. Elles sont chargées de différents rôles comme la formation de biofilms, le développement de structures, la régulation de synthèse, le transfert de l'information… Ces systèmes nécessitent cependant un contrôle minutieux de leur assemblage pour éviter les effets potentiellement toxiques. D'après C. Maury, les organismes auraient évolué en tirant parti de la capacité des protéines à former des amyloïdes. En effet, certaines de leurs propriétés telles que l'avantage d'avoir une forte résistance à la protéolyse, une solidité comparable à celle de l'acier, des caractéristiques d'auto-propagation auraient potentiellement un rôle dans la transmission d'informations, ceci permettant leur évolution.

D’après ce même auteur, les amyloïdes seraient potentiellement à la base de la complexité de l’organisation et de la compartimentation à l’origine de la vie [102].

La vie proviendrait d'un mécanisme complexe entre un ensemble de produits chimiques divers, parmi lesquels on retrouverait les acides nucléiques, les lipides et les peptides. Comprendre l'origine de la vie ne reposerait pas seulement sur une explication de la façon dont ces molécules biologiques sont apparues, mais aussi sur une explication de leur relation structurelle, chimique et thermodynamique. L’hypothèse « d’un monde amyloïdes » gagne du terrain. Selon cette hypothèse, les peptides amyloïdogéniques seraient les précurseurs chimiques de la vie : premières molécules réplicatives et catalytiques. Cette idée amène une nouvelle hypothèse de l’origine de la vie, avec la plus populaire et plus largement acceptée reposant sur l’ARN. Cette idée est soutenue par l’expérimentation. En effet, les amyloïdes peuvent être formées à partir de séquences peptidiques simples et courtes et dans des conditions prébiotiques. Elles sont plus stables et plus résistantes que leurs composants peptidiques isolés. Leurs structures répétitives leur permettent d’interagir avec d’autres molécules telles que les acides nucléiques, les membranes lipidiques et les polysaccharides. Elles peuvent se former spontanément lors de condensation prébiotique contenant des acides aminés et ont différentes activités catalytiques. Toutes ces propriétés en font un bon candidat en tant qu’entité prébiotique à l’origine de la vie [103].

2.2.3. Interaction amyloïde-membrane et amyloïde-ADN

Les amyloïdes ont aussi une capacité intrinsèque à interagir avec les structures répétitives telles que les membranes et l'ADN (plus généralement les acides nucléiques). L’interaction des amyloïdes avec des membranes lipidiques a fait l'objet de nombreuses recherches, car le dépôt de fibrilles à leur niveau est associé à une cytotoxicité impliquée dans de nombreuses maladies humaines. Mais on ne peut pas exclure que cette interaction pourrait aussi avoir des fonctions bénéfiques pour la cellule (export de molécules). Ces interactions se sont avérées à la fois coopératives et perturbatrices : les membranes peuvent favoriser la croissance des fibrilles, mais les fibrilles peuvent déformer, affecter la fluidité ou même perforer les membranes. Par exemple, il a été montré que la protéine Aβ peut se lier aux membranes et former des feuillets β antiparallèles orientés parallèlement à la surface de la membrane [104]. Cette interaction peut amener à un changement de la fluidité membranaire [105]. Une étude récente montre que les amphiphiles sont capables de déclencher l'agrégation d'un peptide soluble [106]. Cette interaction amyloïde-membrane semble être

spontanée. Une autre forme d'interaction est de plus en plus couramment rapportée, celle qui a lieu entre les amyloïdes et l'ADN. En effet, l'ADN favorise la formation d'amyloïdes pour les peptides et protéines de maladies neuro-dégénératives, mais aussi pour des peptides qui n'ont pas ces caractères intrinsèques [107]. Ces interactions sont de plus en plus considérées comme un caractère physiologique et fonctionnel, et que l'on peut aussi retrouver dans les bactéries. Le paragraphe suivant présentera certains exemples d'amyloïdes bactériens.

2.2.4. Exemples d'amyloïdes bactériens

L'exemple de fibres amyloïdes fonctionnelles le plus étudié chez la bactérie est celui des curli et des pili. Les curli sont présents chez des bactéries *E. coli* et *Salmonelle* et sont composés de deux sous-unités principales, CsgA et CsgB [108]. Les curli sont impliqués dans la formation de biofilm ainsi que dans l'adhésion et l'invasion des cellules hôtes. La nature amyloïde des curli est due à la protéine CsgA. Sa nucléation démarre lorsqu'elle est en contact avec son partenaire CsgB. CsgA et CsgB ont quasiment la même séquence qui se répète, mais CsgB contient une partie chargée positivement permettant l'ancrage à la surface/membrane externe sur CsgF (ce dernier étant lié à CsgG). Leurs sécrétions sur la membrane externe sont possibles par un pore généré par CsgG. Ce processus hautement régulé est aidé par plusieurs partenaires impliquant plus de six protéines transcrites par deux opérons. Par exemple, CsgC intervient en retardant la fibrillation amyloïde de CsgA en perturbant sa structuration dans le périplasme, ce qui permettrait de ne pas être auto-toxique ou «auto -destructeur » pour la bactérie [109]. CsgE et CsgH ont aussi été identifiées comme inhibitrices d'amyloïdes mais moins efficaces que CsgC ceci étant dû notamment à leurs structure et caractéristique (charge en surface) différentes, mais ont une structure tertiaire similaire. Cette formation est régie par plusieurs conditions environnementales, notamment la température et la disponibilité de nutriments et de sels [101]. Un système non apparenté génétiquement, mais qui est morphologiquement similaire et qui présente une similitude dans les séquences par rapport au système curli a aussi été identifié et décrit comme étant une amyloïde fonctionnelle chez *Pseudomonas*, appelée Fap [27].

Mais d'autres fonctions ont aussi pu être identifiées pour des protéines amyloïdes bactériennes telles que chez *Bacillus subtilis*, la protéine TasA. Il existe en effet un opéron qui code pour 3 protéines régulant la formation de biofilms chez *subtilis* (protéines SipW, TapA et TasA). Le biofilm formé par *B. subtilis* est principalement constitué de

polysaccharide et de TasA, qui forment un réseau de fibres amyloïdes entourant la surface cellulaire [110]. TasA permet de fournir un soutien structural au biofilm. Les protéines amyloïdes peuvent également être utilisées comme facteur de virulence : chez *Xanthomonas* par exemple, le facteur de virulence HpaG est sécrété et son mécanisme probable est dû à la rupture de la membrane hôte de la plante [111]. Chez *Klebsiella pneumoniae*, la microcine E492 forme des pores dans la membrane cytoplasmique de l'hôte [112]. Enfin, la protéine Rho de *Clostridium botulinum*, dont la fonction est de favoriser la terminaison de la transcription [113]. Ainsi, il existe de nombreux autres exemples qui émergent et la liste des fonctions biologiques pour les amyloïdes bactériennes s'allonge, celles-ci étant notamment liées à des interactions avec les acides nucléiques (ADN/ARN) et aux membranes, comme dans le cas de Hfq [53][71].

Pour utiliser efficacement l'amyloïde, comme celui de type curli, il est essentiel que les organismes possèdent des mécanismes qui contrôlent le processus de formation. Les études sur les amyloïdes fonctionnelles permettraient de mieux comprendre le contrôle de leur formation, ce qui pourrait aider au traitement des maladies à amyloïdose et à la mise au point de nouveaux biomatériaux. Par exemple, CsgC et CsgH se sont révélés capables d'inhiber efficacement l'assemblage d'autres amyloïdes, notamment l'α-synucléine [114].

2.3. Inhibiteurs d'amyloïdes

2.3.1. Introduction : inhibiteurs d'amyloïdes

Afin de trouver un traitement ou voie thérapeutique contre les maladies neurodégénératives liées à la formation de fibrilles amyloïdes, les inhibiteurs d'agrégation d'amyloïdes sont intensément étudiés. Ces recherches ont pour objectif d'empêcher la formation ou détruire les dépôts en place. Ce sont généralement de petites molécules pouvant être de différentes natures: peptides [115], polyphénols, inositols, quinones et leurs dérivés, et molécules chélatrices de métaux. Ces molécules ont été étudiées en utilisant des modèles *in vitro*, *in vivo*, mais aussi des simulations bio-informatiques pour comprendre leur mécanisme d'action ainsi que les caractéristiques structurales communes sous-jacentes au processus d'inhibition. Les molécules inhibitrices étudiées possèdent en général un noyau aromatique qui joue le rôle d'unité de reconnaissance des amyloïdes et un caractère hydrophile qui va agir sur la structure amyloïde afin de les rompre ou bien d'empêcher sa formation. Ce choix de

molécules se justifie par la présence d'acides aminés aromatiques dans les séquences de la plupart des peptides/protéines amyloïdogéniques (Aβ, PrP, Amyline…) [116].

Le mode d'action des molécules inhibitrices ne vise pas forcément la destruction totale des agrégats amyloïdes et peut impliquer la conservation d'une partie de ces structures amyloïdes, afin de piéger et éviter la formation des oligomères amyloïdes toxiques, ou de stabiliser des espèces amyloïdes non toxiques. Il peut réorienter les voies ou modifier la cinétique d'agrégation [117].

2.3.2. Les différents types d'inhibiteurs

Les molécules capables d'interagir avec les fibres amyloïdes peuvent être de différentes natures. Ainsi on pourra différencier :

Les peptides: plus sélectifs, moins chers et plus efficaces que les protéines dites thérapeutiques, les peptides ont été largement utilisés en médecine et en biotechnologie au cours des deux dernières décennies [118]. Les stratégies ont principalement pour but d'empêcher la constitution de fibres en formant par exemple des peptides appelés β « disjoncteurs » qui présentent une séquence comportant des prolines (qui interférent avec l'agrégation) et un groupe sélectif ciblant les amyloïdes [119]. Certains peptides de par leurs charges positives et leur « rigidité » peuvent inhiber la formation de fibres [120]. La caractéristique structurale importante des peptides inhibiteurs est la présence d'une unité de ciblage de l'amyloïde et d'une unité « gêneur ».

Les polyphénols : dans le cas des polyphénols, la taille de la molécule et les fonctions chimiques présentes (groupe hydroxyle par exemple) influencent l'efficacité inhibitrice. Des études ont montré que le groupe aromatique ne suffisait pas et que la présence de la fonction phénol (fonction interagissant avec l'amyloïde) est primordiale. La taille de la molécule séparant deux couches doit faire entre 8 et 16 Å [121]. Le curcumine est un exemple de polyphénol capable de dissocier les fibres amyloïdes et d'empêcher leur formation dans le cas du peptide Aβ [122]. La force d'inhibition viendrait de sa structure compacte et symétrique. Le rouge Congo, colorant des amyloïdes, fait partie de cette famille de molécules. Le gallate épigallocatéchine (EGCG) est aussi une petite molécule polyphénolique présente dans le thé vert qui a beaucoup été étudiée pour son effet d'inhibition sur différentes amyloïdes [123]. L'Apomorphine (APOM) [124], l'acide

tannique, le resvaretrol et les dérivés de EGCG sont tous des polyphénols ayant des effets inhibiteurs [117].

Enfin, d'autres molécules telles que les inositols, représentent une classe de huit stéréoisomères naturels de cyclohexylpolyols, dont trois inhibent la formation d'amyloïdes Aβ et comprennent les myo-, epi- et scyllo-inositols [125]. Les alcools fluorés, les acides fluorés et les nanoparticules fluorées [126] comme l'ester d'acide indolyl-trifluorométhyl-hydroxypropionique et ses dérivés ont significativement inhibé l'agrégation du peptide Aβ et désagrégé les amyloïdes *in vitro* [127]. Les quinones et leurs dérivés (produits oxydés des polyphénols comme par exemple les benzoquinones, les naphtoquinones, les anthraquinones et les phenanthraquinones), présentent d'excellentes propriétés inhibitrices de l'agrégation de diverses protéines amyloïdogéniques [128]. Les molécules de quinone interagissent avec la protéine amyloïdogénique par des liaisons de type hydrogène, des interactions aromatiques et par la forte interaction entre l'anneau central de quinone déficient en électrons, dipôle asymétrique du cycle quinone [117].

Matériel et méthode :

Panorama des techniques utilisées pour l'étude des amyloïdes.

Les techniques spectroscopiques, utilisant l'interaction onde-matière, sont des outils puissants pour sonder la structure 3D des biomolécules. L'interaction onde-matière peut donner lieu à de nombreux phénomènes tels que l'absorption, la réflexion, la diffusion élastique ou inélastique et la diffraction. Ils fournissent une signature spécifique à chaque molécule. Suivant l'énergie de l'onde incidente absorbée ou émise, nous avons accès à différentes informations sur la composition chimique, la structure énergétique, ou la conformation spatiale de la molécule.

De nombreuses techniques spectroscopiques ont été utilisées pour analyser, identifier et caractériser l'auto-assemblage de peptides sous forme de structures amyloïdes. Notamment, nous avons recouru à la spectroscopie infrarouge, ainsi qu'à des techniques utilisant le rayonnement ultraviolet, dans le cas de la spectroscopie de fluorescence et de dichroïsme circulaire. Les rayons X ont aussi été utilisés dans le cas de techniques de diffusion et de diffraction.

Dans ce chapitre, nous rappellerons les notions essentielles de chaque technique utilisée, principalement la spectroscopie infrarouge et le dichroïsme circulaire pour l'étude des structures protéiques. Nous allons aussi décrire certaines techniques d'imageries comme le microscope à force atomique et la microscopie électronique.

1. Introduction - Interaction lumière/matière

Le principe de la dualité onde-corpuscule a été décrit pour la première fois par Albert Einstein en 1905. Ce principe stipule que la lumière est définie à la fois comme une onde et un corpuscule appelé photon. Le photon est porteur d'un « quanta » d'énergie E(J) qui est donné dans le vide par la relation suivante :

$$E = h\nu = h\frac{c}{\lambda} = hc\bar{\nu}$$

Avec : h la constante de Planck (6,626.10^{-34} J.s) et c la vitesse de la lumière (2,998.10^{8} m.s^{-1}) dans le vide. ν (s^{-1}), λ(m) et $\bar{\nu}$(m^{-1}) sont respectivement la fréquence, la longueur et le nombre d'onde du rayonnement, permettant de caractériser les radiations électromagnétiques.

La nature de la lumière (X, UV, visible, IR…) dépend de l'énergie et se caractérise par les valeurs ν, λ ou $\bar{\nu}$ de l'onde électromagnétique. Le spectre électromagnétique (figure 14) présente la décomposition de la lumière en différents domaines d'énergie. On retrouve pour

les faibles énergies les ondes radio, et pour les plus fortes les rayons γ, passant en énergie croissante par l'infrarouge, le visible, l'ultraviolet et les rayons X.

Chacun de ces domaines est associé à des techniques de spectroscopie mobilisant différents types d'interaction (ou d'excitation) avec les molécules étudiées. Pour des raisons historiques et pratiques, les unités utilisées pour décrire l'énergie de l'onde électromagnétique sont différentes selon les techniques. C'est pour cette raison que l'on utilise le nombre d'onde (cm^{-1}) pour l'infrarouge, la longueur d'onde en nanomètre (nm) pour l'Ultraviolet-Visible et les valeurs d'énergie (eV) pour les rayons X.

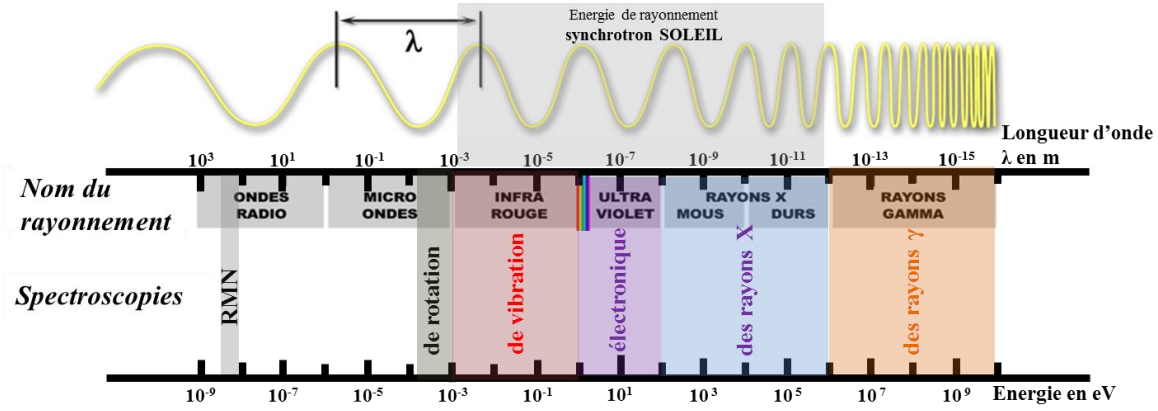


Figure 14 – Domaines d'énergie et spectroscopies associées au spectre électromagnétique.

Lorsqu'une onde électromagnétique interagit avec la matière, elle excite les molécules. Ce phénomène est responsable de l'absorption de photons, parfois aussi de leur émission. La mécanique quantique permet d'expliquer ce qu'il se passe lors de ces interactions. En effet, l'énergie de la matière est quantifiée et, sous l'influence d'un champ électromagnétique, la transition d'électrons entre niveaux électroniques, vibrationnels ou rotationnels est possible. On considère par approximation que l'énergie quantique E_{qu} d'une molécule est la somme de trois énergies. Elle s'écrit :

$$E_{qu} = E_{élec} + E_{vib} + E_{rot}$$

Avec $E_{élec}$ correspondant à l'énergie électronique dont le nombre quantique associé est n. La variation d'énergie électronique $\Delta E_{élec}$ entre un niveau n et n+1 est de l'ordre d'une dizaine d'électronvolts (eV) (~35000cm^{-1}).

Avec E_{vib} correspondant à l'énergie électronique dont le nombre quantique associé est v. La variation d'énergie électronique ΔE_{vib} entre un niveau v et v+1 est de l'ordre du dixième d'eV (~1000cm^{-1}).

Avec E_{rot} correspondant à l'énergie électronique dont le nombre quantique associé est j. La variation d'énergie électronique ΔE_{rot} entre un niveau j et j+1 est de l'ordre du millième d'eV (~30cm^{-1}).

À partir d'une certaine énergie, l'onde électromagnétique peut être ionisante. Les UV et les énergies de photon supérieures ont tendance à être ionisantes et sont donc souvent « destructives ». L'infrarouge à l'inverse est « non destructif » car le rayonnement est de « faible » énergie.

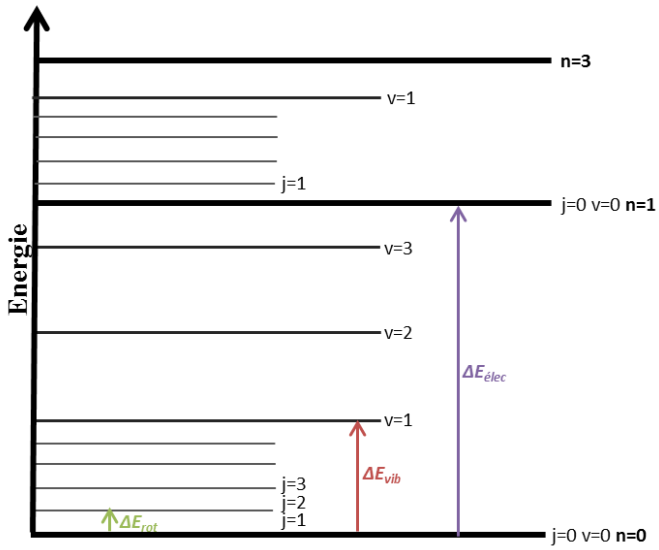


Figure 15 – Diagramme des différents niveaux d'énergie dans le cas d'une molécule diatomique. Représentation des niveaux électroniques (n), vibrationnels (v) et rotationnels (j). Échelle non respectée. (Adapté du diagramme de Jablonski-Perrin).

La Transmission « T » correspond au rapport de l'intensité de photons transmis après le passage du faisceau à travers l'échantillon sur l'intensité incidente. T s'écrit sous la forme :

$$T\ (\bar{\nu}) = \ exp\ [-\alpha(\bar{\nu}).\ l] = \frac{I(\bar{\nu})}{I_0(\bar{\nu})}$$

Le rayonnement subit une décroissance exponentielle. Cette décroissance va être plus ou moins importante en fonction de la nature et de l'épaisseur l de l'échantillon que traverse le

faisceau. $\alpha(\bar{\nu}) = 4\pi k(\bar{\nu})\bar{\nu}$ le coefficient d'absorption, est relié à $k(\bar{\nu})$ le coefficient d'extinction de l'échantillon, $k(\bar{\nu})$ étant la partie imaginaire de l'indice de réfraction complexe de l'échantillon. Le coefficient d'absorption $\alpha(\bar{\nu})$ peut s'écrire, pour une solution, de la manière suivante :

$$\alpha(\bar{\nu}) = \varepsilon(\bar{\nu}).C$$

Avec *C* la concentration de la solution et $\varepsilon(\bar{\nu})$ le coefficient d'extinction molaire de la molécule absorbante en fonction de la fréquence.

L'absorbance *A* (appelée aussi densité optique DO) est plus souvent utilisée car elle est directement proportionnelle en solution à la concentration *C* d'un soluté ou au trajet optique l. Cette loi correspond à la relation de Beer-Lambert suivante :

$$A = -\log T = \varepsilon(\bar{\nu}).C.l$$

L'absorbance mesurées est caractéristique de l'élément ou de l'échantillon étudié. C'est pour cette raison qu'on a recourt aux techniques de spectroscopie. Nous allons voir en particulier dans cette partie les principes et applications de la spectroscopie infrarouge et de la spectroscopie de dichroïsme circulaire. Ces deux techniques apportent des informations précises sur la nature et la structure de molécules notamment biologiques.

2. La spectroscopie infrarouge

2.1. Introduction à la spectroscopie infrarouge

L'origine de l'absorption du rayonnement infrarouge (IR) se trouve dans les vibrations des liaisons entre les atomes. L'infrarouge est divisé en trois domaines d'énergie : le plus énergétique appelé « le proche infrarouge » qui correspond à la gamme spectrale en longueur d'onde comprise entre 0,7 et 2,5µm, « le moyen infrarouge » compris entre 2,5 et 25µm et l'infrarouge lointain, le moins énergétique, de 25 à 1000µm. C'est en général la gamme de l'IR moyen qui est utilisée pour analyser les vibrations des molécules. En effet, la gamme présente une énergie proche de la différence des niveaux d'énergie moléculaire vibrationnelle (ΔE_{vib}).

2.1.1. L'absorption en infrarouge

Une molécule absorbe en infrarouge si elle possède un moment dipolaire μ et la valeur de l'absorption sera proportionnelle à la variation de l'intensité de μ. Les molécules symétriques n'ont pas de moment dipolaire permanent et ne présentent donc généralement pas d'absorption en infrarouge. L'absorption infrarouge est fortement liée à la symétrie des molécules. Une approximation simple permet de déterminer vers quelle fréquence de vibration *v* une liaison chimique d'une molécule diatomique A-B absorbe. Il faut considérer que la liaison chimique entre les deux atomes A et B se comporte comme un oscillateur harmonique idéal liant deux masses m_A et m_B (de façon respective) à ses extrémités. Cette approximation s'écrit sous la relation suivante :

$$\nu = \frac{1}{2\pi}\sqrt{k/m^*}$$

Où $m^* = m_A . m_B /(m_A + m_B)$ correspond à la masse réduite de l'oscillateur et *k* à la constante de raideur du ressort/liaison chimique. Cette relation permet de comprendre simplement certains effets et phénomènes mis en jeu par l'absorption d'un rayonnement infrarouge d'une molécule, sur la nature de la liaison chimique. Par exemple, la fréquence de vibration d'une liaison simple sera environ $\sqrt{2}$ fois plus élevée que la fréquence de vibration d'une liaison double entre deux atomes. C'est en effet le cas pour la vibration de la liaison éthane comparée à celle de la liaison éthylène. D'autres caractéristiques telles que la masse des atomes impliqués peuvent être établies. En effet, plus les atomes sont lourds, plus l'absorption se trouvera vers les basses fréquences. Le marquage isotopique est d'ailleurs souvent utilisé en spectroscopie IR pour identifier des modes vibrationnels.

Du point de vue de la mécanique quantique, E_{vib}, l'énergie vibrationnelle dépend de la fréquence propre ν ci-dessous et est quantifiée d'un nombre quantique de vibration *v* (*v* nombre entier positif). *E* s'écrit sous la forme suivante :

$$E_{vib} = h\nu\ (v + \frac{1}{2})$$

Il existe également une règle de sélection quantique qui stipule que les transitions vibrationnelles possibles sont entre niveaux ayant une différence du nombre quantique v : $\Delta v = \pm 1$.

L'absorption a lieu si le produit scalaire du rayonnement infrarouge (champ électromagnétique) et du moment dipolaire électrique de la molécule est non nul. Toutes les liaisons chimiques qui vibrent n'induisent pas de modification dipolaire électrique lors du passage d'un rayonnement infrarouge incident. Exemples : CH_4, CO_2, H_2 et O_2 sont toutes des molécules qui n'ont pas de moment dipolaire intrinsèque. H_2 et O_2 n'absorbent donc pas en infrarouge, contrairement à CO_2 et CH_4 qui, elles, absorbent en IR car elles induisent instantanément un moment dipolaire.

Le nombre de modes de vibration possible d'une molécule va dépendre du nombre d'atomes N constituant la molécule étudiée et il existe 3N-6 déplacements possibles, sauf pour les molécules linéaires dont le degré de liberté vaut 3N-5.

On peut observer deux types de vibration dans le cas des molécules non linéaires ayant au minimum trois atomes : les vibrations d'élongation notées ν qui impliquent la variation de longueur de liaison interatomique et les déformations notées δ qui correspondent à la variation de l'angle entre deux liaisons. Ces déformations peuvent prendre plusieurs directions, être symétriques ou asymétriques pour les vibrations de type ν. Elles peuvent être de rotation, de cisaillement, de balancement et de torsion pour les vibrations de type δ.

Les vibrations ν demandent moins d'énergie que les vibrations δ. On observe le plus souvent les formes de déformation δ dans les bas nombres d'onde.

La région du moyen infrarouge comprise approximativement entre 4000 cm^{-1} et 400 cm^{-1} est une région particulièrement intéressante, pour la caractérisation des fonctions chimiques. La figure 16 montre quelques régions des bandes d'absorption de certaines liaisons principales (simple, double, triple) entre atomes de carbone, azote et hydrogène. La fréquence d'absorbance dépend aussi des effets inductifs et mésomériques d'une molécule (effets électroniques, couplage avec d'autres vibrations) ainsi que de son environnement chimique.

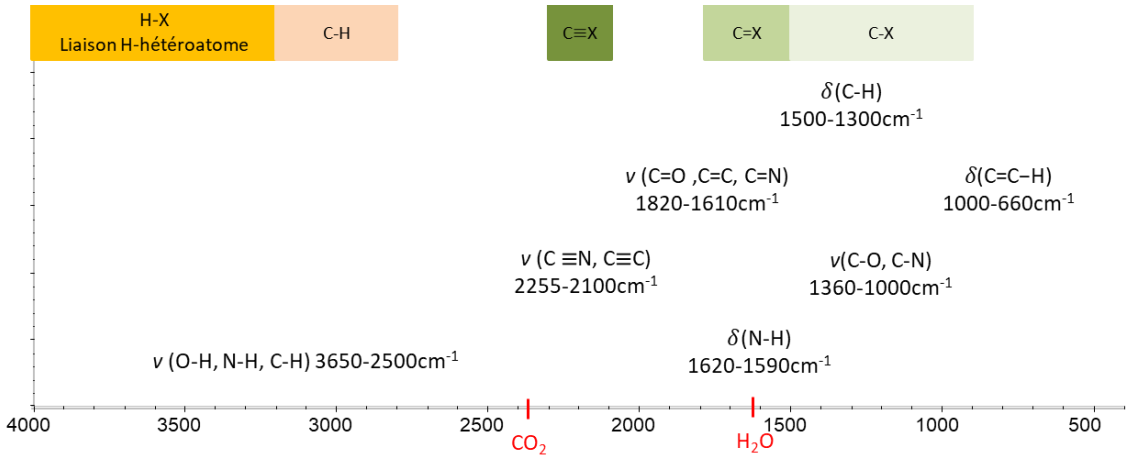


Figure 16 – Régions des groupes fonctionnels en spectroscopie IR.

La spectroscopie infrarouge est donc une technique classique pour déterminer la structure, la composition et l'architecture de molécules. Sa sélectivité selon les liaisons chimiques et sa facilité de mise en œuvre ont fait de cette technique un précieux outil pour déterminer les structures de protéines [129].

2.1.2. La spectroscopie infrarouge à Transformée de Fourier

Le principe de la spectroscopie infrarouge à Transformée de Fourier (IRTF) a été découvert au début du XXème siècle. Mais c'est seulement à partir des années 60, avec l'évolution de l'électronique et des ordinateurs, que les spectromètres à Transformée de Fourier ont remplacé les systèmes à « balayage » (séparation des longueurs d'onde « dispersive » à l'aide d'un prisme ou réseau) dans la plupart des applications. La spectroscopie IRTF a permis de résoudre les inconvénients majeurs des précédents instruments. En effet, l'IRTF résout les problèmes tels que la très longue durée d'acquisition (car l'absorption de chaque longueur d'onde est mesurée de façon indépendante) et l'imprécision sur la valeur de la longueur d'onde du faisceau séquencé des instruments à balayage (de ±0,01 cm^{-1} pour IRTF contre ±2 cm^{-1} pour le système dispersif) [130].

L'analyse et le spectre IRTF sont obtenus grâce à un interféromètre, l'interféromètre de Michelson schématisé en figure 17 [131].

Le principe consiste à séparer en deux le faisceau incident sur une lame séparatrice pour le moyen infrarouge. La première partie du faisceau est réfléchie par un miroir fixe et la seconde, par un miroir mobile. Les faisceaux réfléchis convergent ensuite à l'aide de la lame séparatrice vers l'échantillon et interfèrent. Le signal détecté est appelé un interférogramme.

L'interférogramme est une mesure d'intensité en fonction du retard optique noté δ, c'est-à-dire de la différence de marche entre les deux branches du faisceau dépendant du déplacement du miroir mobile. Des *maxima* sont enregistrés pour δ = nλ et *minima* pour δ = (n +1/2)λ , avec n un entier positif ou négatif.

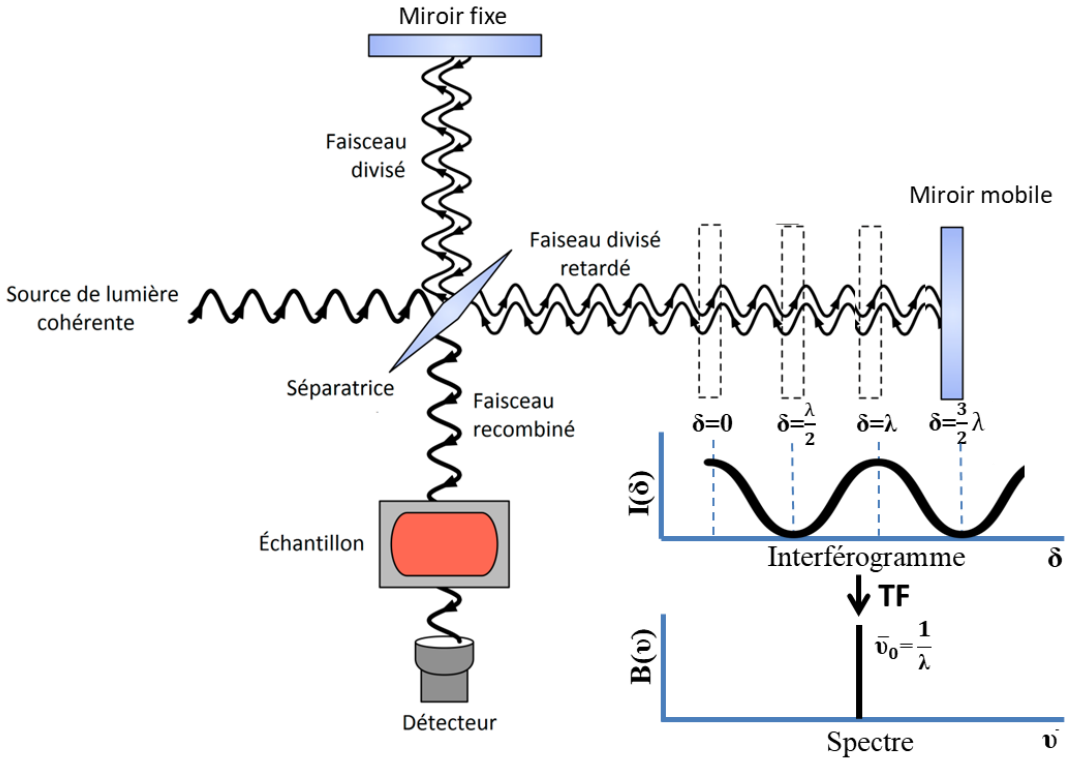


Figure 17 – Schéma d'un interféromètre de Michelson : principe du fonctionnement de la spectroscopie IRTF. Représentation de l'intensité en fonction de différents retards optiques pour une source monochromatique (images tirées du domaine public et modifiées).

Pour un faisceau monochromatique, l'intensité $I(\delta)$ est décrite par une sinusoïdale définie selon la relation suivante :

$$I(\delta) = B(\bar{\nu})\cos(2\pi\delta\bar{\nu})$$

Avec $B(\bar{\nu})$, la fonction représentant l'intensité du spectre en fonction du nombre d'onde.

Lorsque la source polychromatique, il est possible de traiter chaque fréquence comme si chacune produisait un cosinus séparé (chaque cosinus avec sa propre périodicité) et d'ajouter ensuite géométriquement chacun des cosinus pour obtenir la forme de l'interférogramme résultant.

La source infrarouge typique émet un « continuum » qui est habituellement assimilable à un corps noir. On peut décrire l'interférogramme par :

$$I(\delta) = \int_0^{+\infty} B(\bar{\nu})\cos(2\pi\delta\bar{\nu})d\bar{\nu}$$

En appliquant la transformée de Fourier (TF), on obtient la fonction du spectre infrarouge suivante :

$$B(\bar{\nu}) = \int_{-\infty}^{+\infty} I(\delta)\cos(2\pi\delta\bar{\nu})d\delta$$

On collecte et traite ainsi toutes les fréquences de la source infrarouge simultanément, ceci permettant d'obtenir des spectres très rapidement.

La spectroscopie d'absorption IRTF est une technique rapide, précise et non destructive qui permet d'étudier des matériaux organiques et inorganiques. Il existe différents types de systèmes expérimentaux et le choix doit être judicieux en fonction des caractéristiques de l'échantillon à étudier (état, concentration, rugosité…)

2.2. Mode de mesure et échantillonnage

Le choix du mode de mesure est une partie importante et dépend de l'échantillon.

Il existe quatre modes de mesure principaux (et quelques-uns plus spécialisés) très souvent utilisés et chacun a ses avantages et inconvénients (figure 18) :

- **Par transmission :** mode utilisé pour des échantillons minces n'absorbant pas complètement le rayonnement dans le domaine spectral d'intérêt. Il permet d'étudier des sections fines ou des substances en solution en film (dans un solvant organique ou aqueux). La qualité du spectre, c'est-à-dire le rapport signal sur bruit, dépend du trajet optique, de la quantité de l'échantillon mais aussi de la rugosité de l'échantillon et des propriétés optiques des fenêtres d'une cellule.

Les techniques suivantes sont des techniques de réflexion. Elles sont souvent utilisées lorsque la transmission des photons est impossible.

- **Par réflexion spéculaire et réflexion diffuse :** le montage de réflexion spéculaire consiste en un jeu de miroirs permettant d'obtenir un spectre de réflexion. Ce

montage est utilisé pour des échantillons solides réfléchissant comme des miroirs et ayant une très faible rugosité (c'est à dire très inférieure à la longueur d'onde du rayonnement). Dans le cas où la surface est rugueuse et où l'échantillon est partiellement réfléchissant, on utilisera plutôt le mode de réflexion diffuse. En effet, le rayonnement au contact d'une surface rugueuse sera dispersé dans toutes les directions voire pénétrera dans l'échantillon et suivra un trajet complexe avant d'en ressortir sous un angle «aléatoire». Dans ce cas, on utilisera plutôt un miroir (réflecteur) hémisphérique pour collecter le rayonnement issu de l'échantillon. Un traitement des données spécifiques est nécessaire pour ces deux modes car les spectres mesurés tiennent non seulement compte de l'absorption mais aussi d'autres phénomènes optiques comme la réflectivité ou la diffusion (dépendants de l'indice de réfraction complexe, de la permissivité diélectrique, de la granularité et du coefficient d'extinction). Des transformations mathématiques comme celle de Kramers-Krönig (pour la réflexion spéculaire) et celle de Kubelka-Munk (pour la réflexion diffuse) sont nécessaires pour obtenir un spectre exploitable et séparer les différents phénomènes [132].

- **Par réflexion totale atténuée (ATR) :** ce mode de mesure consiste à se placer dans une condition de réflexion totale. Pour cela, on utilise un cristal à fort indice de réfraction dans l'infrarouge (exemples : germanium avec un indice d'environ 4, diamant avec un indice d'environ 2,4) placé au contact de l'échantillon. L'échantillon peut être solide ou liquide. Le faisceau incident est focalisé dans le cristal à un angle supérieur à l'angle critique et est donc totalement réfléchi à l'intérieur du cristal, cependant une onde évanescente est créée à l'interface entre le cristal et l'échantillon. La profondeur de pénétration (dp) de l'onde évanescente varie en fonction de la longueur d'onde λ et de l'indice de réfraction du cristal, modifiant les intensités relatives des pics situés à différentes longueurs d'onde. De ce fait, le spectres obtenu en ATR est différent du spectre obtenu en transmission. Il est donc nécessaire de corriger les spectres ATR. Par ailleurs, la profondeur de pénétration dépend aussi de l'angle θ (en rad) du faisceau incident et des indices de réfraction du cristal (n_c) et de l'échantillon (n_e). La profondeur de pénétration est exprimée par [133] :

$$dp = \frac{\lambda}{\sqrt{2\pi\ (n_e{}^2\ sin^2\ \theta - n_c{}^2)}}$$

Les épaisseurs analysées seront donc plus grandes pour des angles incidents plus faibles, des nombres d'ondes plus faibles ainsi que des cristaux d'indice de réfraction plus faibles. L'ordre de grandeur de profondeur de pénétration est de quelques centaines de nanomètres à quelques micromètres. (La pénétration de l'onde est généralement de l'ordre de 1/5 à 1/20 de la longueur d'onde.) Donc seul l'échantillon au contact du cristal (surface cristal/échantillon) absorbe le rayonnement incident et on obtient ainsi son spectre IR [130].

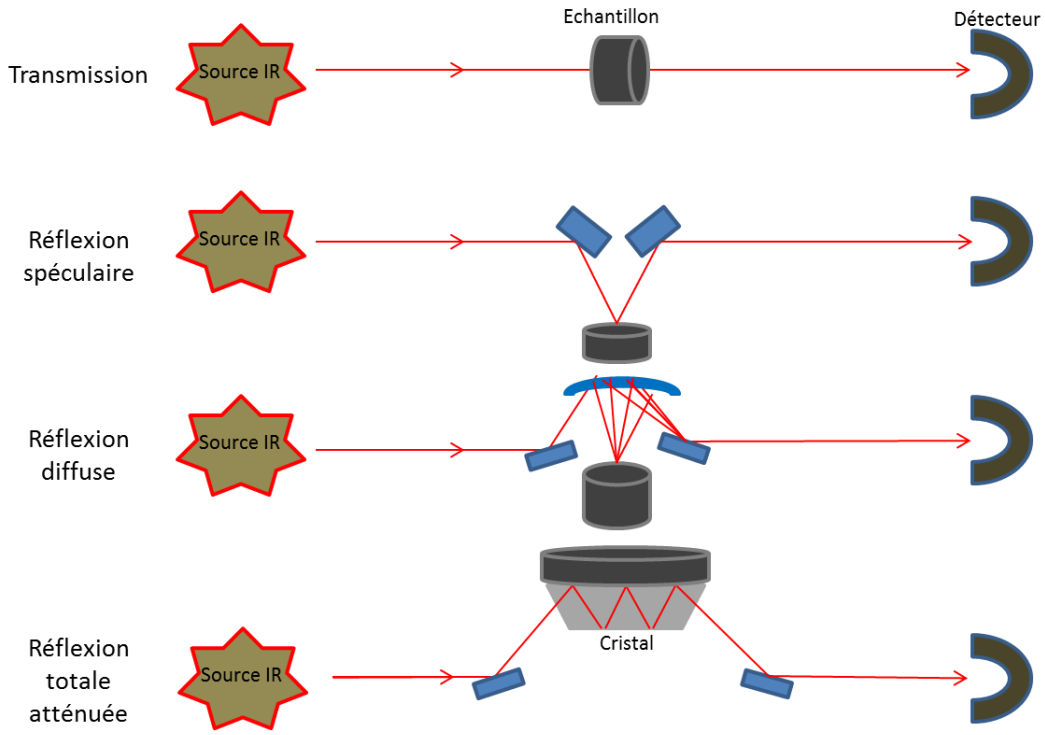


Figure 18 – Schémas des montages infrarouges les plus fréquents. Les montages sont le plus souvent disponibles sous forme d'accessoires adaptables sur les spectroscopes IR commercialisés. (Inspiré de [132]).

Pour une solution contenant des protéines et des bactéries, nous avons souvent utilisé les montages en transmission ou en ATR.

2.3. Imagerie chimique et résolution spatiale

2.3.1. Spectromicroscopie infrarouge en champ lointain

Lorsque la spectroscopie est combinée à la microscopie optique (spectromicroscopie à champ lointain), il est possible d'obtenir une cartographie (image chimique) de l'échantillon c'est à dire avoir une visualisation spatiale des constituants (2D, ou même 3D par tomographie, en infrarouge).

La résolution des images obtenues par spectromicroscopie à champ lointain est limitée par la diffraction. La diffraction est observable par la présence d'une « tache d'Airy » pour un « point » objet. Pour un objet éclairé par une source large telle qu'une source IR, le profil global de diffraction de l'objet sera composé d'une superposition de plusieurs taches de diffraction. Ce phénomène limite la résolution spatiale de nombreuses techniques mais permet de déterminer la forme et la distance entre des objets (techniques de diffraction).

La capacité à mesurer l'image ou le signal d'un objet indépendamment de son environnement est caractérisée par la résolution spatiale du système optique. Cette limite de résolution latérale *d* peut être estimée par le critère de Rayleigh qui spécifie que pour un objet soit résolu, il doit être situé à une distance *d* correspondant au premier minimum du profil de Airy de l'objet le plus proche. La résolution spatiale d'un système optique est décrite empiriquement par l'équation d'Abbe :

$$d = \frac{0{,}61\lambda}{n\sin\theta}$$

Où λ est la longueur d'onde de la source, θ l'angle d'ouverture (la valeur $n\sin\theta$ est aussi appelée « ouverture numérique » de l'objectif) et *n* l'indice milieu dans lequel l'objectif se trouve. Afin d'optimiser la résolution en champ lointain, il faut travailler dans un milieu d'indice de réfraction élevé (principe des objectifs à immersion) ou travailler à de courtes longueurs d'onde. En visible et en infrarouge on a donc une résolution limitée, et l'optimisation des optiques en microscopie permet au mieux d'avoir une résolution proche de λ/2. En microscopie champ visible, la meilleure résolution possible est de d = 160 nm. La limite de résolution latérale en microscopie infrarouge, allant de 2 µm à 75 µm selon la longueur d'onde prise en compte, est atteinte en utilisant un rayonnement synchrotron [134]–[137]. En effet, le rayonnement synchrotron est une source permettant d'avoir une meilleure brillance (de plus de cent fois, c'est à dire une plus grande densité de photons par unité

angulaire). Il permet en outre d'atteindre une résolution spatiale de l'ordre de la longueur d'onde ou mieux, jusqu'à 3 µm pour les protéines, à l'aide d'un microscope confocal [134].

Un microscope confocal permet d'améliorer les résolutions latérales et axiales en éliminant une grande partie des figures de diffraction causées par le dispositif optique et par l'échantillon. Le premier diaphragme permet de réduire la taille du faisceau sur l'échantillon. Le second permet de s'affranchir de la lumière des plans adjacents et de « détecter » uniquement celle du plan focal de l'objectif. Cela entraîne un meilleur contraste car la profondeur de champ est diminuée. Ce dispositif améliore également la résolution latérale en x et en y. Pour reconstruire l'image complètement avec ce dispositif, il est nécessaire de balayer la surface [138].

En ce qui concerne la spectromicroscope confocale IRTF (figure 19), la lumière passe d'abord par un interféromètre de Michelson puis est focalisée sur un premier diaphragme suivi de l'objectif qui focalise la lumière sur l'échantillon. Le diaphragme définit la taille de la région mesurée et donc la résolution optique du système mais occasionne une perte de photons limitant le rapport signal sur bruit et ultimement la résolution spatiale de la mesure. Le signal peut être détecté en transmission ou en réflexion. Le spectre d'absorption correspondra au point focal. En balayant et mesurant pas à pas l'échantillon, on obtient une cartographie chimique. Les données sont dites hyperspectrales car elles possèdent deux dimensions spatiales et un grand nombre de dimensions spectrales correspondant chacune à l'une des longueurs d'onde du spectre : la position de l'éclairage ou de l'échantillon (x,y) et les valeurs d'absorption aux différentes longueurs d'onde mesurées (A,λ). La distribution spatiale des différents composants de l'échantillon peut être mesurée grâce aux spectres enregistrés en tous points de l'échantillon.

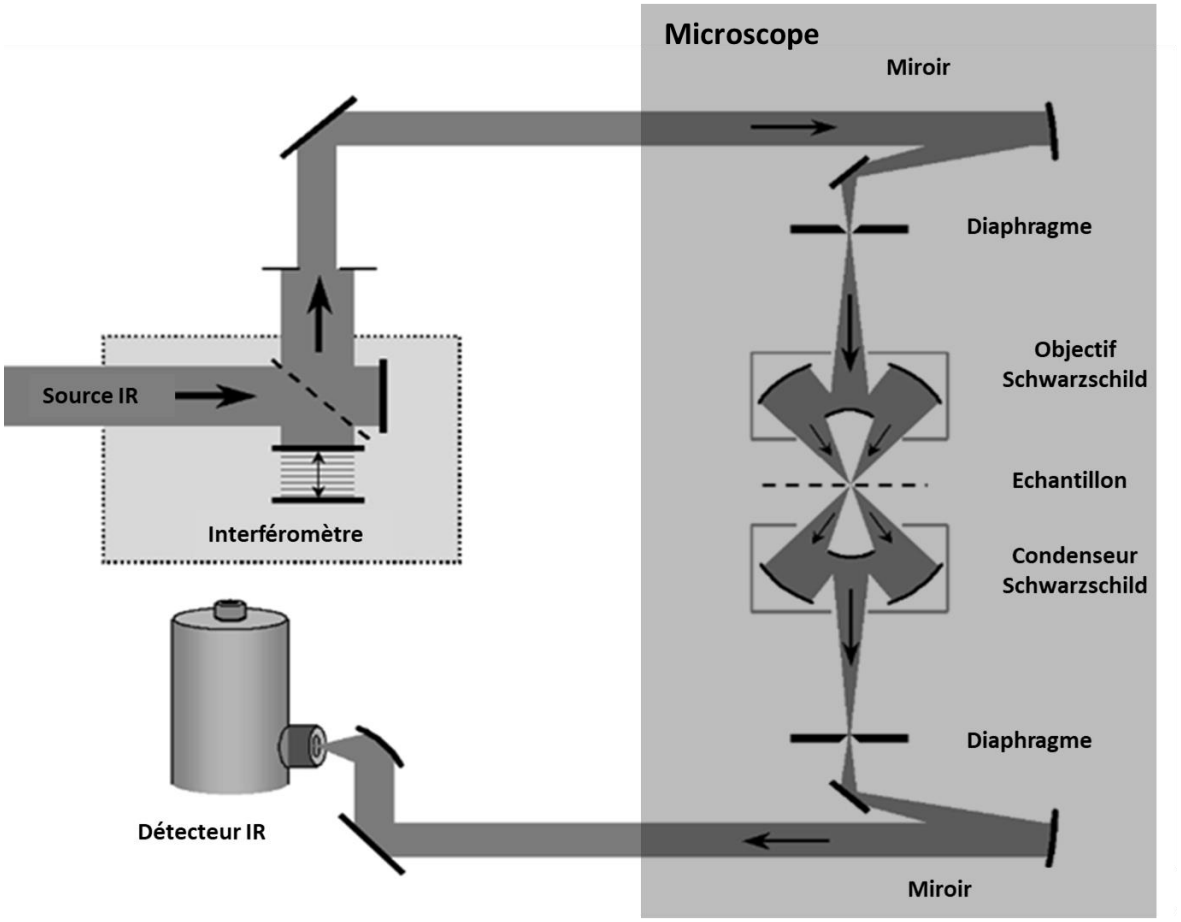


Figure 19 – Schéma d'un spectromicroscope confocal IRTF. Modifiée de [139].

Il est possible d'atteindre des résolutions de l'ordre de 25-50 µm avec un microscope IRTF moderne équipé d'une source conventionnelle. Au moyen d'un microscope couplé à une source synchrotron, les résolutions peuvent atteindre jusqu'à 3 µm pour les protéines [134].

Afin de passer outre les limites de résolution de la microscopie optique, des techniques en champ proche ont été développées afin d'imager des objets de taille inférieure au micron. L'association des techniques champ proche à la spectroscopie permet d'obtenir des informations chimiques locales à l'échelle du sub-micrométrique.

2.3.2. Microscopes en champ proche

Lorsqu'une onde électromagnétique interagit avec un objet, deux types de champs diffractés se forment : un champ « lointain » et un champ « proche ». Le champ « lointain » tient son nom de son caractère propagatif et est détecté plus loin à une distance plus grande que la longueur d'onde. Les ondes diffusées par l'objet peuvent se propager jusqu'à l'infini en absence d'absorption. Le champ « proche » reste au voisinage de l'échantillon. À quelques nanomètres de l'objet pour l'infrarouge/visible, l'amplitude des ondes évanescentes n'est pas négligeable devant celle des ondes propagatives.

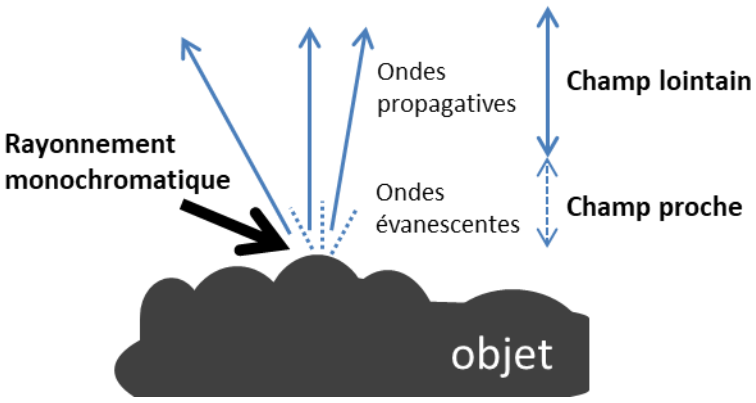


Figure 20 – Représentation des champs lointain et proche respectivement par des ondes propagatives et évanescentes d'un objet diffractant un rayonnement monochromatique.

Les microscopes en champ proche utilisent des sondes qui balayent la surface. Ils sont désignés par SPM pour S*canning Probe Microscopes*. Depuis les années 1980, de nombreux types de SPM ont vu le jour. Chronologiquement sont apparus le microscope à effet tunnel, le microscope à force atomique et le microscope optique à champ proche.

Le premier microscope à sonde locale a permis l'observation d'atomes à la surface d'un conducteur solide (ou semi-conducteur). Le principe de fonctionnement du microscope à effet tunnel a été découvert par deux chercheurs de la société IBM : G. Binnig H. Rohter qui ont reçu pour leurs travaux un prix Nobel de physique en 1986 [145][146]. L'inconvénient de la technique est que les échantillons doivent être conducteurs pour pouvoir les observer. Pour soulever ce problème, la microscopie à force atomique (AFM de *Atomic Force Microscope*) a été développée quelques années plus tard par la même société [142]. Elle permet d'imager des échantillons isolants et organiques.

2.3.3. Microscope à force atomique

Pour sonder la surface, la microscopie à force atomique (AFM) utilise une sonde locale, comme tout SPM. La sonde locale de l'AFM est un levier flexible micrométrique se terminant par une pointe souvent en forme conique/pyramidale d'une hauteur de quelques micromètres et ayant un rayon de courbure de quelques nanomètres.

L'AFM permet d'obtenir une topographie de surface en étudiant l'interaction entre la pointe et la surface de l'objet sondé. Les interactions étudiées sont des forces qui peuvent être répulsives (forces coulombiennes) ou attractives (forces coulombiennes ou de van der Waals). Le contrôle de la distance pointe-surface s'obtient par le suivi de la déflexion du bras de levier. Pour suivre cette déflexion, un faisceau laser est pointé au bout du levier (au niveau de la pointe). Le laser est alors réfléchi et centré vers une photodiode à quatre cadrans. Lors d'une interaction pointe/surface, le levier se déforme. Cette déformation minuscule (déflexion) est mesurée par le déplacement du laser sur la photodiode. Les mouvements détectés en fonction de la force appliquée sur le levier peuvent être des « flexions » et/ou des « torsions ». Lors du balayage de la surface, les signaux sur la photodiode sont directement analysés et une boucle d'asservissement rétablit la position centrée du laser sur le cadran en ajustant la distance pointe-échantillon par un piézoélectrique (Z). Des piézoélectriques permettent aussi le balayage de l'objet (sur un plan en X et Y).

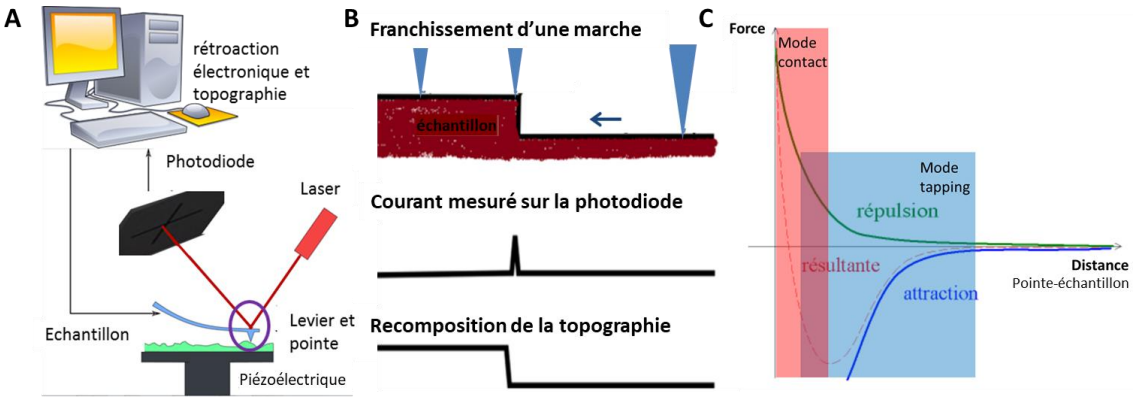


Figure 21 – Principe et fonctionnement d'un microscope à force atomique.
(A) Schéma de montage de l'AFM. (B) Schéma pour l'imagerie en mode contact. (C) Courbe de potentiel interatomique de Lennard-Jones, force en fonction de la distance pointe-échantillon et les modes AFM associés. (Images tirées du domaine public et modifiées, CC).

Il existe trois modes de fonctionnement en AFM : un mode statique (mode contact) et deux modes dynamiques (mode tapping [143] et mode « non » contact [144]). Nous décrirons brièvement les modes contact et tapping car ce sont les principaux modes pour faire de la topographie.

- Mode contact : il est le premier mode utilisé en AFM et le plus simple. Dans ce mode, la pointe est en « contact » avec la surface et donc dans un régime répulsif. Il consiste à maintenir la force entre la surface et la pointe constante (appelée technique « à force constante ») à une certaine valeur de consigne en agissant sur la hauteur pointe-échantillon. La topographie est obtenue par le suivi des variations en hauteur.

- Mode tapping : il consiste à créer un contact pointe/échantillon par intermittence. Pour cela, on fait osciller le levier à sa fréquence de résonance. Son principe de fonctionnement est fondé sur la variation d'amplitude et de phase de l'oscillation du bras de levier durant le balayage de l'échantillon, et donne ainsi accès à sa topographie. Ce mode de fonctionnement assure une sécurité lors du balayage d'échantillons fragiles et non lisses.

La bonne résolution latérale (de l'ordre d'une dizaine de nanomètres) et verticale (de l'ordre de l'angström) a fait de l'AFM un outil incontournable pour étudier la topographie/défauts de surface à l'échelle nanométrique. La résolution spatiale n'est limitée uniquement que par l'apex de la pointe de la sonde AFM.

Un grand nombre de techniques basées sur AFM ont été développées pour identifier les matériaux sur la base de diverses propriétés. La pointe utilisée peut être composée de différents matériaux et / ou fonctionnalisée, ce qui permet d'utiliser l'AFM dans de nombreuses applications et domaines. Elle permet notamment de mesurer l'adhésion, l'élasticité des matériaux, l'observation d'interactions entre protéines, faire de l'imagerie par mesure de champs électrostatique et magnétique…

Afin d'identifier les différents constituants (ou molécules) d'un objet sur sa topographie, des techniques à champ proche ont été couplées avec la spectroscopie afin d'obtenir des cartographies chimiques d'une résolution sub-micrométriques. Nous allons voir dans la partie qui suit ces techniques et en particulier l'AFM-IR, utilisé durant la thèse.

2.3.4. Spectromicroscope infrarouge en champ proche

2.3.4.1. Introduction

L'intérêt d'utiliser le champ proche en spectromicroscopie infrarouge est d'obtenir une cartographie chimique c'est-à-dire une information spatiale à très haute résolution (de l'ordre sub-micrométre jusqu'au nanomètre) des constituants de l'objet. Pour cela, la microscopie en champ proche doit être couplée à la spectroscopie infrarouge. Plusieurs techniques permettent de faire de la spectromicroscopie infrarouge. Elles peuvent être optiques ou thermiques. Nous allons exposer ci-dessous les techniques thermiques, dont l'AFM-IR en fait partie.

2.3.4.2. Techniques en champ proche thermique

Il existe deux types de technique de spectromicroscopie thermique en champ proche. Le premier est à détection thermique/température (PTMS de *PhotoThermal MicroSpectroscopy*) le second est à détection mécanique (AFM-IR).

La PTMS se base donc sur la mesure de l'évolution de la température locale d'un échantillon absorbant un rayonnement IR. En effet, lors de l'absorption les vibrations moléculaires sont excitées puis se relaxent de façon non radiative, et émettent des ondes thermiques. La sonde est un thermocouple nanométrique (un capteur de température constitué de deux matériaux ayant une conductivité thermique différente, une variation de température génère un courant) placé en contact de l'échantillon. Cette sonde se comporte comme un thermomètre nanométrique. Le signal obtenu par le thermocouple est envoyé à un spectromètre infrarouge à Transformée de Fourier classique afin d'obtenir un spectre-infrarouge. Il a été montré que l'utilisation d'une source synchrotron permet d'améliorer le contraste des images [145].

Hammiche a montré qu'il est possible de réaliser des images chimiques et des spectres avec une résolution de l'ordre de la sub-longueur d'onde. La limite de résolution possible est proche d'une centaine de nanomètres correspondant à la taille minimum des sondes de thermocouples. Mais cette résolution reste difficile à atteindre [31][32][147].

Pour l'analyse de l'auto-assemblage de fibres amyloïdes nous avons utilisé principalement AFM-IR (*Atomic Force Microscope InfraRed*) qui utilise et mesure des propriétés mécaniques.

2.3.4.3. L'AFM-IR : présentation et fonctionnement

Dans le but d'établir une technique permettant des analyses chimiques par spectroscopie infrarouge avec la résolution spatiale de l'AFM, l'AFM-IR a été développé en 2004 par Alexandre Dazzi au Laboratoire de Chimie Physique de l'Université Paris-Sud

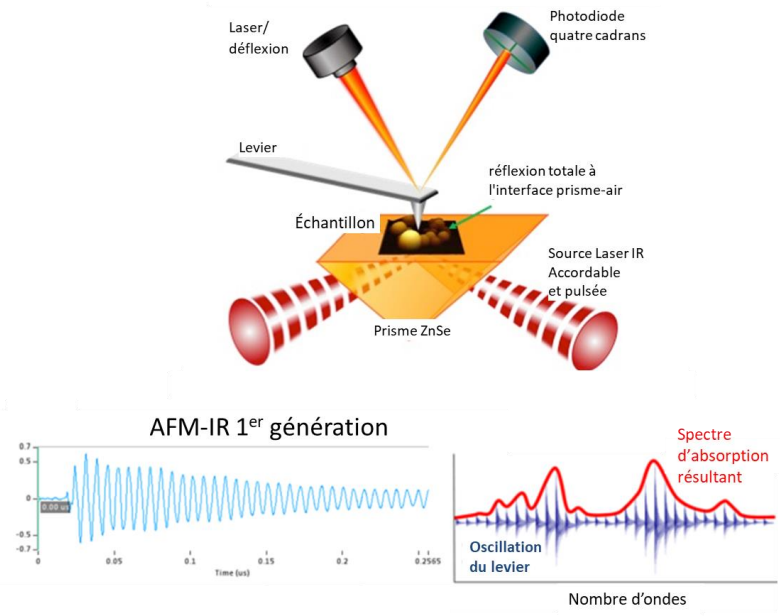


Figure 22 – Schéma de principe de l'AFM-IR première génération : éclairage par-dessous. Une source laser accordable pulsée et focalisée éclaire un échantillon à travers un prisme ZnSe. Lorsque le laser est accordé à une fréquence d'absorbance de l'échantillon, la lumière absorbée entraîne une expansion photo-thermique dilatant l'échantillon. Cette dilatation entraine une oscillation décroissante ou amortie du levier. La mesure de l'amplitude de l'oscillation du levier pour chaque nombre d'onde fournit un spectre d'absorption local. Adapté avec autorisation de [148]. Copyright (2017) American Chemical Society.

La technique consiste à utiliser un laser infrarouge pulsé et accordable couplé à un microscope à force atomique. L'échantillon absorbant le rayonnement infrarouge va générer un effet photo-thermique. La pointe de l'AFM-IR va détecter une dilatation de l'échantillon due à l'augmentation locale de la température. Cela provoque une force d'impulsion sur la pointe se trouvant à proximité de l'objet dilaté. Cette impulsion crée des mouvements d'oscillation aux fréquences propres du levier. Ces oscillations sont observées sur la photodiode quatre cadrans du microscope AFM. La pointe AFM agit ainsi comme le détecteur IR.

Une étude théorique [149] a montré que l'amplitude des oscillations du levier est directement proportionnelle à l'absorption par l'échantillon. La technique permet d'avoir des spectres d'absorption localisées ou des cartographies chimiques (en même temps qu'une acquisition topographique).

Le spectre d'absorption est obtenu en fixant la sonde sur un point de l'objet. L'amplitude d'oscillation du levier est mesurée pour chaque longueur d'onde et ainsi le spectre d'absorption IR est reconstruit pas à pas (longueur d'onde par longueur d'onde).

La cartographie chimique est obtenue en fixant le nombre d'onde du laser et en balayant la pointe sur la surface. L'amplitude d'oscillation du levier est mesurée en chaque point de l'échantillon en même temps que sa topographie. Ce qui nous permet d'avoir à la fois la topographie de l'échantillon et la distribution des absorbances à la longueur d'onde sélectionnée. Si la longueur d'onde sélectionnée correspond à l'absorption spécifique d'une molécule, on obtient une carte montrant la distribution de cette molécule dans l'échantillon.

La résolution latérale et la sensibilité de la mesure axiale de l'AFM-IR dépendent de différents paramètres : la dimension de la pointe, la taille et la forme de l'échantillon mais aussi le signal détecté. Cependant, le signal en régime d'oscillation amortie reste difficilement exploitable car faible pour des objets de quelques nanomètres. Pour surmonter cette restriction, des évolutions récentes de l'AFM-IR ont été développées et vont être présentées ci-dessous.

2.3.4.4. Évolution et amélioration de AFM-IR

Des améliorations conséquentes ont été réalisées afin d'une meilleure résolution verticale permettant d'imager et d'obtenir des informations chimiques d'un objet à des dimensions inférieures à 5 nm.

Pour améliorer la résolution verticale deux principes ont été exposés par les équipes de F. Lu et M. Belkin (Figure 23) [161][162].

- Mise en œuvre d'une résonance forcée du levier AFM
- Un effet d'exaltation par amplification de l'intensité du champ électrique incident par l'utilisation de métaux nobles.

Le premier microscope AFM-IR développé par A. Dazzi *et al.* [152] utilise un laser pulsé de type OPO (de l'anglais *optical parametric oscillator*) qui possède une fréquence de tir de 1khz. F. Lu et M. Belkin proposent l'utilisation d'un laser QCL (de l'anglais *quantum cascade laser*). Ce laser a une plus grande fréquence de tir de 500khz permettant la synchronisation de la fréquence de tir du laser avec l'un des modes de résonnance d'oscillation du levier sur l'échantillon. Par cette nouvelle façon de procéder, on passe d'une oscillation transitoire décroissante ou amortie à une oscillation entretenue augmentant la sensibilité. « L'AFM-IR amélioré par résonance » fournit une oscillation continue du levier qui permet une détection beaucoup plus efficace de l'absorption IR et donc des mesures plus sensibles [150].

L'équipe de M. Belkin a aussi apporté une amélioration sur la résolution verticale en jouant sur l'exaltation du champ électrique [151]. Le montage a été modifié pour un éclairage de l'échantillon par le dessus. La différence par rapport à la première génération de l'AFM-IR est que l'éclairage de l'échantillon ne passe plus par un prisme à réflexion totale (ZnSe) mais se fait directement sur l'échantillon. Ce qui a pour avantage de pouvoir déposer l'échantillon sur n'importe quel substrat du moment qu'il n'absorbe pas en IR et est suffisamment lisse. Une contrainte supplémentaire de l'éclairage par le dessus est que la pointe ne doit pas absorber en IR. Une pointe métallique (ou recouverte d'un métal) et un substrat métallique permettent de générer une amplification du champ électromagnétique local (effet d'exaltation) focalisé sous la pointe. Pour bénéficier de cet effet d'exaltation, le métal choisi recouvrant les pointes et les substrats est généralement l'or.

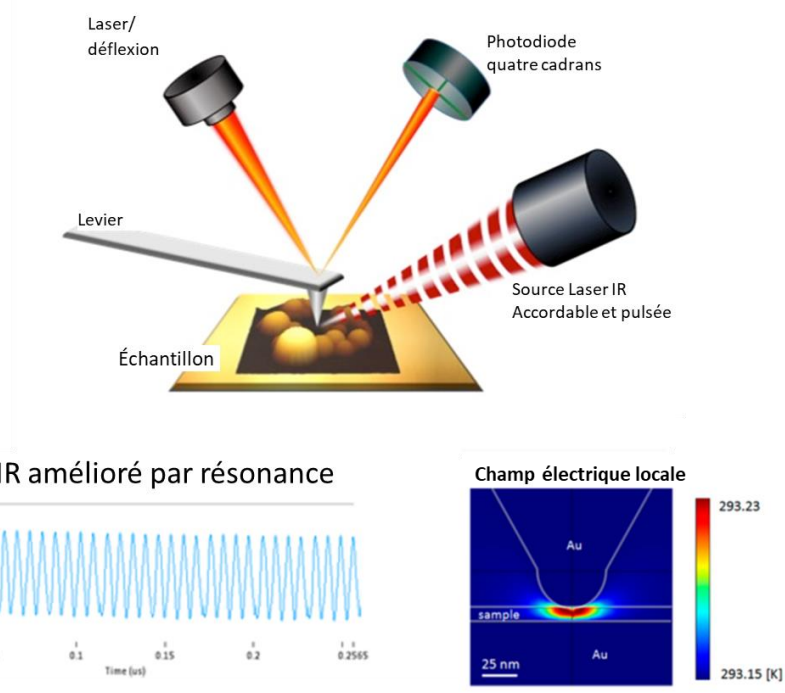


Figure 23 – Schéma et principe de l'AFM-IR deuxième génération : éclairage par-dessus. L'AFM-IR amélioré par résonance permet une oscillation continue forcée du levier. L'augmentation locale du champ électrique avec une pointe et un substrat d'or permettant une augmentation de la sensibilité de l'AFM-IR. Adapté avec autorisation de [148]. Copyright (2017) American Chemical Society.

Ces microscopes nous ont permis d'obtenir une sensibilité axiale et une résolution verticale inférieures à 10 nm. À l'aide de l'AFM-IR, nous avons donc pu étudier à l'échelle nanométrique la structure des fibres amyloïdes formées par le peptide CTR-Hfq.

2.4. Étude de la structure secondaire des protéines par IRTF

2.4.1. Les bandes d'absorption polypeptidique

La spectroscopie IRTF est une technique classique pour analyser les structures secondaires des protéines et des peptides en solution ou déshydratés. La liaison peptidique est une fonction amine et présente 9 bandes d'absorption actives en infrarouge nommées bandes amides A, B et I à VII. La région d'intérêt pour l'analyse des structures protéiques est comprise entre 1750-1400 cm^{-1} et contient deux fortes bandes d'absorbance sensibles à l'angle de la liaison peptidique : les bandes amides I et amides II qui sont les plus fortes et les plus utilisées pour identifier la structure des protéines. Le tableau ci-dessous représente les modes les plus forts pour une molécule polypeptidique.

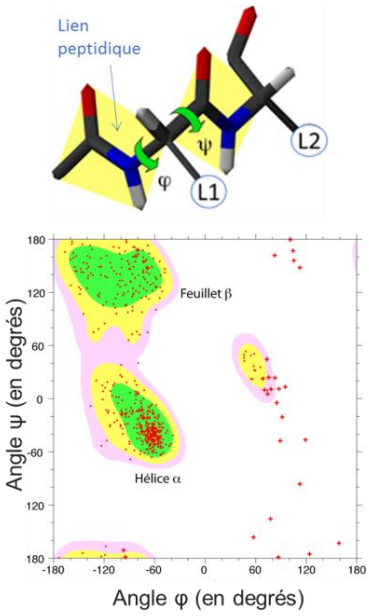
Mode	Nombre d'ondes (cm^{-1})	Association
Amide A (et B)	~3300 cm^{-1} (~3100 cm^{-1})	La bande amide A est due à plus de 95% à l'élongation N-H. Ce mode de vibration ne dépend pas de la conformation du squelette mais est très sensible à la liaison hydrogène.
Amide I	1600 -1700 cm^{-1}	La bande d'absorption la plus intense et la plus large des protéines. Elle est principalement due aux vibrations d'élongation des groupes C=O (70-85%) et d'élongation C-N (10-20%). La position exacte de la bande est déterminée par la conformation/ structures secondaires (liaison H).
Amide II	1510 -1580 cm^{-1}	Généralement deux fois moins intense que l'amide I. L'amide II provient principalement de la déformation N-H dans le plan (40-60%). L'élongation C-N (18-40%) et l'élongation C-C (environ 10%) contribuent aussi.
Amide III	1250-1400 cm^{-1}	Etant de très faible intensité en IR, cette bande est limitée pour l'extraction d'informations structurelles. Assez complexe, ce mode est dépendant principalement de la vibration (déformation/pli dans le plan) de N-H avec un effet de C-C, C-N et C=O.

Tableau 2 – Résumé des caractéristiques des bandes Amides à symétrie dans le plan dans une liaison polypeptidique [153].

2.4.2. Détermination de la structure secondaire par la bande amide I

La bande Amide I, absorbant vers 1650 cm^{-1}, correspond principalement à la contribution de la vibration d'étirement C=O. Mais cette vibration est fortement délocalisée sur les liaisons amide C-N et C-H donnant lieu à un couplage fort : le couplage trans-dipolaire. La fréquence exacte du pic d'absorption est sensible à l'angle de la liaison amide, qui dépend de la structure locale et secondaire de la protéine. Peu influencée par les chaines latérales cette région est donc la plus utile pour la caractérisation structurale. Les hélices α, les feuillets β, les coudes β, la pelote statistique non périodique (aléatoire) et les hélices 3_{10} sont chacun restreints à un ensemble d'angles différents comme le montre le diagramme de Ramachandran (figure 24),

déterminé par la périodicité et la position des liaisons hydrogène. La fréquence d'absorption de la bande amide I est extrêmement sensible à ces angles donnant lieu à des pics amides situés à des fréquences différentes en fonction de la structure secondaire d'un domaine peptidique (tableau 3). Le spectre de la protéine ou du peptide est la somme des contributions de ses différents domaines. L'utilisation de l'IRTF pour la caractérisation des structures secondaires est reconnue et une banque de données spectrales se remplit jour après jour [154].



Structure	φ en degrés	ψ en degrés
Hélice α	-57°	-47°
Hélice 3_{10}	-49.0°	-26.0°
Hélice π	-57.1°	-69.7°
Hélice II / gauche	-79.0°	+145.0°
Feuillet β antiparallèle	-139.0°	+135.0°
Feuillet β parallèle	-19.0°	+113.0°

Figure 24 – Angles dièdres et structures secondaires caractéristiques. Schéma présentant les angles dièdres φ et ψ dans une chaîne protéique. L1 et L2 indiquent les chaînes latérales de deux acides aminés. Le Diagramme est celui de Ramachandran pour une protéine. Les zones énergétiquement favorables sont signalées par des contours colorés. Chaque acide aminé est représenté par un point rouge. Les croix correspondent aux acides aminées glycines. Le tableau présente les Angles dièdres moyens des structures protéiques. (Images tirées du domaine public)

L'information structurale est difficilement exploitable directement par l'observation visuelle du profil de la bande amide I. L'approche la plus courante pour déterminer la structure secondaire consiste à une première évaluation qualitative des structures secondaires par l'étude de la dérivée seconde du spectre IRTF ou par une déconvolution de bande par « *Fourier self-deconvolution* » (FSD) [155] suivie d'une analyse d'ajustement de courbe (*curve fitting*) [156]. La déconvolution et la dérivée seconde peuvent être utilisées pour trouver le nombre et la position exacte des pics se chevauchant dans la bande Amide I.

L’ajustement des courbes consiste à adapter une série de pics modèles aux données expérimentales. Les fonctions utilisées comme modèles sont habituellement des profils Lorentzien, Gaussien ou Lorentzien-Gaussien ou une combinaison donnant le profil de Voigt. La méthode la plus utilisée est celle d’une somme de fonctions gaussiennes représentant chacune la contribution d’une structure secondaire. L’ajustement de pics se fait par un algorithme itératif/empirique, et se poursuit jusqu'à ce que la somme des fonctions minimise la différence entre le spectre théorique et le spectre expérimental. Un test d’adéquation (test du Chi-2 ou fonction F) est généralement utilisé pour juger de la convergence entre données et théorie et évaluer la qualité de l’ajustement de courbe. L'intégration de la zone des bandes de composants utilisée pour générer l'ajustement de la courbe finale peut fournir une estimation des pourcentages relatifs de la structure secondaire. Le tableau 3 indique la correspondance entre fréquence/nombre d’onde des pics de la bande amide I et les caractéristiques de structures secondaires. Brièvement, un pic près de 1645 cm^{-1} indique la présence d’une pelote statistique, à 1655 cm^{-1} d'hélice α, et 1630 cm^{-1} et 1690 cm^{-1} de feuillet β. En revanche, des protéines agrégées montrent souvent une bande proche de 1620 cm^{-1} [157]. En effet une liaison hydrogène plus forte entraîne un déplacement vers des nombres d'onde plus faibles et, par conséquent, les fibrilles amyloïdes présentent souvent des pics de feuillets β absorbant à des nombres d’onde inférieurs à 1630 cm^{-1} [92]. Cependant il reste difficile de discriminer les feuillets β parallèles et antiparallèles.

Nombre d'ondes (cm^{-1})	Attribution
1615-1630	Feuillets β intermoléculaire (en particulier les vibrations des brins β agrégés fortement liés intermoléculaires, y compris les amyloïdes [92]).
1630-1640	Feuillets β intramoléculaires natifs.
1640-1660	Pelote statistique/ séquence peptidique désordonnée.
1650-1660	Hélices α.
1660-1690	Coudes β.
1675-1695	Feuillets β intermoléculaire (en particulier feuillets β agrégées antiparallèles/oligomères de fibrilles amyloïdes [158]).

Tableau 3 – Affectations empiriques pour les bandes infrarouges (IR) d'Amide I caractéristiques des structures secondaires des protéines. Tirée de [159].

Les superstructures (fibres amyloïdes, les tonneaux β, faisceau d'hélices/enroulements super hélicoïdaux…) ou les plus longues structures d'hélice α ou de feuillet β ont tendance à absorber à des nombres d'onde plus bas [129]. On peut noter que certains des pics caractéristiques des différentes structures peuvent se chevaucher dans certaines régions spectrales rendant l'attribution parfois ambigüe.

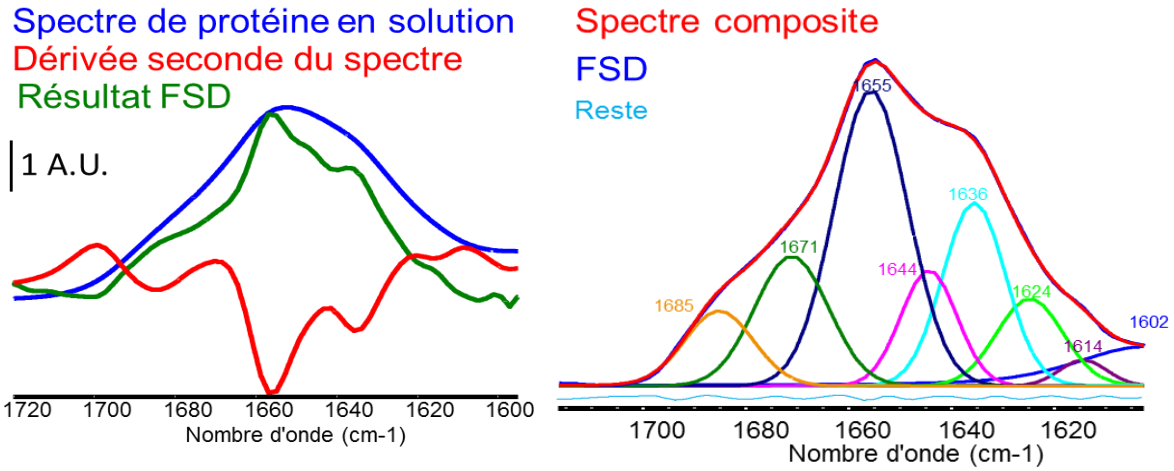


Figure 25 – Déconvolution par FSD et dérivées secondes (à gauche) et ajustement de courbes (à droite) d'un spectre générique de protéine.

Une des limites de l'étude des structures des protéines en solution est la forte bande de vibration de déformation de H_2O se trouvant vers 1640cm^{-1}, en plein milieu de la bande Amide I. L'eau absorbant très fortement l'infrarouge, le pic d'absorption de l'eau est généralement plus intense que celui des protéines en solution et sature pour un trajet optique

supérieur à 10 µm. La manière la plus simple pour analyser la structure de protéines en solution est d'utiliser une cellule de mesure avec un trajet optique inférieur à 8 µm en transmission composée d'un espaceur entre deux lames transparentes en infrarouge. Cela permet d'éviter de saturer. La soustraction du spectre de l'eau est néanmoins nécessaire et peut entrainer des artefacts.

Un autre moyen est de travailler en eau lourde (deutérée $^{2}H_2O$ ou D_2O). Ce qui permet de déplacer les bandes de vibration de $^{1}H_2O$ vers des nombres d'onde plus bas en dehors de la bande amide I. La bande Amide I, renommée Amide I', quant à elle ne se déplace peu car elle dépend principalement des vibrations d'étirement C=O et C-N. Elle a tendance à abaisser de quelques nombres d'ondes la fréquence de vibration des structures secondaires. [129] La bande amide II est quant à elle fortement impactée par échange isotopique de N-H en N-D.

La spectroscopie en mode ATR est une autre technique utilisée pour s'affranchir d'une trop forte absorption de l'eau. Comme nous l'avons vu, la spectroscopie ATR utilise une onde évanescente qui a une profondeur de pénétration limitée dans l'eau. Pour un cristal de ZnSe avec un angle de 42° la pénétration de l'onde évanescente dans l'eau n'est que de 0.54 µm à 1650 cm^{-1}. Cela permet de mesurer une solution protéique sans saturer la bande déformation O-H. Une plus forte concentration en protéines est néanmoins nécessaire pour compenser le trajet optique réduit.

2.4.3. Analyses structurales avec les autres bandes amides

L'utilisation de l'eau lourde permet d'étudier les changements de conformation de protéines par l'étude de la bande Amide II (appelée Amide II' dans $^{2}H_2O$) [160]. En effet, l'Amide II est sensible à l'échange $^{1}H/^{2}H$ en raison de la contribution principale de la vibration de déformation N-H. La bande amide II' se retrouve vers 1450 cm^{-1} (contre 1550 cm^{-1} pour la bande amide II). Cet échange permet d'étudier la flexibilité des protéines, puisque le taux d'échange des protons légers et lourds dépend de leur accessibilité, les résidus enfoncés dans la protéine étant moins accessibles. Il faut néanmoins noter que la force des liaisons hydrogène entre acides aminés va perturber fortement l'échange H/D. La liaison hydrogène peut-être protégée de l'échange même lorsqu'elle est exposée au solvant d'eau lourde. Cela a été démontré pour les petites molécules pour lesquelles la liaison hydrogène peut ralentir les taux d'échange de près de six ordres de grandeur. Le temps d'échange dépend de la

protéine. L'échange se fait plus rapidement pour les chaines latérales polaires que pour celles de la chaine peptidique. Les facteurs de protection/d'échange peuvent être déterminés par calcul et donnent une information sur la structure secondaire [175][176]. Mais l'échange H/D peut également entrainer une légère dénaturation de la protéine affectant son fonctionnement ce qui peut limiter son intérêt dans les études d'enzymes (remplacement des liaisons hydrogène par des liaison deutérium) [153].

Il est aussi possible de déterminer les structures secondaires à l'aide de la bande amide III qui y est extrêmement sensible et cela pourrait même permettre une meilleure prédiction de la structure surtout pour les tours β et la pelote statistique en la combinant avec les résultats de la bande amide I [163]. Mais la faible intensité de la bande amide III (quatre à cinq fois moins intense que la bande amide I) dégrade fortement son rapport signal sur bruit, elle est donc rarement étudiée.

2.4.4. Détermination de la structure *in vivo*

Les spectres dans le moyen infrarouge sont comme nous l'avons vu riches en information. En effet, les fréquences de vibration des molécules dépendent à la fois de leur structure mais aussi de leur environnement chimique. Pour l'étude de tissus ou de bactéries, l'infrarouge est particulièrement adapté, car cette technique ne cause pas de dommages grâce à la faible énergie des rayonnements contrairement aux rayons X ou UV. L'IRTF est d'ailleurs un outil utilisé en milieu hospitalier pour par exemple diagnostiquer des pathologies urolithiases [164] ou évaluer la quantité de lipides dans des foies en attente de transplantation [165]. La microspectroscopie IRTF est également utilisée en recherche pour évaluer son potentiel à discriminer différentes pathologies induisant des modifications chimiques (ou biochimiques) de cellules ou de tissus comme la présence de cellule cancéreuses aux marges de résection [166] ou un état hypoxique induisant une résistance aux anticancéreux [167]. L'IRTF associé à la microfluidique permet aussi de faire des examens sanguins rapides [168].

Par cette technique, il est aussi possible d'étudier la différence entre différentes souches bactériennes, de les identifier ou de les classifier [169][170]. En effet, les principales bandes observées dans les systèmes biologiques sont celles des protéines, des acides nucléiques, des lipides, des glucides [171]. Les bandes des glycoprotéines, des polysaccharides et des glucides sont souvent caractéristiques de la souche ou de l'espèce car les parois bactériennes et leurs déterminants antigéniques sont composés majoritairement de glycoprotéines.

Les protéines étant le type de biomolécule majoritaire dans la cellule bactérienne (70-80% en masse), le spectre infrarouge des bactéries est dominé par les bandes amides I et II. Les acides nucléiques qui présentent des bandes C=N et C=O dans la même région spectrale que la bande amide I sont en concentrations trop faibles pour interférer. Il est donc possible d'étudier les structures secondaires des protéines majoritaires dans la cellule bactérienne ou les différences entre deux souches [172][173][174]). D'autres groupes fonctionnels (acides nucléiques, lipides, sucres/glucose, …) chez la bactérie ont des bandes observables et sont répertoriés dans le tableau 4. En pratique il est même possible de faire l'empreinte métabolique d'une bactérie et de l'identifier à partir de son spectre infrarouge [175].

Nombre d'ondes du pic (cm^{-1})	Intensité relative	Attribution
~3500	Moyen	νO−H des groupes hydroxyles
~3200	Moyen à fort	νN-H amide A des protéines
~2959	Faible	ν_aC-H de -CH3
~2934	Faible	ν_aC-H de >CH2
~2921	Moyen	ν_aC-H de >CH2 dans les acides gras
~2872	Moyen	ν_sC-H de -CH3
~2852	Moyen	ν_sC-H de >CH2 dans les acides gras
~1740	Faible	νC=O des esters
~1715	Très faible	νC=O des acides carboniques, acides nucléiques
~1660	Fort	νC=O amide I des protéines
1550–1520	Fort	δN-H amide II des protéines
~1515	Moyen	Bande de "Tyrosine" (cycle aromatique)
~1498	Faible	Bande de "phénylalanine"
~1468	Faible	δC-H de >CH2
~1400	Moyen	νC=O (sym) de COO^-
1310-1240	Faible	δN-H/νC-N amide III des protéines
1250-1220	Faible à moyen	ν_aP=O de $>PO_2^-$ phospodiesters
~1085	Faible à moyen	ν_sP=O de $> PO_2^-$ phospodiesters
1200-900	Moyen	C-O-C, C-O dominé par le cycle polysaccarides C-O-P, P-O-P

Tableau 4 – Bandes d'absorption des groupes fonctionnels bactériens dans le moyen infrarouge. Fréquences, intensités et modes de vibration. Inspiré de [171].

2.5. Conclusion sur la spectroscopie infrarouge

La spectroscopie infrarouge est une technique d'analyse rapide, performante et fiable. Elle permet de caractériser des composés et des fonctions chimiques. Elle permet aussi de déterminer la structure et l'environnement d'une fonction chimique ou d'une molécule entière. Ce dernier point est fondamental pour l'étude des structures de polypeptides ou de

protéines. Diverses structures secondaires ont une fréquence de vibration distincte contribuant à la bande d'amide I. La déconvolution et l'ajustement de courbe permettent de déterminer et de quantifier chacune des structures présentes. Cette technique peut être appliquée à des protéines en solutions, des protéines déshydratées ou directement à des protéines surexprimées dans la cellule bactérienne.

La spectroscopie IRTF couplée à la microscopie est capable de donner la même information à l'échelle locale et de dresser des cartes chimiques pour discriminer des régions dans une cellule ou dans un tissu, avec une résolution de l'ordre de quelques micromètres en utilisant une source synchrotron. Les dernières avancées ont permis de dépasser la limite de diffraction en travaillant avec des sondes locales.

La technique AFM-IR fait partie des techniques en sonde locale infrarouge fournissant des informations et des images chimiques d'une résolution nanométrique. Cela en fait un outil indispensable pour imager et acquérir des informations spectrales de bio-nanostructures, comme les fibres amyloïdes individuelles qui ont un diamètre variant de quelques nanomètres à quelques dizaines de nanomètres (voire quelques centaines de nanomètres pour les super-enroulements).

3. Dichroïsme circulaire

3.1. Introduction

La spectroscopie de dichroïsme circulaire (CD de l'anglais *Circular Dichroism*) est une technique d'absorption sensible très utilisée pour l'étude de structures biomoléculaires comme les protéines ou les acides nucléiques (ADN et ARN) en solution. Les premières études de solutions par lumière polarisée ont été effectuées, en 1811, par le physicien astronome et mathématicien français Jean-Baptiste Biot. Contrairement aux techniques de Résonance Magnétique Nucléaire (RMN) ou de cristallographie de rayon X, le DC ne donne pas de détails à l'échelle du résidu ou de l'atome. Il a cependant des avantages, comme la possibilité d'étudier la structure secondaire avec une faible quantité de matériel et dans différents milieux. Le CD est une des meilleures méthodes pour suivre toutes les altérations structurales qui pourraient résulter de changements de conditions telles que l'environnement chimique, le pH, la température, la force ionique ou la présence d'un ligand.

Nous avons utilisé le rayonnement synchrotron pour faire de la spectroscopie de dichroïsme circulaire (SRCD de l'anglais *Synchrotron Radiation Circular Dichroism*) sur la ligne de

lumière DISCO au synchrotron SOLEIL. L'avantage de la SRCD, par rapport à un spectromètre de laboratoire, est qu'il permet d'élargir la gamme spectrale jusqu'à 168nm en solution aqueuse et jusqu'à 120nm pour des films hydratés, avec un meilleur rapport signal sur bruit. Les informations de structures obtenues par SRDC sont donc plus nombreuses et plus précises [176]. Nous allons voir au cours de cette partie le fonctionnement et les informations obtenues par le CD.

3.2. Le principe de la spectroscopie CD

La spectroscopie de dichroïsme circulaire mesure la différence d'absorption d'un échantillon par un rayonnement polarisé circulaire droit et circulaire gauche.

Le rayonnement électromagnétique se compose d'un champ électrique (E) et magnétique (B) qui oscillent perpendiculairement l'un à l'autre suivant la direction de propagation k. Contrairement à une lumière polarisée linéairement où le vecteur de champ électrique n'oscille que dans un même plan, une lumière polarisée circulairement apparaît lorsque la direction du vecteur de champ électrique tourne autour de l'axe de propagation, comme une hélice, en conservant le vecteur E à amplitude constante. Le tour complet du cercle correspond à la longueur d'onde.

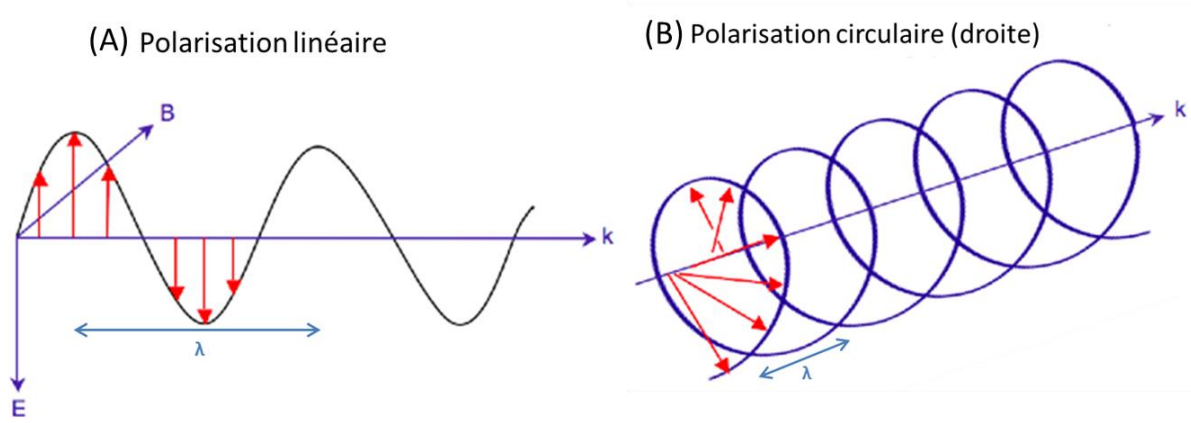


Figure 26 – Illustration des différentes polarisations d'un champ électromagnétique. Polarisation (A) linéaire et (B) circulaire de la radiation électromagnétique. E et B correspondent respectivement aux champs électrique et magnétique, k au vecteur de direction de propagation et λ la longueur d'onde. Adaptée avec authorisation de [177]. Copyright 1998 American Chemical Society.

L'absorption d'une lumière polarisée par un échantillon dépend, comme toute spectroscopie, du coefficient d'extinction molaire de l'échantillon epsilon, notée ε. Pour un échantillon dont

les molécules ont une énantiomérie (c'est-à-dire qui sont chirales par la présence d'atomes asymétriques, comme le carbone) la lumière polarisée circulaire droite ou gauche est absorbée différemment. Ces échantillons sont donc dits optiquement actifs et ont un coefficient d'extinction molaire distinct pour les lumières polarisées gauche, noté ε_g, et droite, noté ε_d. La différence entre les deux extinctions molaires, gauche et droite, est donnée par $\Delta\varepsilon$. La spectroscopie CD, comme toute autre spectroscopie d'absorption électronique, suit la loi de Beer-Lambert, et c'est la différence d'absorption entre la lumière circulaire polarisée gauche et droite (respectivement A_g et A_d) qui est mesurée et est notée ΔA. Le signal CD est donné par :

$$\Delta A = A_G - A_D = \Delta\varepsilon . c . l$$

Avec c la concentration de l'échantillon et l le trajet optique à travers l'échantillon.

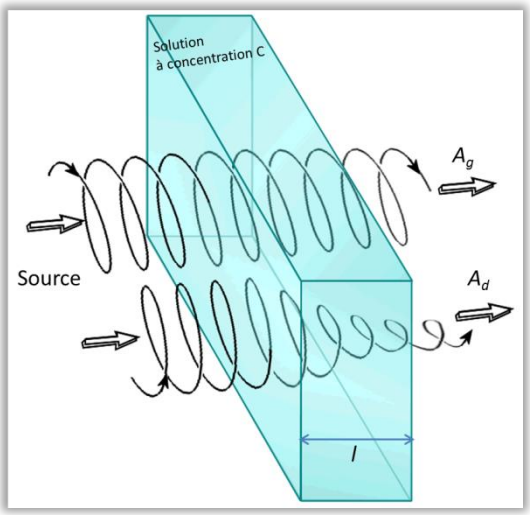


Figure 27 – Illustration de l'effet du dichroïsme circulaire. Deux rayons lumineux polarisés circulaire gauche et circulaire droite dont le rayonnement est absorbé différemment. (Respectivement A_g et A_d). Tirée et modifier de « Understanding Circular Dischroism (2017). In ScienceAid. Retrieved Aug 8,2018, from https://scienceaid.net/Understanding_Circular_Dischroism.

Lorsque Ag et A_d sont différents, on dit que le rayonnement résultant possède une polarisation elliptique. Cette différence d'absorbance ΔA, des deux composantes, est une mesure de CD, mais elle est généralement rapportée en termes d'ellipticité (θ - Theta) en degrés. Il existe une relation numérique simple entre le ΔA, et l'ellipticité en degrés qui est la suivante :

$$Ellipticité = \theta = tan^{-1}\left(\frac{b}{a}\right) = \frac{180ln10}{4\pi}(A_G - A_D) \text{ en degrés}$$

$$\theta \approx 3298.2(A_G - A_D) \text{ en millidegrés}$$

Avec *b* et *a*, les axes mineur et majeur de l'ellipse résultante.

Le spectre CD est obtenu en mesurant donc la différence d'ellipticité en fonction de la longueur d'onde.

Les mesures sur des molécules biologiques sont possibles du fait qu'elles présentent naturellement des carbones asymétriques d'un seul énantiomère. Les acides aminés des protéines existent presque uniquement sous la forme L (pour lévogyre, du latin *laevus* = gauche et *gyrare* = tourner). Dans le cas d'un mélange racémique (équimolaire) des deux énantiomères, on constaterait une absence de signal. À noter, des acides aminés peuvent être retrouvés naturellement sous leur forme D (pour dextrogyre, du latin *dexter* = droite), notamment chez les bactéries (dans le peptidoglycane ou toxines). Mais ils restent rares.

3.3. Détermination des structures secondaires par dichroïsme circulaire

3.3.1. Échantillons de protéines ou de peptides

Les peptides et les protéines sont observables dans les régions UV lointaines et proches.

L'absorbance, prédominante en CD dans les UV lointains compris entre 170 et 250 nm, est soumise à la contribution des liaisons peptidiques des fonctions amides, cette absorbance étant due entre autres à la présence des électrons délocalisés entre les atomes. Cette absorption correspond principalement à l'excitation de deux transitions : des orbitales n-π* et π-π*. La transition n-π*, dite transition magnétique, se produit généralement autour de 220 nm avec un faible signal et est polarisée dans la direction de la liaison carbonyle, mettant en jeu une paire d'électrons non liantes de l'oxygène. Cette transition est très sensible à l'environnement et donc aux conformations protéiques. Quant à la transition électrique π-π*, elle est généralement enregistrée autour de 190 nm et 210 nm et est polarisée dans la direction donnée par la liaison carbone azote [178].

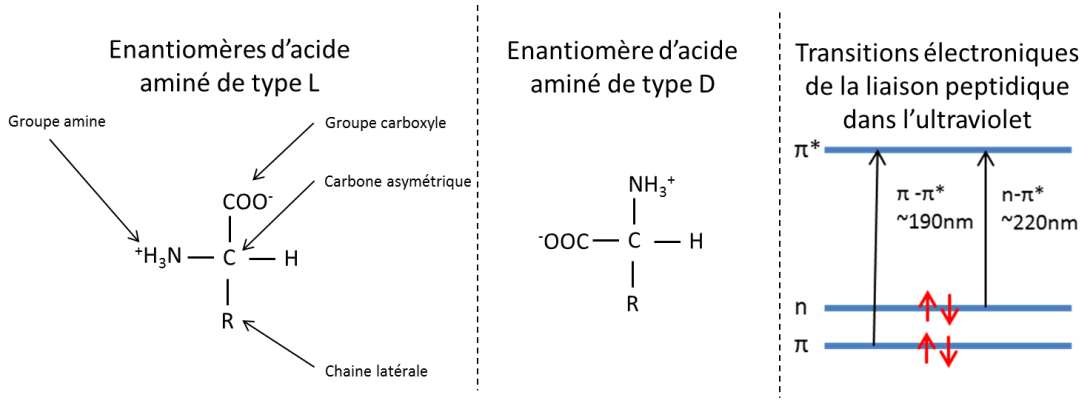


Figure 28 – Présentation des énantiomères (L) et (D) des acides aminés, ainsi que du diagramme des transitions électroniques de la liaison peptidique dans l'UV.

Les orbitales liante π, anti-liante π* et non liante n sont impliquées.

Les chaînes latérales contenant des acides aminés aromatiques et les ponts disulfures, quant à eux, absorbent dans le proche UV (entre 260 et 320nm) avec un faible signal CD.

Les données CD sont précieuses pour observer la conformation des peptides et protéines et les étudier en présence « de contraintes » (ligands, température, sels) [179]. À partir de ces spectres, il est possible de déterminer la proportion des différents éléments de la structure secondaire (hélices α, feuilles β, etc.) dans la protéine. La structure secondaire des protéines peut être déterminer en utilisant une base de données de spectres de protéines de référence. Il est aussi possible de procéder par calculs *ab initio*, de nombreux algorithmes existent pour l'estimation de la composition de la structure secondaire à partir des spectres DC, utilisant le plus souvent la décomposition en valeurs propres. L'algorithme Bestsel (pour *best structure selection*) développé en collaboration entre la ligne DISCO et le laboratoire de recherche en neuro-immunologie de Université Eötvös Loránd (Budapest,Hongrie) est utilisable en ligne sur internet (http://bestsel.elte.hu/). Cet algorithme a été développé pour déterminer quantitativement la structure secondaire et principalement la présence de feuillets β, plus compliquée à déterminer en raison de la grande diversité morphologique et spectrale des structures β. En effet l'orientation parallèle / antiparallèle et la torsion des feuillets β provoquent une diversité de réponses spectrales en CD. En tenant compte de la torsion des structures β, Bestsel sélectionne avec précision la structure secondaire pour une large gamme de protéines, en particulier pour les protéines riches en feuillets β et les fibrilles amyloïdes [180].

Cependant il est assez simple de faire une première estimation lorsqu'une structure domine (figure 29). Pour la pelote statistique, un pic négatif d'ellipticité vers 200nm est observé. Un

pic positif et deux négatifs (respectivement à 193 nm, 208 nm et 222 nm) sont observés pour les hélices α. Deux pics moins intenses, un positif vers 195nm et un négatif vers 218 nm sont observés pour les feuillets β [181] [178].

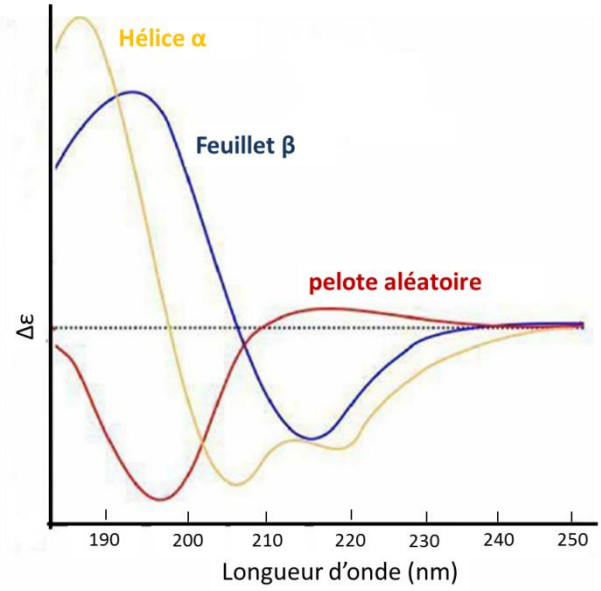


Figure 29 – Réponses des différentes structures secondaires "pures" obtenues en CD. Tirée de [181].

3.3.2. Échantillons d'acides nucléiques

La région UV utilisée pour l'étude des ADN se situe entre 180 à 300 nm, là où les bases de l'ADN absorbent la lumière (10). La spectroscopie CD dans la région UV fournit des informations sur les propriétés conformationnelles des acides nucléiques en solution.

La lumière absorbée par l'ADN est principalement due à la transition π-π* des bases. Intrinsèquement, les bases de l'ADN (purine et pyrimidine) n'ont pas de signal CD car les bases sont planes et donc achirales. Cependant, les bases acquièrent une asymétrique dans leurs transitions électroniques par le couplage avec les riboses présents dans les ADN. La différence de chiralité entre la purine et la pyrimidine est observable respectivement par le signal CD négatif et positif vers 270 nm.

Néanmoins, l'utilisation de la spectroscopie CD pour l'étude de la conformation de l'ADN reste principalement comparative ou empirique, pour l'étude d'interaction ou de dénaturation. Par exemple, il est possible de distinguer la conformation de différents ADN qui incluent les structures de la famille B, A, Z, les quadruplexes guanine, et certaines triple hélices car chacune présente des spectres CD caractéristiques [182]. Des méthodes

chimiométriques ont permis la classification de ces structures d'ADN à partir de données de dichroïsme circulaire. Pour ce faire, un ensemble de données comprenant 50 spectres expérimentaux correspondant à différentes structures d'ADN (bobine aléatoire, duplex, épingle, triplex inversé et normal, G-quadruplex parallèle et antiparallèle…) a été analysé par des méthodes non supervisées (ou non discriminantes) comme l'analyse en composantes principales [183].

3.4. Conclusion

Principalement utilisé pour étudier l'évolution d'un système en solution, le CD permet également d'étudier la structure secondaire des biomolécules. Le CD est un outil efficace et sensible pour obtenir rapidement des informations sur l'effet d'interaction entre molécules ou de dénaturation par l'augmentation de température, de changement de pH ou de force ionique.

4. Techniques complémentaires pour sonder les amyloïdes

4.1. La microscopie électronique en transmission

Comme pour la microscopie à force atomique (AFM), la microscopie électronique à transmission (MET) permet d'étudier la morphologie des fibres amyloïdes qui sont généralement rectilignes, non ramifiées, d'environ 10 nm de diamètre et qui peuvent présenter différents niveaux de surenroulement.

La microscopie électronique en transmission a été mise au point par Max Knoll et Ernst Ruska vers 1931. Pour cette invention, Ernst Ruska a reçu en 1986 le prix Nobel de physique. Le principe consiste à focaliser un faisceau d'électrons traversant un échantillon. L'énergie du faisceau d'électrons (entre 40 à 400 keV) améliore la résolution en diminuant la distance « *d* » observable entre deux objets d'un facteur 10^3 par rapport à un microscope optique. Par exemple, on peut distinguer deux objets d'une distance de 0.2 µm par la microscopie optique et de 0.2 nm par la microscopie électronique. Bien qu'opérant avec des électrons, la microscopie MET suit les mêmes lois d'optique et la résolution spatiale obéit à l'équation d'Abbe rappelée partie matériel et méthode au paragraphe 2.3.1.

Le faisceau d'électrons est absorbé ou transmis, ce qui permet d'obtenir une image contrastée et d'observer des objets nanométriques. Les électrons peuvent être déviés ou diffusés (phénomène inélastique), ils peuvent être rétrodiffusés (phénomène quasi-élastique). Des électrons secondaires peuvent aussi être observables par microscopie électronique à balayage.

Le faisceau d'électrons est produit par effet thermo-ionique au niveau d'une cathode (généralement un fil de tungstène) et est accéléré par une anode. Des lentilles électrostatiques et électromagnétiques concentrent les électrons. Le faisceau d'électrons sort de l'échantillon et donne des informations sur la structure de l'échantillon qui sont amplifiées par le système de lentilles de l'objectif du microscope. Les électrons sont absorbés par les molécules gazeuses et atmosphériques, ce qui impose de travailler sous vide. L'image peut être obtenue au moyen d'un système optique ou d'une fibre optique vers le capteur d'une caméra CCD (*Charge Coupled Device*).

Les méthodes de préparation d'échantillon en MET nécessitent une intention particulière, notamment pour fixer l'échantillon sur une surface, généralement une grille avec un revêtement de carbone. Pour les fibres amyloïdes, il est préférable de réaliser une décharge

luminescente à l'aide d'un *plasmacleaner* (processus d'ionisation des gaz donnant au film de carbone un caractère hydrophile) afin d'optimiser l'adsorption de l'échantillon. Pour leur observation, on utilise en général un agent de contraste. S'il est vrai que la coloration négative est l'outil de choix pour une caractérisation rapide, cela implique la déshydratation de l'échantillon, ce qui peut conduire à des artefacts, agrégations de fibres ou modifications structurales.

Une récente avancée de la technique consiste à utiliser la cryo-microscopie électronique (Cryo-EM), ce qui permettrait d'éviter ce problème d'artéfact dû à l'utilisation du colorant, mais aussi de gagner en résolution. Trois scientifiques ont reçu le Prix Nobel de Chimie en 2017 pour avoir développé cette méthode : Jacques Dubochet, Joachim Frank et Richard Henderson. Cette méthode permet de déterminer la structure à haute résolution des protéines en solution et est basée sur une congélation rapide, permettant la formation de glace vitreuse transparente aux électrons contrairement à la glace. La Cryo-EM permet de la figer la structure des protéines dans de l'eau vitrifiée. Cette technique a permis de déterminer à haute résolution (inférieure à 4 Å) de nombreuses structures comme la structure du virus Zika [184] où récemment, la structure de protofilaments amyloïdes du peptide Aβ et la protéine tau impliqués dans la maladie d'Alzheimer [185].

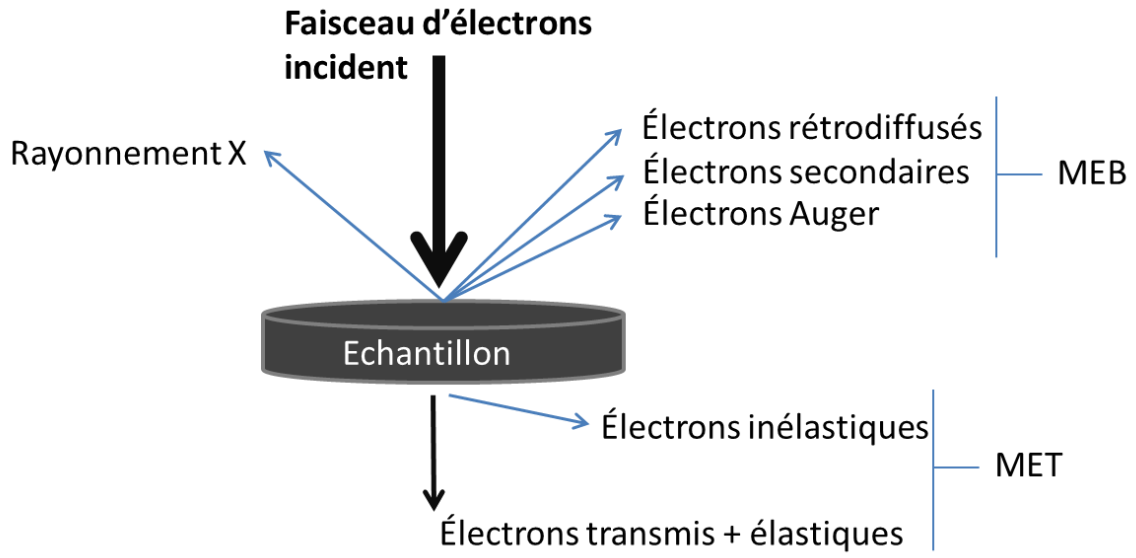


Figure 30 – Interaction électrons/matière. Types d'électrons observés après interaction entre le faisceau d'électrons incident et l'échantillon étudié par microscopie électronique à balayage (MEB) et à transmission (MET).

4.2. La diffusion de rayons X aux petits angles (SAXS)

Dans le domaine de la biologie structurale, la diffusion de rayons X aux petits angles (SAXS) a connu une évolution considérable au cours des deux dernières décennies. Le rayonnement synchrotron a joué un rôle clé dans cette évolution, en fournissant des faisceaux de rayons X intenses de haute qualité optique qui ont rendu possible l'acquisition relativement rapide et l'enregistrement de données statistiquement significatives à partir de solutions biologiques faiblement dispersées.

Le principe physique du SAXS est basé sur la diffusion élastique de photons X interagissant avec les nuages électroniques de l'échantillon L'observation fournit des informations sur la fluctuation des densités électroniques. Le motif de diffusion est collecté en utilisant un détecteur bidimensionnel placé derrière l'échantillon et perpendiculaire au faisceau incident. L'intégration radiale du motif de diffusion (enregistré en image 2D), est effectuée afin d'obtenir des courbes SAXS. Les courbes représentent l'intensité en fonction de « q », le vecteur d'onde : $q = \frac{4\pi sin\vartheta}{\lambda}$ dont ϑ est l'angle de diffusion et λ est la longueur d'onde. L'intensité de diffusion diffractée dépend de plusieurs variables et s'exprime de façon simplifiée :

$$I(q) = \Delta\rho^2 \,.\, \Phi V \,.\, VP \,.\, P(q) \,.\, S(q)$$

$\Delta\rho^2$ est le contraste, ΦV est la fraction volumique des particules, VP le volume d'une particule sèche, $P(q)$ est le facteur de forme de la particule qui ne dépend que de sa géométrie et $S(q)$ est le facteur de structure qui décrit les corrélations entre particules.

Lorsque des macromolécules isolées s'assemblent en structures régulières, de nouvelles corrélations apparaissent dans la densité électronique de la structure entière, donnant lieu à des modulations spécifiques de l'intensité de diffusion. Mais lorsque les motifs sont répétés le long d'un axe défini avec une distance périodique d, les ondes de photons diffusés subissent de fortes interférences constructives aux valeurs « q », données par $q = n\frac{2\pi}{d}$, où n est un entier (loi de Bragg), c'est le cas pour l'assemblage de feuillets β amyloïdes. Le phénomène est appelé diffraction. Lorsque des unités identiques, par exemple un peptide ou une protéine, sont régulièrement espacées avec des périodicités typiquement de l'ordre du nanomètre, une diffraction est observée dans la gamme des petits angles.

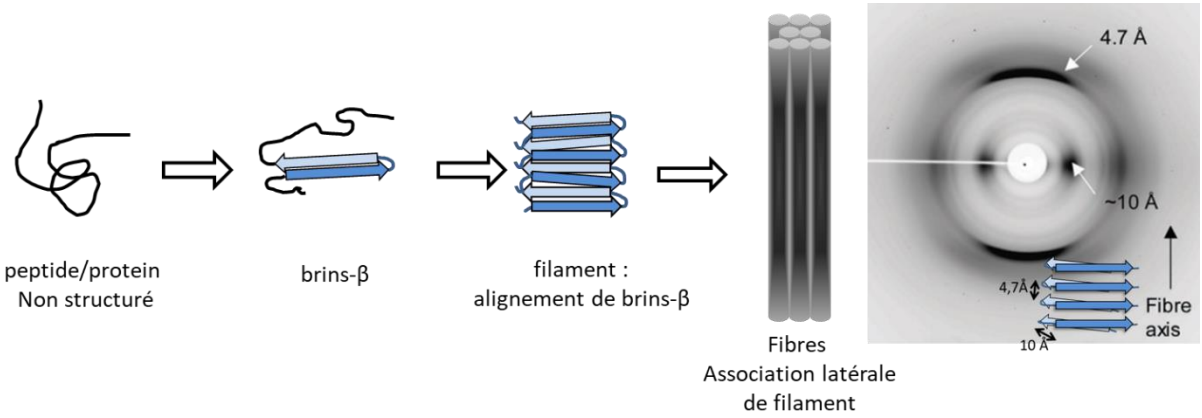


Figure 31– Motif typique de diffraction des fibres amyloïdes par les rayons X. Les réflexions anisotropes sont indicatives d'une fibre partiellement alignée, perpendiculaire au faisceau de rayons X. Les réflexions à d = 4,7 Å et à d = 10 Å correspondent respectivement à l'espacement entre les brins β et entre les feuillets β. Tirée de [186].

4.3. Spectroscopie de fluorescence : Thioflavine T colorant des amyloïdes

La détection de présence d'agrégats amyloïdes est possible en étudiant la fluorescence indirecte de la ThT. L'intensité du signal augmente avec la quantité de fibres. Cette façon d'étudier est historiquement l'une des premières analyses utilisées pour sonder la cinétique de l'agrégation des protéines et demeure l'une des approches expérimentales les plus utilisées aujourd'hui. Cependant, malgré son utilisation répandue, il reste difficile d'obtenir des données suffisamment quantitatives par les tests de fluorescence ThT. Cette difficulté est due en grande partie au fait que les caractéristiques de la liaison de la ThT aux fibres amyloïdes ne sont pas complètement connues et que l'intensité de fluorescence est susceptible de subir des perturbations dues à la présence d'impuretés ou d'agrégats de protéines amorphes et se déposant au fond de la solution. Cependant le signal de fluorescence dépend linéairement de la concentration des agrégats dans des conditions optimisées selon chaque étude.

Une autre méthode de suivi de concentration en agrégats est possible en spectroscopie de fluorescence sans marquage en suivant la modification spectrale spécifique et intrinsèque lors de la formation d'amyloïdes [187][188]. Mais le signal reste faible et la méthode n'est pas toujours aisée à mettre en œuvre.

Études : résultats et discussion

Chapitre I : caractérisation de la structure amyloïde du domaine CTR de Hfq

1. Identification de l'amyloïde du domaine C-terminal d'Hfq

1.1. Identification par un ensemble de techniques physico-chimiques

Le premier chapitre de cette thèse a pour but de présenter les différentes techniques pouvant être utilisées pour caractériser la nature amyloïde de l'auto-assemblage du domaine CTR de Hfq. Ces techniques ont été décrites dans un chapitre de livre de la série « Methods in Molecular Biology » des éditions Springer en 2018 (ISBN:978-1-4939-7633-1, DOI:10.1007/978-1-4939-7634-8_18) dans le cadre de cette thèse. Le chapitre de livre est disponible en **annexe A**.

Brièvement, les fibres du CTR de Hfq étudiées sont obtenues en utilisant un peptide synthétisé chimiquement (trente-huit derniers acides aminés de la protéine Hfq de *E. coli*, synthèse de type Fmoc, réalisée par la société Proteogenix©, notée 38aa CTR-Hfq). La synthèse chimique de cette séquence est le seul moyen d'étudier cette portion *in vitro*. Celle-ci est en effet impossible à purifier à partir de la bactérie, car elle est instable et dégradée par des protéases lorsqu'elle est exprimée seule (ce qui a été vérifié par Western Blot).

Après quelques jours en solution à une concentration assez forte (50mg/ml), les peptides formaient des fibres amyloïdes que ce soit dans l'eau milliQ ou dans une solution tampon à pH 7-8 (tampon phosphate ou HEPES). Par la suite, nous avons préféré travailler dans l'eau afin d'éviter les artéfacts de mesure pour l'ensemble des techniques détaillées ci-dessous, en particulier pour les mesures de dichroïsme circulaire très sensibles à la présence de sels. Les résultats présentés dans ce chapitre ont été obtenus avec différentes techniques complémentaires :

(A) Les techniques d'imagerie nanométrique comme la microscopie électronique en transmission (MET) avec coloration (coloration négative) ou sans (Cryo-EM), et l'utilisation de microscopie à force atomique (AFM solide ou liquide). Ce sont des techniques assez rapides pour vérifier la présence, les caractéristiques structurales (longueur, diamètre, entrelacement) ainsi que l'homogénéité des fibres.

(B) Les techniques spectroscopiques telles que la spectroscopie infrarouge et le dichroïsme circulaire. Elles permettent principalement d'étudier l'agrégation des

feuillets β et la présence d'un signal amyloïde au niveau moléculaire avec un signal intrinsèque ne nécessitant pas de perturber l'échantillon ou de le marquer.

(C) La technique de diffusion des rayons X aux petits angles (SAXS). Cette technique met en évidence un motif de diffraction de fibres typiques et la présence de facteur de forme, en diffusion, de long tube. À noter, la diffusion de neutrons peut aussi être utilisée mais les résultats que nous avons obtenus sont moins probants et nécessitent des mises au point.

(D) La microscopie et spectroscopie par marquage avec une sonde fluorescente de type Thioflavine-T (ThT).

L'ensemble de ces techniques nous a permis de caractériser les fibres formées. Après 30 jours en solution à une concentration de 50mg/ml, nous avons pu observer par microscopie de longues fibres de quelques micromètres de longueur et d'une dizaine de nanomètres de diamètre en moyenne ainsi qu'une réponse en spectroscopie IRTF et SRCD de feuillets β dits agrégés (figure 32).

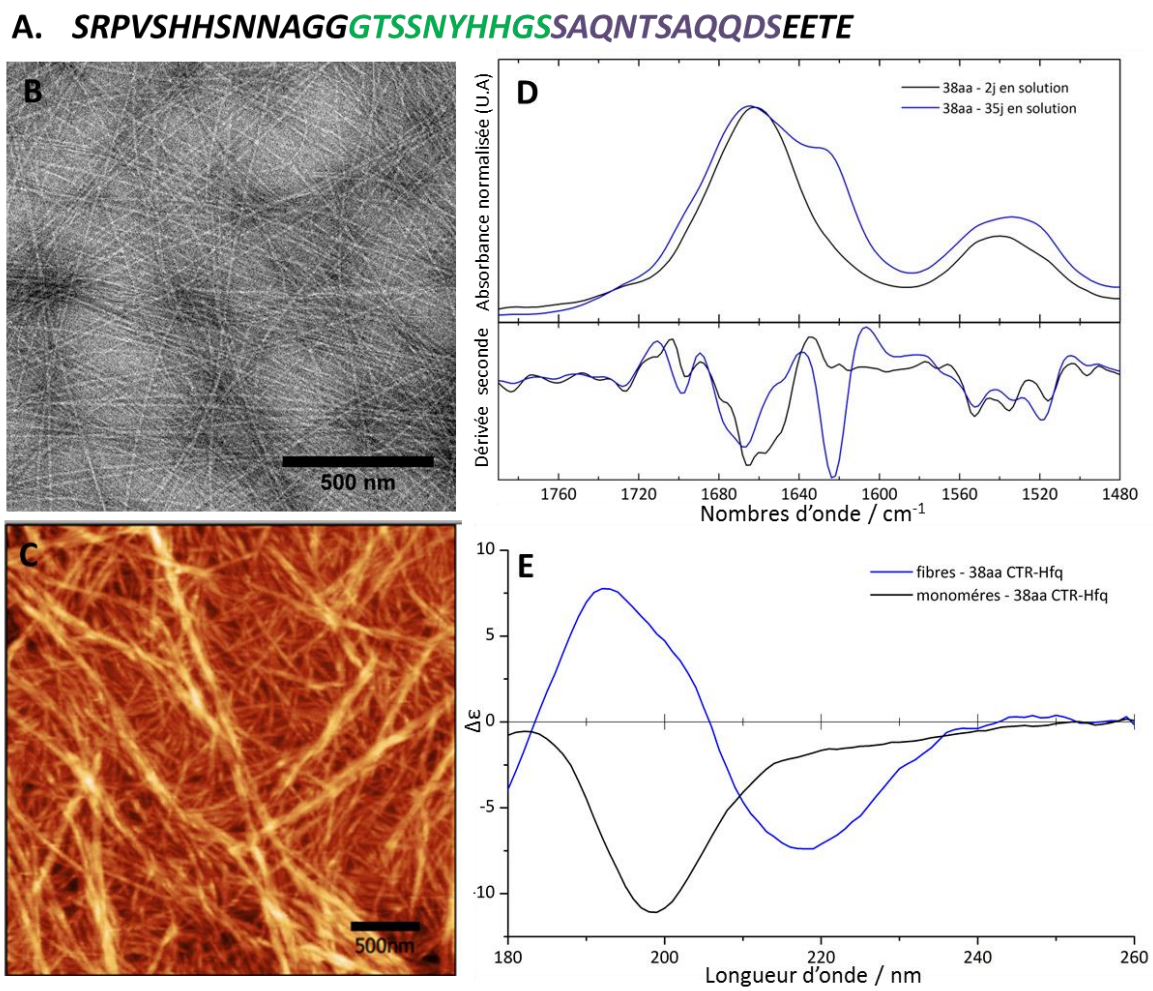


Figure 32 – Caractérisation de la structure amyloïde du domaine CTR-Hfq par quatre techniques complémentaires. (A) Séquence de trente-huit acides aminés du domaine CTR-Hfq. La séquence d'acides aminés en vert correspond à la zone déterminée d'interaction avec l'ADN. La séquence en violet correspond quant à elle à la zone indispensable pour la formation d'amyloïde (11aa). (B) Image MET obtenue en coloration négative du peptide 38aa CTR-Hfq. (C) Image AFM obtenue en mode contact. (D) Spectres IRTF (et dérivée seconde du domaine de 38aa) avant et après fibrillation. L'épaulement du spectre d'absorption et le pic de la dérivée se trouvent à 1624 cm^{-1}, signal typique des amyloïdes. (E) Spectres SRCD du CTR avant et après la fibrillation. *Remarque : à l'époque de la publication du livre nous n'avions pas encore pu mettre en évidence le signal amyloïde du 38aa CTR-Hfq en SRCD à 220 nm.*

Outre les analyses présentées dans ce livre, d'autres sont présentées brièvement ci-dessous.

1.2. Mesures cinétiques

1.2.1. Test en présence de Thioflavine T

La principale méthode pour étudier la cinétique de formation des fibres amyloïdes est en général le suivi de l'intensité de fluorescence de la Thioflavine T (ThT) en présence de fibres amyloïdes par spectroscopie de fluorescence. Nous avons donc utilisé cette méthode en première intention sur des fibres déjà formées pour observer la présence de signal. Cette analyse a d'ailleurs été publiée dans la référence Fortas *et al.* en 2015 [6].

Cependant, dans nos conditions expérimentales, nous avons observé une très rapide décroissance de l'intensité de fluorescence de la ThT en présence de fibres amyloïdes de Hfq-CTR (en quelques secondes) rendant la technique impraticable pour la réalisation de cinétiques d'agrégation. Nous formulons l'hypothèse que les fibres amyloïdes formées par Hfq-CTR sont dans un état d'équilibre monomère/fibre favorisant la formation de monomères et que les fibres sont déstabilisées lors de la dilution accompagnant l'ajout de ThT. La méthode ne serait utilisable que pour mesurer la concentration de solutions âgées contenant une haute concentration de fibres amyloïdes superenroulées et donc plus stables.

La faible quantité de peptide Hfq 38aa, disponible en matière et en volume, nous a conduits à utiliser des microplaques permettant de faire des mesures avec un volume minimal de suspension (10-20 µl) pour économiser le peptide. Lors de l'ajout de ThT, on diluait la solution entre 5 et 10 fois. Un fort signal de fluorescence à 480 nm était observé durant quelques secondes, caractérisant l'interaction de la ThT et des fibres amyloïdes. L'intensité de fluorescence diminuait ensuite rapidement entre quelques secondes à quelques dizaines de secondes. Cette extinction pourrait être causée par un quenching de la fluorescence dû à une concentration trop élevée mais des tests avec des dilutions plus fortes n'ont pas permis de réduire ce phénomène. Une autre hypothèse est que la ThT puisse désagréger les fibres ce qui a d'ailleurs été observé et décrit pour d'autres types d'amyloïdes [189][190]. Une autre hypothèse est que la simple dilution puisse déstabiliser les fibres et favoriser la formation de monomères d'Hfq-CTR. Si c'est le cas, la méthode ne pourrait être appliquée que pour des faisceaux de fibres déjà âgés, fortement polymérisés (présence de surenroulements) et très stables, comme ceux décrits par Fortas *et al* [6].

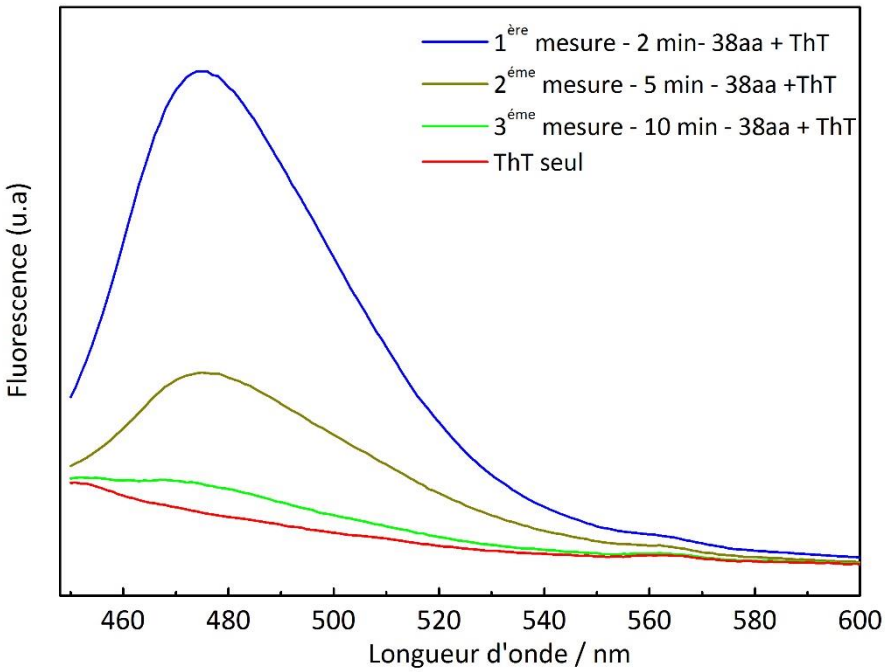


Figure 33 – Spectres de fluorescence d'une solution contenant des fibres du domaine CTR-Hfq en présence de ThT.

Nous avons donc étudié l'effet de la dilution par IRTF et par SRDC (figure 34) à partir de deux solutions initiales contenant une forte concentration en fibres amyloïdes. Nous observons bien la perte du signal cross-β de manière proportionnelle à la dilution.

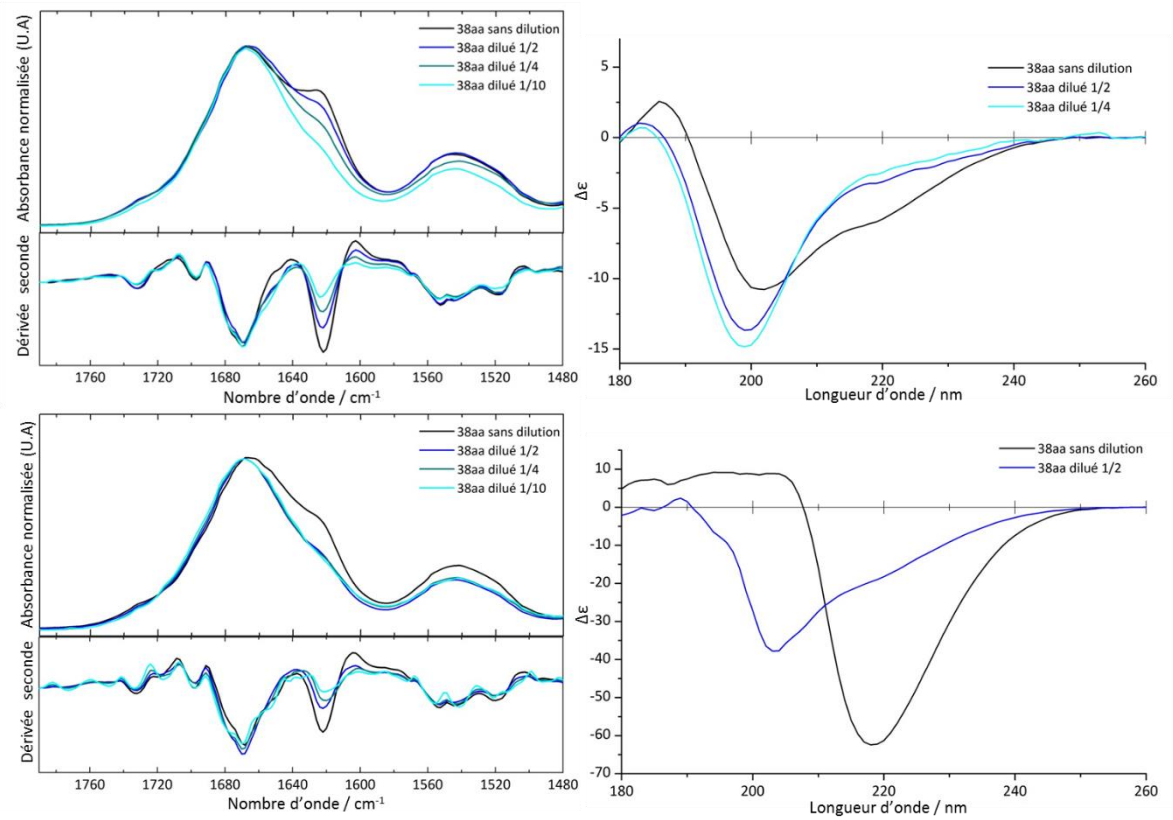


Figure 34 – Spectres IRTF et SRDC de solutions de fibres du peptide 38aa CTR-Hfq à différentes dilutions.

Nous avons évalué l'effet de la présence de ThT sur la formation de fibres amyloïdes Hfq-CTR par méthode SAXS. Nous observons la perte du pic de diffraction caractérisant la structure de feuillets β croisés en présence de ThT (figure 35).

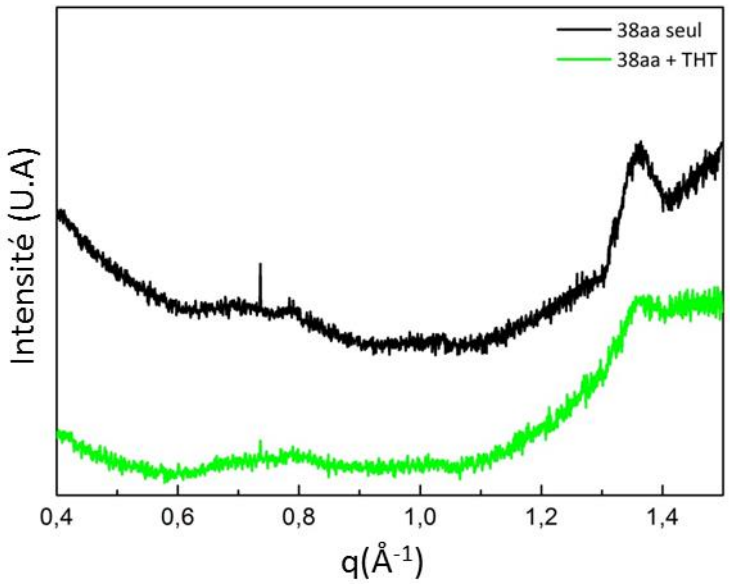


Figure 35 – Courbes SAXS du peptide CTR-Hfq en présence ou absence de ThT. Pics de feuillets β observés à 0,79 Å-1 et 1,37 Å-1 correspondant à une structure β croisée observée pour le peptide du domaine CTR (incubé pendant six semaines à 20 mg/mL). On observe, dans les mêmes conditions d'incubation, (concentration en peptide et durée) que la présence de ThT a un effet sur la structure en empêchant la formation des feuillets β.

Nos observations montrent que les fibres amyloïdes formées par le peptide Hfq-CTR 38aa peuvent être facilement déstabilisées par simple dilution ou par ajout de molécules inhibitrices de l'agrégation. La ThT qui est généralement utilisée comme un marqueur de la concentration des fibres amyloïdes (mais dont l'effet déstabilisant a aussi été rapporté) n'est donc pas adaptée pour suivre la cinétique de formation de ces fibres.

Nous avons donc choisi d'utiliser la spectroscopie IR pour mesurer la cinétique de formation des fibres en raison de la simplicité de la méthode (bien que plus ardue à mettre en œuvre en solutions), de la spécificité de la signature amyloïde en IR et de sa sensibilité élevée.

1.2.2. Cinétiques en spectroscopie IRTF

Les mesures à sec ont été réalisées à différentes concentrations de peptide 38aa, à différentes forces ioniques par ajout de sel (NaCl) et à différentes températures (25°C et 4°C). Dans les conditions sans sel et à température ambiante, nous observons, pour une concentration de 50mg/ml, une augmentation au cours du temps du signal amyloïde avec un pic ou épaulement à 1624 cm^{-1}. À 100mg/ml, les fibres se forment encore plus rapidement (figure

36). En revanche, pour les concentrations 10 et 20 mg/ml, aucun signal amyloïde n'a été observé durant le laps de temps de l'expérience c'est à dire jusqu'à 45 jours en solution.

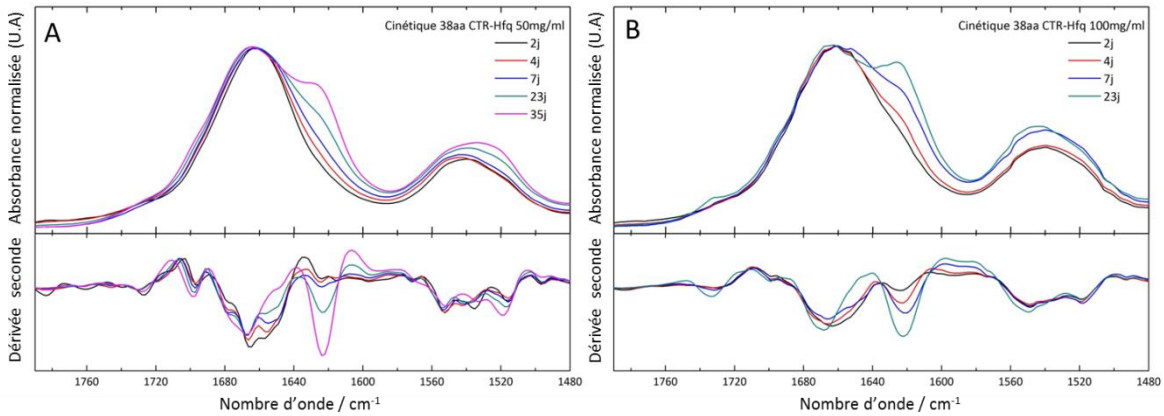


Figure 36 – Suivi par IRTF de la cinétique de formation des fibres amyloïdes du domaine CTR-Hfq durant 35 jours.

L'évolution de l'épaulement d'absorption et celle du pic pour la dérivée seconde à 1624 cm^{-1} sont observables sur 35 et 23 jours respectivement pour les concentrations de 50 mg/ml (A) et 100 mg/ml (B).

Le tableau ci-dessous montre la présence (en vert) ou l'absence (en rouge) du signal vers 1620 cm^{-1} après 35 jours en solution de peptide 38aa dans différentes conditions, mesurées par IRTF.

Concentration / Conditions	10mg/ml	20mg/ml	50mg/ml	100mg/ml
4°C				
4°C avec 50mM de NaCl				
4°C avec 250mM de NaCl				
25°C				
25°C avec 50mM de NaCl				
25°C avec 250mM de NaCl				

Tableau 5 – Présence ou absence des fibres amyloïdes dans différentes conditions de température et concentrations en sels. Détection *in vitro* des structures amyloïdes par spectroscopie infrarouge du peptide CTR-Hfq.

Ces mesures nous ont permis de confirmer que le domaine CTR-Hfq forme des structures de type amyloïde en solution. La concentration, la température et la pression osmotique semblent également favoriser la cinétique de formation de fibres par ce domaine. Cependant, nous avons vu que les fibres sont fragiles et qu'elles peuvent facilement se désagréger, par dilution ou par la présence d'une molécule interférant avec leur formation comme la ThT.

1.2.3. Analyses *in vivo*

1.2.3.1. Souche BL21(DE3)

Suite aux résultats montrant par immunofluorescence que Hfq forme des structures sous-membranaires, nous avons voulu vérifier si l'on pouvait détecter *in vivo* le signal IRTF associé aux structures β amyloïdes. Pour cela, nous avons comparé trois souches de la bactérie *E. coli*. Ces trois souches sont délétées du gène *hfq* et ont été transformées par un plasmide permettant la synthèse de Hfq (promoteur inductible). Ces plasmides apportent tous deux la résistance à l'ampicilline, mais les promoteurs de transcription en amont du gène *hfq* sont différents. Les premières souches transformées sont de type BL21(DE3), permettant une forte surexpression de Hfq à partir du promoteur T7 en amont du gène *hfq* dans le plasmide. À noter, la souche BL21(DE3) est elle-même génétiquement modifiée (voir ci-dessous). Les bactéries BL21(DE3) transformées sont cultivées en présence d'ampicilline (pour maintenir la pression de sélection, maintenir le plasmide pTE608) : la culture doit être menée jusqu'en phase exponentielle. La ARN polymérase T7 est spécifique d'un promoteur de transcription et ne transcrit que l'ADN en aval de ce promoteur T7 (de

séquence TAATACGACTCACTATAG) ; elle est de plus extrêmement efficace. La particularité de la souche BL21(DE3) de *E. coli* est d'avoir dans son génome cette polymérase T7 (qui est d'origine virale) sous contrôle du promoteur P_{Lac} inductible par l'IPTG (IsoPropyl-β-D-1-ThioGalactopyranoside). On induit alors l'expression de l'ARN polymérase T7 avec de l'IPTG : l'ajout d'IPTG va permettre indirectement la surproduction de la protéine Hfq puisque dans le pTE608 le gène *hfq* codant pour la protéine Hfq est sous contrôle du promoteur T7. Après induction (environ 3 heures) on récupère les cellules ayant surexprimé fortement Hfq. Différents plasmides sont utilisés :

(A) La souche BL21(DE3) Δ*hfq* + pTE608. Ce plasmide permet d'exprimer fortement la protéine Hfq-WT d'*E. coli* grâce au promoteur viral T7.

(B) La souche BL21(DE3) Δ*hfq* + pTEHfq-NTR72, contenant un plasmide où seule la partie NTR (les 72 premiers acides aminés, notée Hfq-NTR72) de la protéine Hfq est exprimée et peut être induite.

(C) La souche BL21(DE3) Δ*hfq* + plasmide vide (souche contrôle). Absence totale de Hfq.

Pour l'analyse par spectroscopie IR, nous avons acquis des spectres IR en transmission, en déposant et laissant sécher les cellules sur une surface de CaF_2. Pour chaque dépôt, nous avons effectué une trentaine d'acquisitions en utilisant un microscope infrarouge Thermo Scientific IN10 avec une ouverture de 100x100 µm². D'autres mesures ont été effectuées en mode ATR sur un cristal en diamant de plusieurs mm² de façon à mesurer une large population de plusieurs millions d'individus. L'expérience a été répétée pour plusieurs cultures indépendantes.

Nous avons effectué sur l'ensemble des spectres une analyse des composantes principales (ACP) à l'aide du logiciel « The Unscrambler X 10.3 » pour déterminer les variabilités entres les différentes souches. L'ACP est une méthode d'analyse statistique multivariée permettant d'analyser la variance d'un ensemble d'observations (ici des spectres) contenant de multiples variables (les nombres d'onde) de façon objective. L'ACP permet de résumer la structure (la présence de groupes, la similarité entre spectres) de données multidimensionnelles sous la forme de graphiques en 2 ou 3 dimensions montrant la distance entre les spectres. L'ACP permet aussi d'extraire l'information spectrale expliquant la structure des données sous la forme de composantes principales, des vecteurs qui sont des combinaisons linéaires des

variables d'origine. Ces composantes sont orthogonales et permettent de capturer des sources de variabilité indépendantes.

Les ACP réalisées sur les spectres des 3 cultures indépendantes des 3 souches BL21(DE3) montrent une variabilité entre les souches A surexprimant Hfq WT et les souches B et C. Cette variabilité est expliquée par la composante principale 1 qui capture environ 90% de la variance spectrale et présente systématiquement un pic à 1624cm^{-1}. Il est associé à la souche A et caractéristique des amyloïdes obtenues pour le peptide 38aa CTR-Hfq. Les résultats ont été obtenus pour trois cultures indépendantes et sont présentés dans la figure 37. La position du pic amyloïde à 1624 cm^{-1} est exactement la même que dans le cas des peptides produits *in vitro* avec le peptide CTR seul. Cela suggère une structure très semblable et élimine la possibilité de la détection d'une agrégation protéique intermoléculaire non-amyloïde (il arrive en effet que les protéines fortement surexprimées dans les bactéries précipitent en s'agrégeant de façon non spécifique [172]). Ces résultats montrent qu'il est possible de détecter la formation de fibres amyloïdes directement dans les bactéries surexprimant fortement la protéine Hfq WT.

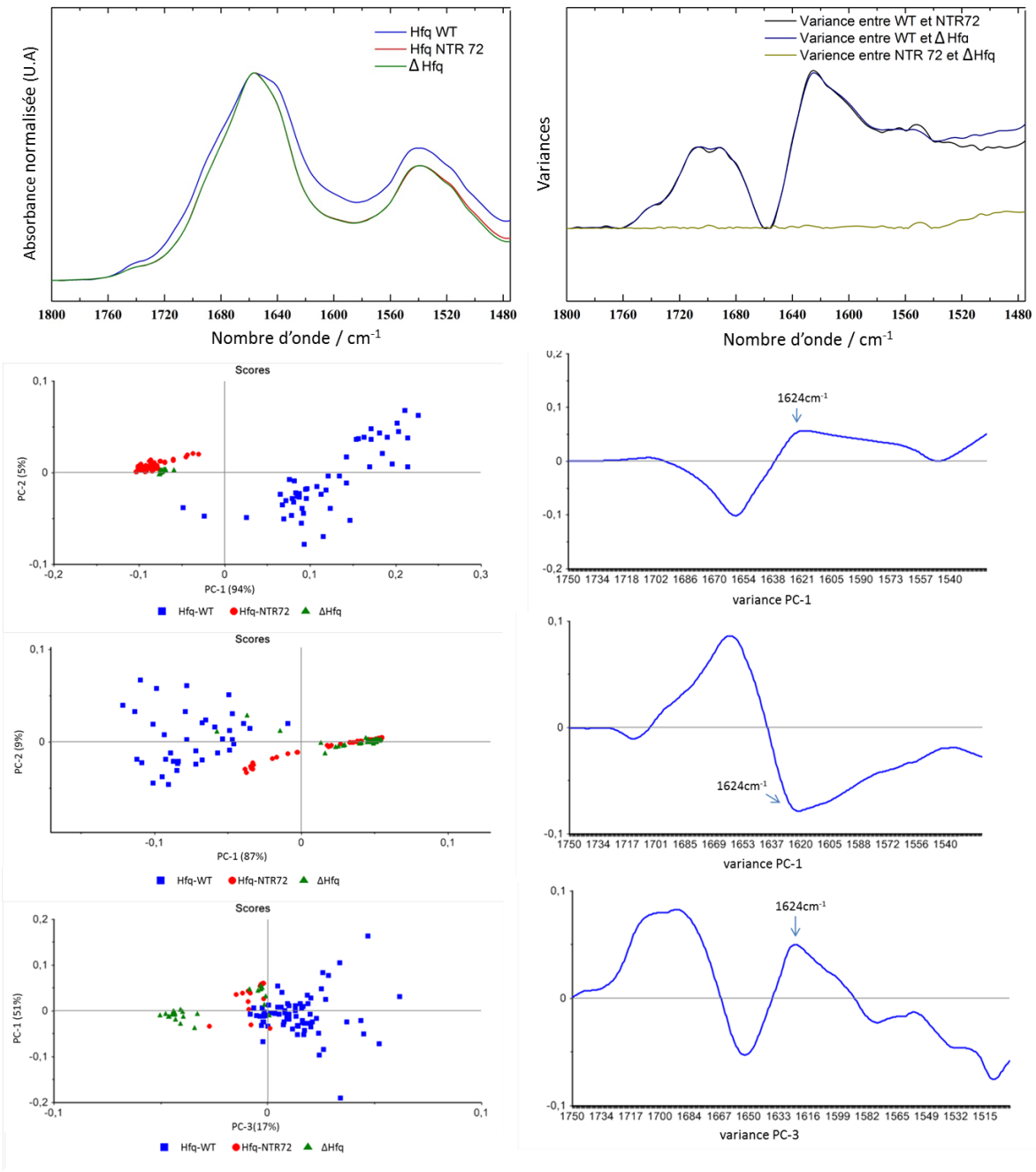


Figure 37 – Moyennes, variances et ACP des spectres infrarouges pour les différentes souches BL21(DE3). Chaque point correspond à un spectre. Les points bleus correspondent à Hfq-WT, les points rouges à Hfq-NTR72 et les points verts à la souche dépourvue de Hfq (Δhfq).

1.2.3.2. Souche MG1655

La protéine Hfq est fortement surexprimée dans la souche BL21(DE3), beaucoup plus que dans ses conditions natives, ce qui induit des changements morphologiques de la cellule, probablement dus à la dérégulation de l'expression de la protéine FtsZ [24][191]. Nous avons donc souhaité utiliser une souche exprimant moins fortement Hfq, pour être plus proche des conditions physiologiques. Ceci est possible grâce à l'utilisation d'un plasmide de type pBAD contenant un promoteur inductible à l'arabinose. On peut ainsi induire la production de Hfq WT ou NTR72, à des niveaux proches des conditions physiologiques en choisissant la concentration adaptée d'arabinose [67]. Les spectres IR de 3 souches MG1655 exprimant Hfq-WT ou Hfq-NTR72 ou n'exprimant pas Hfq ont été mesurés par spectroscopie ATR en présence d'arabinose pour obtenir une expression modérée d'Hfq n'induisant pas de modification morphologique des cellules. Nous n'avons pas observé de différences spectrales entre les deux souches par examen visuel des spectres et nous n'avons pas non plus observé de variation avec l'ACP ou par spectroscopie de différence. Une toute petite variation a été observée mais elle n'est pas forcément significative (figure 38).

2.

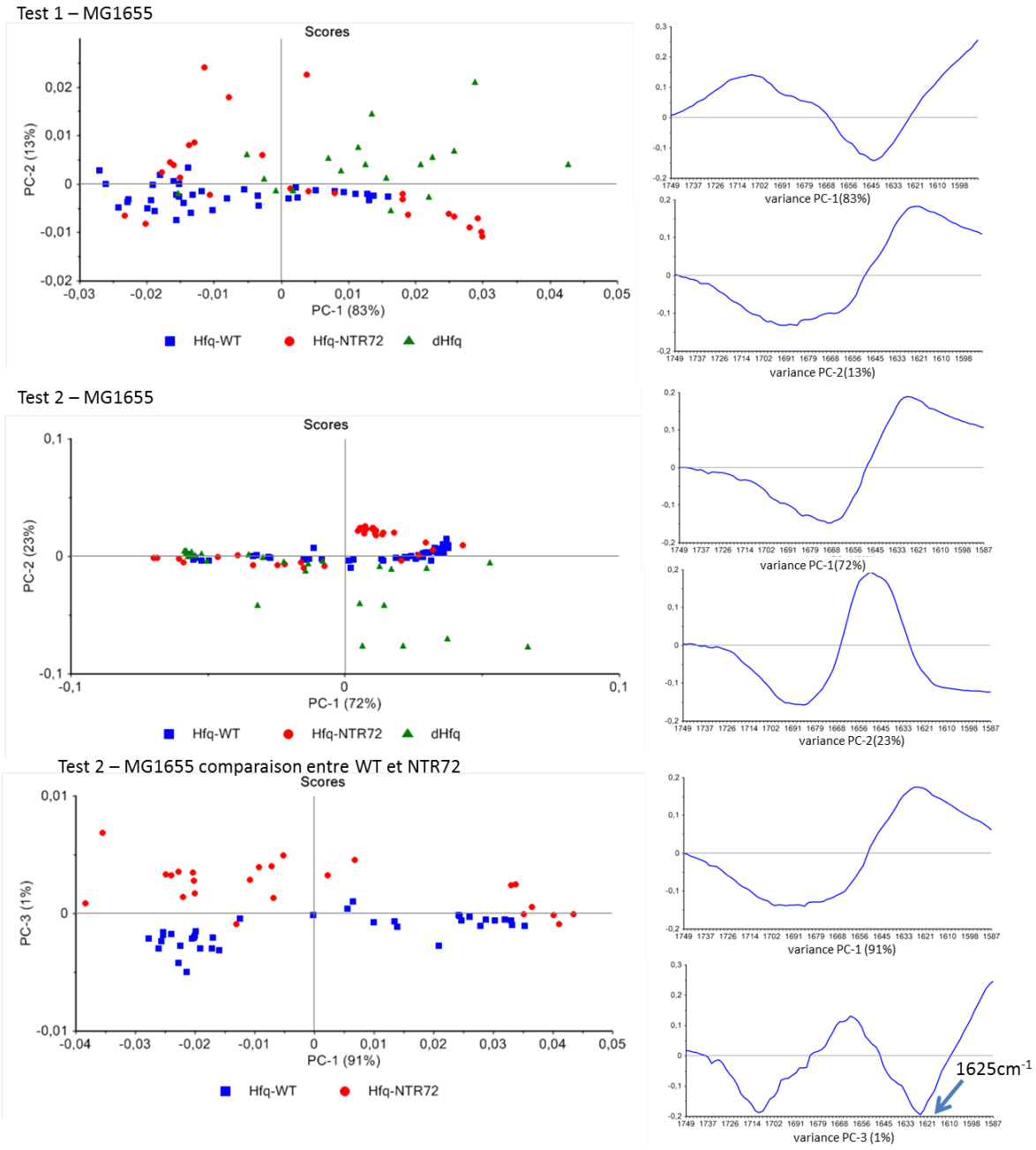


Figure 38 – ACP des spectres infrarouges pour les différentes souches MG1655. Chaque point correspond à un spectre. Les points bleus correspondent à Hfq-WT, les points rouges à Hfq-NTR72 et les points verts à Δhfq.

2. Discussion

Ces séries d'expériences nous ont permis de confirmer la capacité du domaine CTR-Hfq à s'auto-assembler en fibres amyloïdes *in vivo* et *in vitro* en solution aqueuse. Le domaine CTR de Hfq de *E. coli* est donc capable de se structurer spontanément. Les nouveaux résultats IRTF renforcent l'hypothèse émise sur la nature amyloïde des fibres observées précédemment par immunofluorescence [6]. Cependant, on peut se demander comment ces feuillets cross-β se forment spontanément en conditions physiologiques. En effet, il a fallu de fortes concentrations (forte expression de Hfq ou forte concentration de peptide) pour obtenir une réponse amyloïde. Bien que des études montrent que Hfq se concentre près des membranes ou dans le nucléoïde [69][67], l'on peut supposer que la concentration de Hfq serait insuffisante pour former ce genre de structures en absence d'un cofacteur.

Hfq est capable d'interagir *in vivo* avec de nombreuses autres molécules comme des protéines, les ARNm et sARN. Comme montré plus récemment, elle est capable d'induire la condensation de l'ADN au-delà d'une concentration seuil. Cette concentration seuil peut être abaissée par la présence de NaCl [19]. La formation d'amyloïdes par Hfq étant aussi influencée par la concentration en NaCl, on peut faire l'hypothèse que les fibres pourraient être impliquées dans la condensation de l'ADN induite par Hfq et que la présence d'un acide nucléique pouvait en retour favoriser la formation de ces structures. L'étude de cette hypothèse constitue le chapitre III (partie Étude) de cette thèse.

En conclusion à ce chapitre, nous avons mis en évidence que :

- Thermodynamiquement, la formation des fibres amyloïdes par le domaine Hfq-CTR suit la loi classique, c'est-à-dire que la formation d'amyloïdes est favorisée par la concentration de peptides/protéines.

- La présence de sel (NaCl) favorise la formation de fibres amyloïdes. Ce résultat va dans le sens d'une précédente étude [19] montrant que la concentration seuil de Hfq provoquant la compaction de l'ADN est plus faible en présence de NaCl.

- La dilution et l'ajout de ThT semble désagréger les fibres amyloïdes, peut-être particulièrement les fibrilles individuelles initiales.

En examinant de plus près la séquence du peptide du domaine CTR, il est difficile de faire une hypothèse concernant la formation de ces fibres. Avec seulement quatre acides aminés hydrophobes sur 38, le peptide est majoritairement hydrophile avec une première partie

positive (composée de quatre histidines et d'une arginine) et une zone C-terminale négative (composée d'un aspartate et de trois glutamates). Notre équipe a montré qu'une partie contenant 11 acides aminés séparant ces deux zones chargées est la séquence minimale pour la formation de fibres. Cette séquence est extrêmement amyloidegénique [71], et crée l'interaction centrale stabilisée par les charges (liaisons statiques). La partie positive N-terminale se trouvant près du tore ne devrait pas jouer un rôle majeur sur la formation des fibres dû à l'encombrement stérique. La zone formant des feuillets se situe juste après une zone flexible.

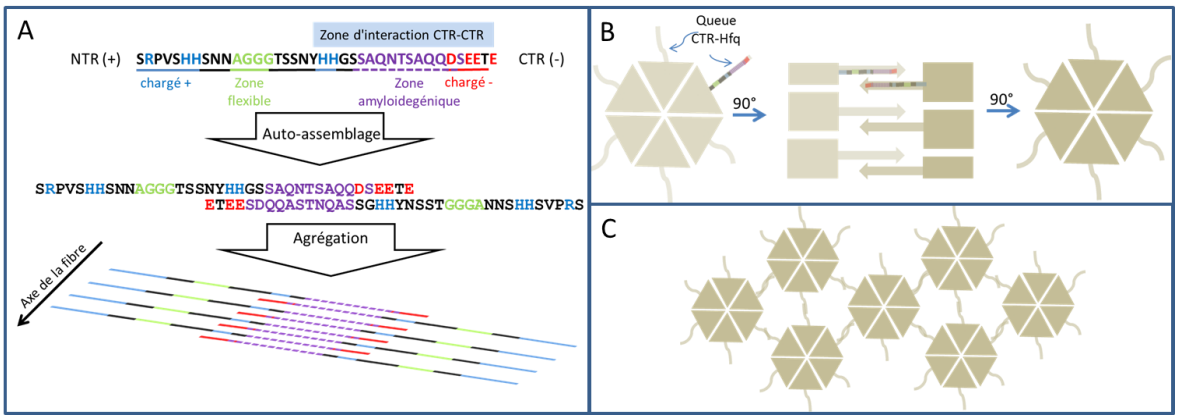


Figure 39 – Modèles et hypothèses de l'auto-assemblage Hfq par le domaine CTR. (A) Processus d'agrégation en fibres amyloïdes du domaine CTR-Hfq (d'après [192]). (B) et (C) sont deux modèles d'interaction d'Hfq via le contact du domaine CTR-Hfq. (B) Engrenage des queues CTR des hexamères entre deux protéines d'Hfq (d'après [193]). (C) Réseau protéique d'Hfq structuré par l'interaction de la queue CTR.

Les expériences menées dans le cadre de cette étude semblent indiquer que Hfq peut facilement se dissocier par dilution, ce qui suggère que l'interaction entre les queues CTR est relativement faible, au moins avant que les fibrilles isolées ne s'assemblent sous forme d'un faisceau de fibres. L'ajout de marquage peut fortement changer le comportement de la formation amyloïde de Hfq, ce qui remet en cause la possibilité d'observation par de nombreuses techniques de microscopie haute résolution. Cependant, cette faible interaction permet peut-être de jouer un rôle plus complexe en termes de régulation, par exemple quand Hfq *via* sa région CTR s'agrège sur les membranes, avec une préférence pour les membranes hydrophiles.

Chapitre II : mise en place d'une technique corrélative pour l'étude de nano-objets : application sur les fibres amyloïdes

Ce chapitre résume les résultats présentés dans l'article suivant : « Correlative infrared nanospectroscopy and transmission electron microscopy to investigate nanometric amyloid fibrils: Prospects and challenges » par **David Partouche***, Jérémie Mathurin*, Antoine Malabirade, Sergio Marco, Christophe Sandt, Véronique Arluison, Ariane Deniset-Besseau & Sylvain Trépout (* eq contribution), soumis à la revue *Journal of Microscopy* en aout 2018. Joint en **annexe B**.

1. Présentation de l'article

Cet article décrit l'observation de la même fibre à l'aide de deux techniques d'imagerie nanométrique : la microscopie électronique à transmission (MET) et l'AFM-IR. Cette approche de microscopie corrélative a été possible en respectant les contraintes de chaque technique. La faible densité de fibres déposées sur le substrat est le premier défi à relever. Le deuxième est d'avoir une statistique importante sur des fibres uniques et isolées. Enfin et surtout le choix du support est déterminant, car il doit être à la fois transparent aux électrons, aux rayonnements infrarouges, et avoir une faible rugosité de surface afin de faire des topographies AFM d'objets ayant une hauteur inférieure à 5nm.

Cette analyse tri-corrélative est possible avec un support à base de nitrure de silicium Si_3N_4 qui est de faible rugosité, transparent aux électrons et aux rayonnements infrarouges. Sa conception de forme (figure 40) permet de retrouver facilement la même zone afin de faire l'analyse sur le même échantillon sans ajouter de marqueur (figure 41). L'information apportée par chaque technique permet de mieux caractériser la fibre étudiée. La microscopie électronique permet de mesurer précisément la largeur des fibres (mais pas leur hauteur ou leur composition), la microscopie AFM intégrée dans l'AFM-IR permet de mesurer leur épaisseur précisément (mais est moins précise que la MET sur la largeur) et la spectroscopie nano-IR permet d'avoir une information sur la composition (protéique) et la structure secondaire locale (feuillets beta amyloïdes) des fibrilles.

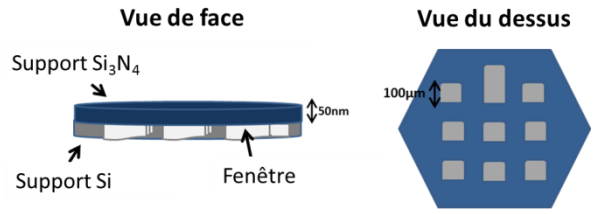


Figure 40 – Schéma du support de nitrure de silicium.

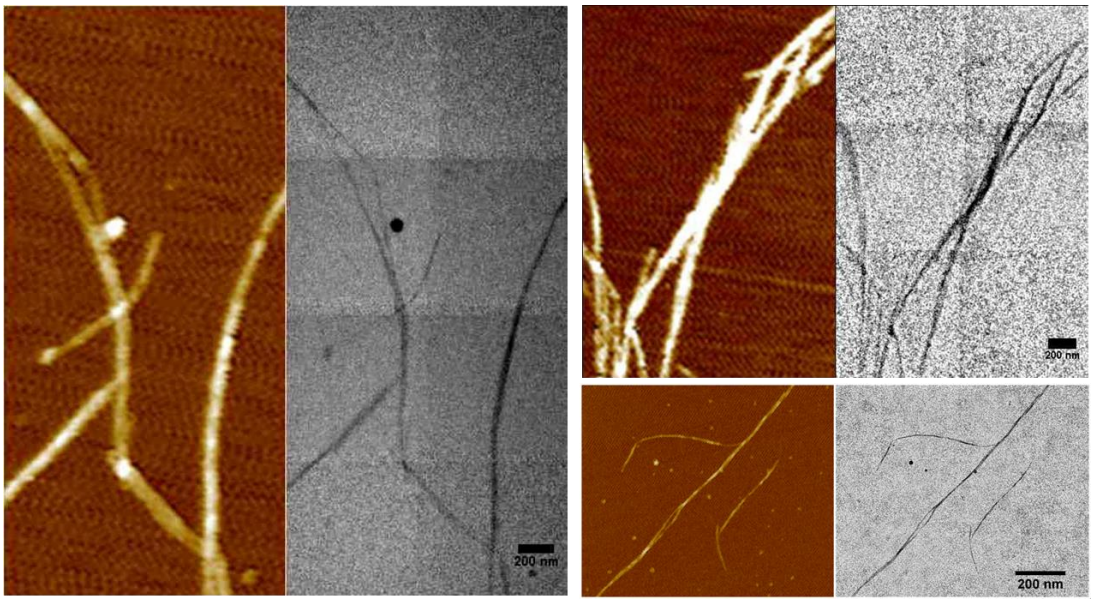


Figure 41 – Quelques exemples d'images des fibres CTR-Hfq obtenues par la méthode corrélative MET/AFM sur les fenêtres Si_3N_4. Les images MET sont reconstruite à l'aide du plugins MosaicJ sous images J [194].

2. Discussion

La corrélation entre MET et AFM-IR a été réalisée avec l'observation des fibres amyloïdes du CTR de Hfq. Grâce à l'AFM et à la MET, nous avons pu déterminer que les fibrilles individuelles avaient un diamètre de 4 nm et nous avons pu mesurer le spectre IR locale de quelques une de ces fibrilles avec une résolution latérale de quelques dizaines de nm. Ces objets sont à notre connaissance les plus petits jamais mesurés en nano-IR. L'acquisition de spectre IR localisé obtenu par AFM-IR pour un objet ayant de telles dimensions est une avancée et montre la grande sensibilité de la technique AFM-IR. Plusieurs difficultés techniques ont été identifiées et ont dû être résolues pour parvenir à réaliser ces mesures. Des améliorations sont encore possibles pour maximiser le signal AFM-IR :

- La qualité du signal est extrêmement dépendante de la qualité de la pointe AFM et de son revêtement en or. Il est pour l'instant nécessaire de sélectionner les meilleures

pointes pour réussir ces mesures. Un meilleur contrôle de la forme et de la taille de la sonde, permettant de travailler à de faibles puissances laser permettrait d'améliorer le rapport signal sur bruit et d'améliorer la consistance des mesures.

- Nous avons identifié que le film de Si_3N_4 pouvait se mettre à vibrer sous les impulsions du laser QCL utilisé comme source pour les mesures nano-IR et perturber ces mesures. Des adaptations doivent être envisagées : utiliser des films Si_3N_4 plus épais, travailler à des puissances laser faibles, éviter les effets de levier en travaillant près des bords du film.

Nous avons rencontré de grandes difficultés pour opérer le AFM-IR en mode imagerie et si le mode spectroscopie s'est révélé faisable sur le film, nous n'avons pas réussi à réaliser des cartes chimiques directement sur le film Si_3N_4. Cette imagerie n'a été possible qu'en dehors des fenêtres (figure 42) sur le support en silicium autour des films en Si_3N_4. Ces améliorations seront donc les clés pour généraliser l'imagerie tri-corrélative : topographique, électronique et chimique. Une imagerie est peut-être déjà faisable sur des objets de plus grande taille.

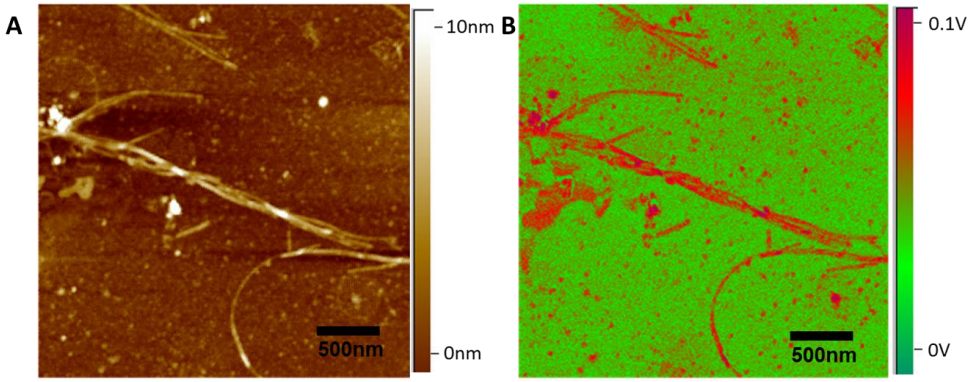


Figure 42 – Imagerie corrélative d'une solution séchée de peptide CTR-Hfq obtenue par AFM-IR sur support de Si_3N_4 (en dehors des fenêtres). (A) Topographie d'une zone 3µm x 3µm en AFM (B) Imagerie chimique infrarouge obtenue à une longueur d'onde de 1620 cm^{-1}.

Cependant, la méthode corrélative MET/AFM-IR pourrait trouver une application dans de nombreux domaines, non seulement dans le domaine des amyloïdes, mais en général pour les nanostructures auto-assemblées ou synthétiques. Une question importante, par exemple, porte sur l'analyse de fibres amyloïdes complexées à des acides nucléiques [195], ce que cette technique permettra d'analyser.

Chapitre III : rôle de la formation amyloïde d'Hfq sur la topologie de l'ADN

Comme indiqué précédemment, la cinétique d'association lente du peptide pose la question de la pertinence de sa formation *in vivo* (plusieurs jours contre des temps de division de 30 minutes en moyenne pour une bactérie). Nous avons donc cherché à savoir en quoi certains cofacteurs pourraient accélérer cet assemblage. Les résultats sont présentés dans l'article suivant : « Revised role for Hfq bacterial regulator on DNA topology » par Antoine Malabirade*, **David Partouche***, Omar El Hamoui, Florian Turbant, Frédéric Geinguenaud, Pierre Recouvreux, Thomas Bizien, Florent Busi, Frank Wien et Véronique Arluison. (* eq contribution) soumis à *Scientific reports* fin juillet 2018. Joint en **annexe C**.

1. Présentation de l'article

Cet article décrit en première partie l'étude de l'interaction entre le domaine CTR de Hfq et l'ADN. Nous établissons que l'ADN induit et favorise l'amyloïdogénèse du CTR, ce qui entraîne un changement local de la conformation de l'ADN. Nous avons ensuite étudié comment l'interaction de Hfq avec l'ADN affecte la topologie de l'ADN, en prenant un plasmide comme modèle.

Pour répondre à ces questions, nous avons fait la série d'expériences suivante :

Une analyse en dichroïsme circulaire par rayonnement synchrotron (SRCD) a permis de suivre l'agrégation en feuillets β dans une structure amyloïde (changement spectral significatif passant d'une réponse d'absence de structure à une réponse de présence de structure de type cross-β). Nous avons confirmé ce résultat en utilisant la spectroscopie infrarouge à transformée de Fourier (IRTF) (en D_2O). Nous observons clairement une bande/réponse correspondant à la formation de la structure amyloïde en présence d'ADN. Enfin, la formation de la structure de feuillets β croisés a également été étudiée par diffusion de rayons X aux petits angles (SAXS).

Pour étudier l'influence de Hfq et de sa queue terminale sur la topologie de l'ADN, nous avons étudié l'effet de Hfq purifiée sur la topologie de plasmides (pHSG298) *in vitro*. On observe la même migration du plasmide, en électrophorèse sur gel d'agarose, que ce soit en présence ou en absence d'Hfq. Cela signifie que Hfq n'influence pas directement les super-enroulements du plasmide. Pour confirmer ce résultat, nous avons utilisé un système de pinces magnétiques permettant de mesurer la force et la torsion appliquées sur un double-

brin d'ADN unique. L'analyse montre que Hfq se lie à l'ADN double brin en stabilisant les plectonèmes. Enfin, pour mieux comprendre le mode de liaison Hfq avec l'ADN, la compaction de l'ADN a également été étudiée par des mesures de sa longueur en fonction du temps. Une transition dynamique et réversible entre les états compacts et étendus de l'ADN a été observée et peut être interprétée comme la rupture / formation d'interactions entre les protéines Hfq et la molécule d'ADN double brin. La liaison entre protéines d'Hfq est susceptible d'être responsable de la stabilisation de cet état de compaction.

Des expériences sur l'étude de la compaction du plasmide pHSG298 ont également été réalisées *in vivo*, en extrayant de différentes souches les plasmides (sauvage WT, délétée *Δhfq* et délétée juste de la partie CTR *hfq72*). Nous n'avons vu aucune différence entre les souches WT et *hfq72* sur la topologie des plasmides. Pour la souche Δ*hfq*, le plasmide est plus relaxé, ce qui avait déjà été observé précédemment [23]. Ceci semble indiquer que Hfq induit un changement de topologie de l'ADN mais plutôt de manière indirecte. Une autre différence a en outre été observée : l'estimation du nombre de copies du plasmide pHS298 nous informe que la compaction par Hfq influence de manière défavorable l'efficacité de la réplication plasmidique. En effet, dans la souche où Hfq est présente (de manière entière ou tronquée), aucune différence significative n'est observée alors qu'en absence de Hfq, il y a un plus grand nombre de plasmides. Si l'efficacité de la réplication dépend de la compaction, la suppression du CTR dans *hfq72* devrait modifier l'efficacité de la réplication. Ainsi, nos résultats suggèrent que l'effet sur l'efficacité de la réplication est directement lié au changement de super-enroulement de l'ADN, mais est indépendant de compaction de l'ADN induite par le CTR.

2. Discussion

Avec cette série d'expériences, nous avons mis en lumière l'effet de l'acide nucléique sur la cinétique d'assemblage amyloïde du domaine CTR de Hfq. À noter, le fait que la protéine Hfq soit abondante dans les cellules et qu'environ un cinquième se retrouve dans le nucléoïde, fait qu'elle peut recouvrir une région de celui-ci (probablement avec les régions riches en adénosines de l'ADN), mais ça ne semble pas suffisant pour changer la topologie globale de l'ADN.

Nous établissons aussi à nouveau que la région amyloïde permet à Hfq de compacter l'ADN. Elle permet le pontage de séquences éloignées dans l'ADN double brin. Cette propriété a

déjà été observée pour des protéines associées au nucléoïde comme H-NS. Mais H-NS utilisant un motif de type « *coiled-coil* » alors que Hfq utilise un motif amyloïde. Ainsi, Hfq agrandit la famille des protéines associées au nucléoïde pontant l'ADN [62].

Enfin, le surenroulement (topologie) de l'ADN observé *in vivo* n'est pas affecté par le CTR, bien que ce domaine compacte fortement l'ADN. Cette étude amène donc à de nouvelles perspectives sur la relation qui existe entre pontage, surenroulement et compaction de l'ADN avec d'importantes conséquences sur la régulation génétique. Bien qu'il reste des zones d'ombre sur le mécanisme précis, Hfq joue un rôle important dans la structure du chromosome soit directement, soit indirectement en régulant la synthèse d'autres protéines.

Dans la poursuite de cette recherche, nous avons montré que l'auto-assemblage des fibres amyloïdes Hfq est favorisé par la présence de brins d'ADN, et qu'un complexe peptide-ADN est formé. Reste cependant une question, à savoir si l'acide nucléique fait partie du complexe amyloide ou si il permet simplement d'accélérer sa formation, comme dans le cas de RepA [195]. Nous pensons utiliser la technique de nanoIR (voir chapitre II, partie Étude) pour répondre à cette question. La technique corrélative vue dans le chapitre précèdent permettra aussi de déterminer la conformation de l'ADN dans le complexe le cas échéant. L'information IR permettra de déterminer la structure fine de l'ADN au niveau de la fibre, en analysant par exemple la conformation des sucres (C2' endo/C3' endo).

Enfin, il sera aussi intéressant de voir si ces effets sont aussi observés avec de l'ARN. Un nouveau *proposal* sur la ligne DISCO de SOLEIL a été déposé en ce sens.

Chapitre IV : interaction entre le domaine C-terminal de la protéine Hfq et des inhibiteurs d'amyloïdes

Il existe un grand nombre d'inhibiteurs d'amyloïdes connus. Comme vu précédemment, ils peuvent soit empêcher la formation des fibres ou désagréger les fibres. La première étape de mon travail a été de trouver des inhibiteurs pour notre système. Le but est de trouver les molécules permettant de gêner ou de dissocier la formation des fibres obtenues pour Hfq, dans l'espoir que ces molécules inhibent l'action de Hfq et affectent la virulence de la bactérie.

Pour cette partie nous avons principalement travaillé avec la séquence amyloïde minimale du domaine C-terminal contenant 11 acides aminés (on appellera le peptide correspondant CTR11) mais nous avons aussi travaillé avec le CTR entier de Hfq (qu'on notera CTR38, comprenant les 38 derniers acides aminés de Hfq de *E. coli*). La première approche consiste à utiliser un ensemble d'inhibiteurs déjà développés dans le cadre de pathologies neurodégénératives pour affecter la fibrillation de Hfq, et vérifier qualitativement par imagerie MET l'effet de ces inhibiteurs sur la structure de la fibre. Après une recherche bibliographique, nous avons identifié les molécules suivantes : R-Apomorphine Hydrochloride (APOM) [124], (-)-Epigallocatechin Gallate (EGCG) [196], Acide Fuchsin [197], Acide Caféique [198], Curcumine [199], Rifamycine et Silibinine [200].

Nous avons effectivement observé des propriétés d'inhibition ou de désagrégation des fibres amyloïdes d'Hfq avec certaines de ces molécules. Nous avons également observé que certains de ces inhibiteurs ne fonctionnaient qu'avec certains types de fibres (CTR11 ou CTR38). Un autre problème rencontré est le polymorphisme dans les échantillons amyloïdes obtenus *in vitro* [201]. De plus, une simple dilution lors de l'ajout d'inhibiteurs, qui ont souvent une solubilité assez faible et doivent être dissous dans de grands volumes, peut provoquer une disparition des fibres ce qui a rendu la tâche plus difficile.

Dans la liste des inhibiteurs testés, deux ont retenu notre attention et nous semblaient les plus prometteurs : l'APOM et l'EGCG. Les autres semblaient aussi avoir un effet et ne sont pas à écarter.

1. Études de APOM et EGCG sur le peptide CTR38

Pour étudier l'effet des inhibiteurs sur les fibres obtenues avec le CTR38, nous avons utilisé l'imagerie MET, le SAXS et le SRDC. En MET, nous constatons clairement une perte en nombre et en forme de fibres en présence de 5mM de EGCG et APOM. L'effet de EGCG semble déjà se produire à 1mM mais être plus efficace à 2,5mM (figure 43).

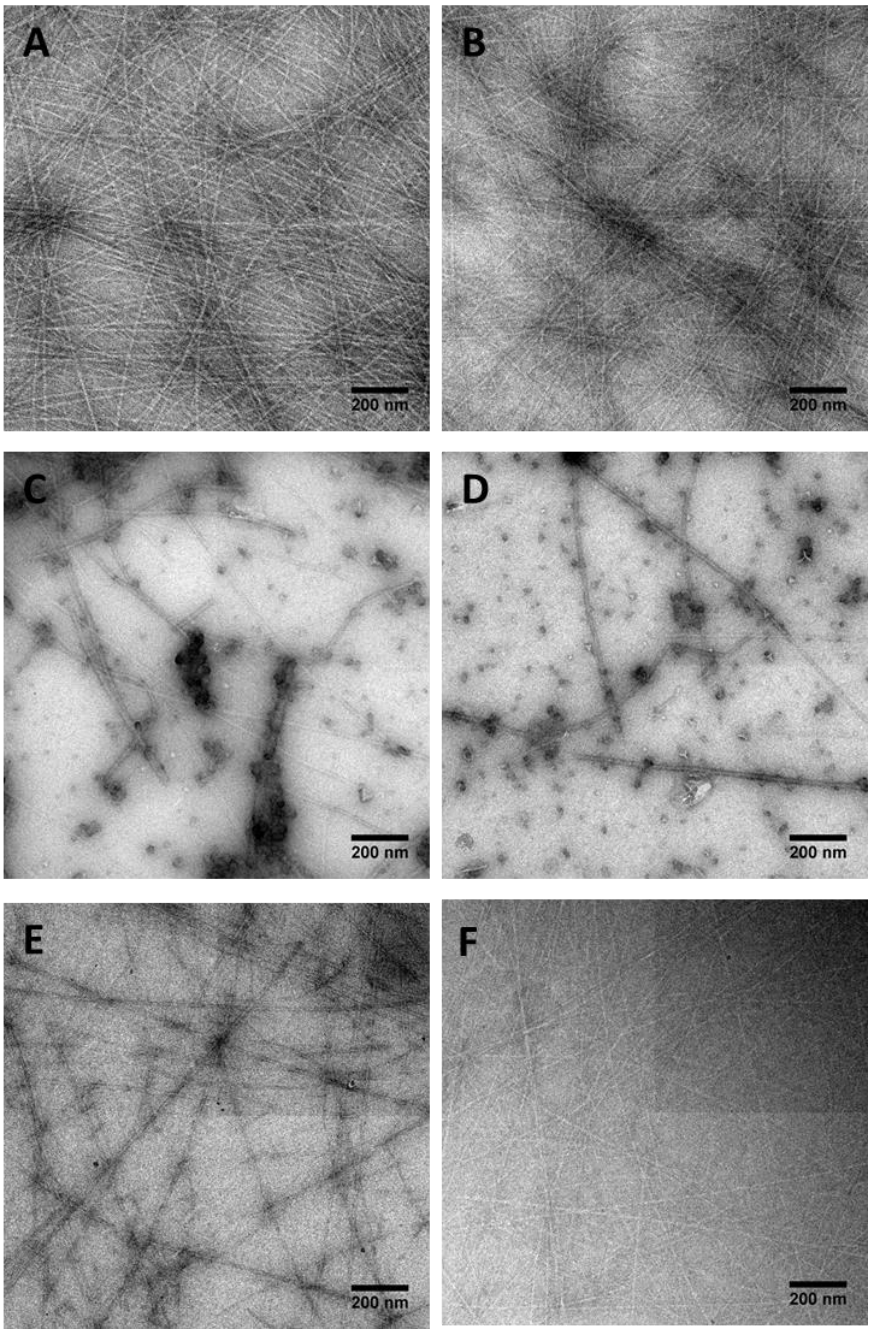


Figure 43 – Images MET de fibres amyloïdes en présence et absence d'inhibiteurs. Les images A et B correspondent aux fibres amyloïdes formées par le domaine CTR-Hfq 38aa sans inhibiteur. Les images C et D correspondent aux fibres

et agrégats amyloïdes formés par le domaine CTR38aa en présence d'inhibiteur d'EGCG à 5mM. Toutes les incubations ont duré 24h. Les images E et F correspondent aux fibres et agrégats amyloïdes formés par le domaine CTR38aa en présence d'inhibiteur d'APOM à 5mM, incubé durant 24h.

Nous avons également observé l'effet sur la structure par SAXS sur la ligne SWING du synchrotron SOLEIL (figure 44). On peut observer la diminution du pic amyloïde, la présence d'un pic plus faible suggère que des fibres subsistent comme on peut le voir aussi en imagerie MET.

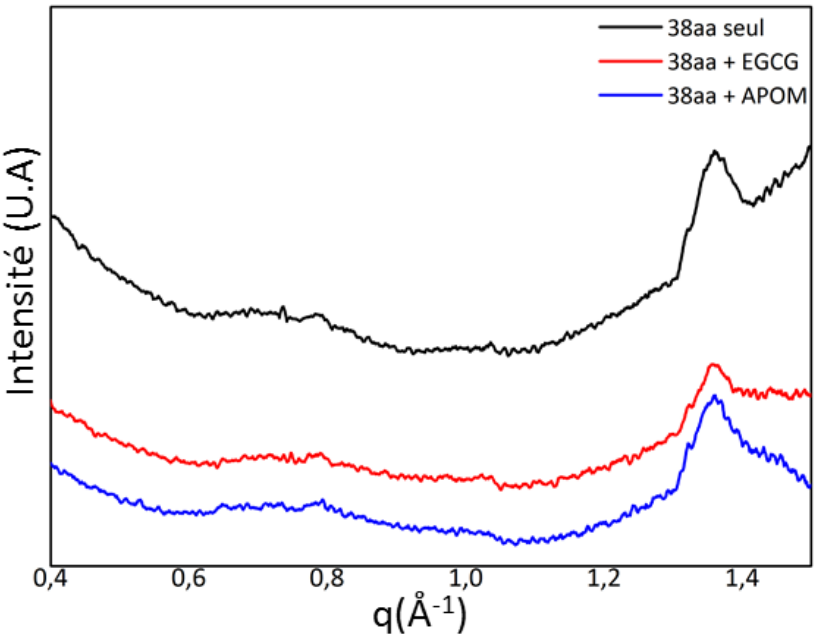


Figure 44 – Courbes SAXS de fibres amyloïdes en présence et absence d'inhibiteurs. Les courbes SAXS des fibres obtenues par le domaine CTR-Hfq 38aa seul (courbe noire) à 20mg/ml, en présence de 5mM de EGCG (courbe rouge) ou de 5mM de APOM (courbe bleue) incubé pendant 24h.

Enfin une cinétique en SRDC (figure 45) en présence de 5mM de chacun des inhibiteurs a été effectuée.

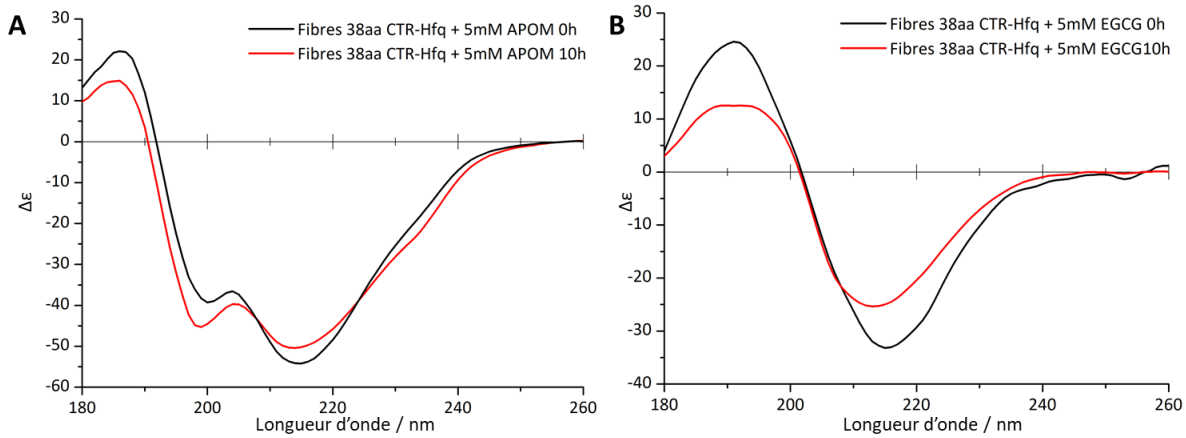


Figure 45 – Cinétiques montrant l'effet des inhibiteurs en SRCD sur la structure cross-β amyloides. (A) spectres en présence d'APOM et (B) d'EGCG. Seuls les spectres initiaux et à 10h sont présentés.

Ces séries d'expériences nous montrent clairement l'effet des inhibiteurs utilisés ici sur les fibres amyloïdes obtenues avec CTR38. EGCG semble plus efficace que APOM. En effet, en présence de 5mM de chacun de ces composés, on observe moins de fibres en présence de EGCG que d'APOM. Mais dans les deux cas il y a moins de fibres en présence de molécules inhibitrices qu'en leur absence. Ce résultat est confirmé par le pic de diffraction qui est moins haut pour APOM et beaucoup moins haut pour EGCG. Concernant le résultat SRDC, nous remarquons une diminution de structures de type feuillets β et un shift vers le pic à 195 nm correspondant à un peptide non structuré.

2. Études de l'effet EGCG sur le peptide 11aa

À la suite des résultats précédents, nous nous sommes plus particulièrement intéressés à l'effet de EGCG et nous avons pris comme modèle la partie minimale du domaine CTR composée d'une séquence de 11 acides aminés, le peptide CTR 11aa. Nous utilisons ce peptide et non le CTR 38aa car il présente une moins grande variabilité de morphologie des fibres et de cinétique d'association.

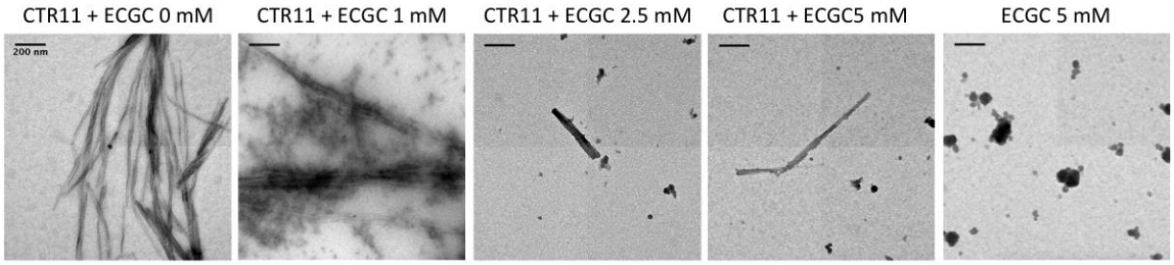


Figure 46 – Visualisation en MET de l'effet d'une gamme d'ECGC sur les fibres CTR11. La présence de 1 mM d'ECGC suffit à dissocier les fibres, même si l'effet n'est pas complet (durée incubation 24h).

L'analyse de l'effet d'une gamme d'ECGC sur les fibres CTR11 vu en MET nous montre que même 1 mM d'ECGC suffit à dissocier les fibres, même si l'effet n'est pas complet (figure 46). Augmenter la concentration augmente cet effet, même si quelques rares fibres subsistent sur la grille. Par la suite nous avons donc voulu confirmer cet effet en présence de 5mM de EGCG par une analyse en SRDC. Deux types d'analyses ont été effectués, une cinétique sur la nuit et une mesure de la température de fusion des fibres.

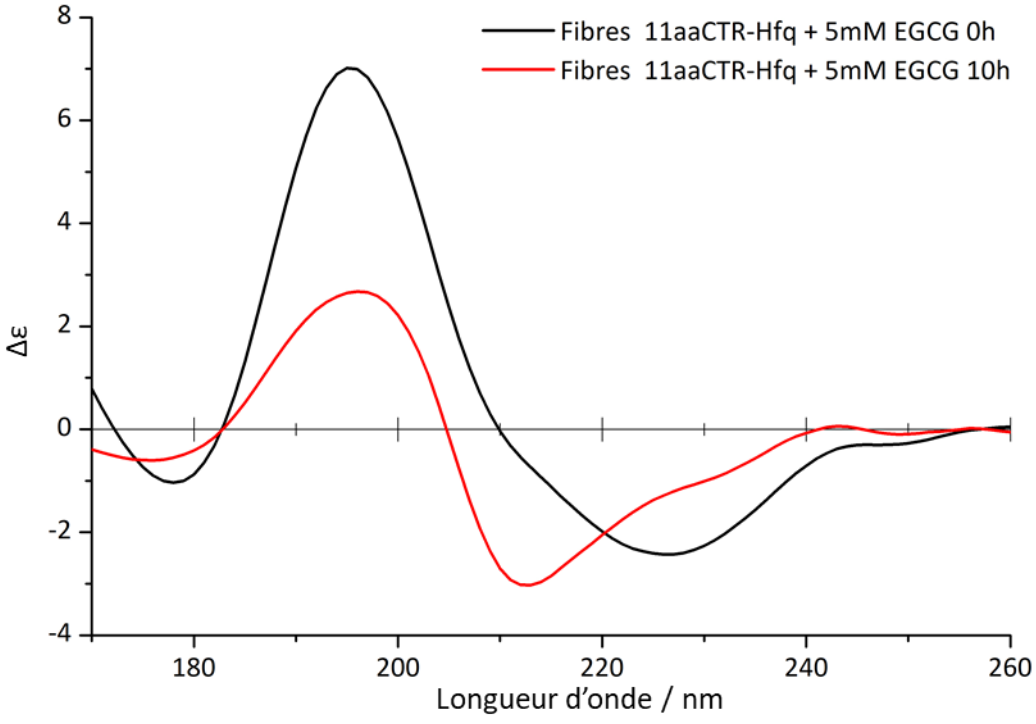


Figure 47 – Courbes cinétiques SRDC de fibres en présence de 5mM ide EGCG 5mM incubé durant 10h. Seuls les point initiaux et finaux sont montrés.

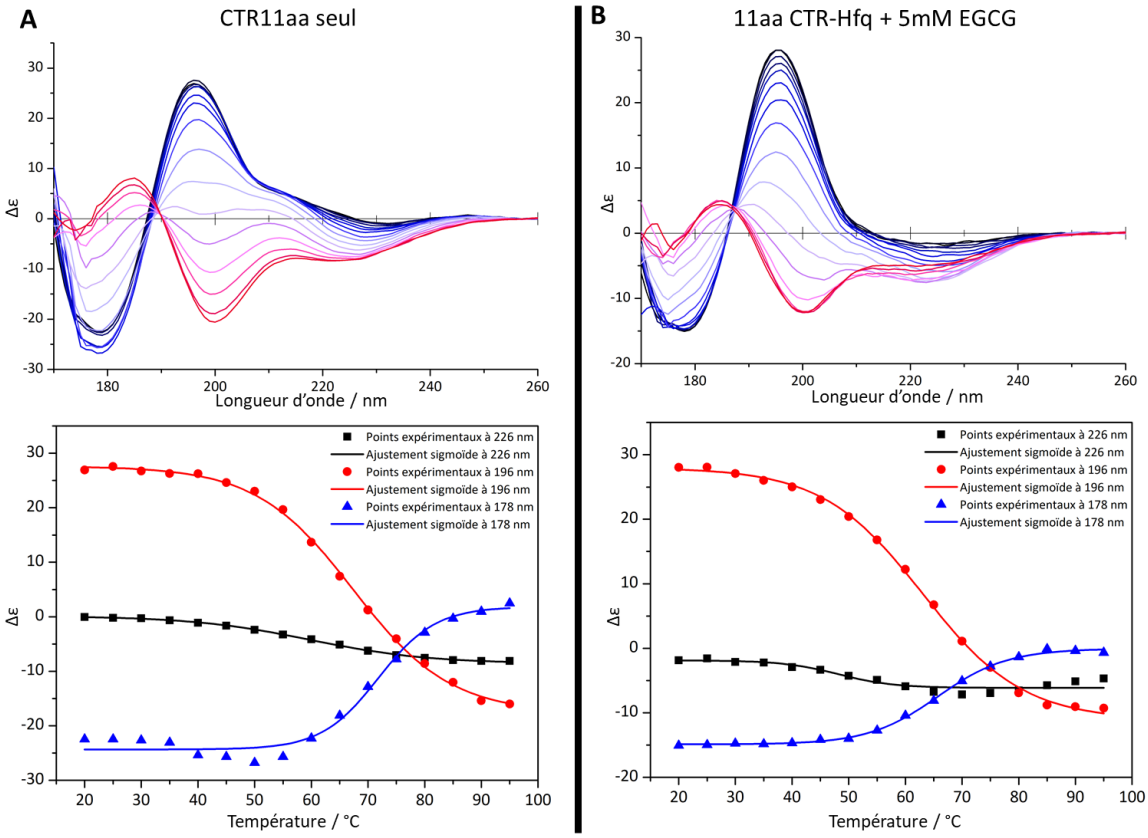


Figure 48- Détermination des températures de fusion par SRCD. Les analyses ont été faites sur des fibres 11aa CTR-Hfq en (A) l'absence et (B) présence de EGCG**.** Les points CD ont été mesurés tous les 5 °C pour les longueurs d'onde suivantes : 226 nm, 196 nm et 178 nm. Les résultats de la température de fusion pour la longueur d'onde de 226 nm, pour CTR seule ou en présence d'EGCG sont respectivement : Tm = 59,7 +/- 2,7 °C et Tm = 48,1 +/- 2,7 °C. Pour la longueur d'onde de 196 nm, les températures de fusion sont : Tm = 67,4 +/- 0,4 °C et Tm = 63,1 +/- 0,4 °C et pour la longueur d'onde de 178 nm, nous obtenons : Tm = 71,8 +/- 1,0 °C et Tm = 65,7 +/- 0,3 °C.

La cinétique par SRDC du CTR11 en présence de 5mM EGCG montre une dissociation de la structure amyloïde par un décalage du pic négatif de 223 à 212 nm ainsi qu'une diminution des pics. Concernant la température de fusion, nous obtenons une différence de température de fusion moyenne inférieure à 7 °C pour la solution contenant EGCG (66,3 °C) comparé à une solution n'en contenant pas (58,9 °C).

Dans les deux analyses, on remarque un effet de l'ECGC sur la stabilité des fibres. Nous avons souhaité confirmer l'effet de l'ECGC par une analyse en microcalorimétrie (DSC),

mais hélas l'ECGC donne un fort signal au même endroit que la fusion du peptide et il n'a pas été possible de mener à bien cette analyse.

3. Études de l'effet EGCG *in vivo*

Enfin, dans une dernière partie nous avons souhaité voir l'effet de l'EGCG sur les bactéries. Cette dernière analyse est cependant complexe car l'EGCG est déjà connu pour affecter la croissance et la virulence bactérienne [202]. Il fallait donc une approche permettant de savoir si l'EGCG jouait un rôle via Hfq d'une part et d'autre part si ce rôle était lié au CTR amyloïde. Les souches dépourvues de Hfq (Δ*hfq*) sont en effet généralement plus sensibles à l'ajout de composés, dû à une déficience à la réponse en stress via des sARN. Nous avons donc effectué des mesures de survie des bactéries en fonction de concentrations croissantes d'inhibiteur et déterminé la concentration qui donne 50% de CFU (colony-forming-unit). Pour ce faire, les cultures soumises aux inhibiteurs ont été étalées sur boites et les colonies ont été comptées. Le travail a été fait sur les souches WT, Δ*hfq* (totalement délétée du gène *hfq*) et *hfq72* (délétée juste de la partie CTR, voir chapitre I). Cette concentration, pour chaque inhibiteur, sera appelée IC50. Les résultats ont montré que, comme prévu, la souche Δ*hfq* était très sensible à l'ajout du composé. Afin de confirmer que cet effet était au moins en partie dû à l'effet sur le CTR, nous avons donc mesuré des CMI pour la souche *hfq72*. Les résultats ont été obtenus avec différentes dilutions des cultures et en triplicata. Nous avons pu estimer que les C50 sont de 0,67 ± 0,02 mM pour la souche WT contre 0,81 ± 0,05 mM pour souche délétée du CTR. Ceci montre, que même si l'effet est plus fort pour la souche totalement dépourvue de Hfq (probablement dû à la régulation de sigma E par RybB) [202], un effet dû à l'absence de CTR uniquement est aussi observé. Nous chercherons donc par la suite à analyser cet effet plus en détail. À noter, la IC50 mesurée n'est cependant pas très élevée, et il faudra probablement optimiser le composé en le modifiant chimiquement pour cibler le CTR de Hfq (« *drug design* »). Cette analyse est aussi en cours au laboratoire, en collaboration avec le laboratoire de P. Tuffery à l'université Paris Diderot.

Chapitre V : caractérisation structurale des fibres amyloïdes/ADN et analyse des propriétés rhéologiques du système.

1. Introduction

Dans ce dernier chapitre, nous présenterons les premiers éléments d'une étude encore en cours montrant l'influence de la formation du complexe amyloïde/ADN sur la dynamique du système. Pour ce faire, nous avons d'une part caractérisé la structure des fibres CTR38/ADN en Cryo-EM, puis utilisé une technique de rhéologie. La rhéologie est un domaine de la physique permettant de déterminer les propriétés mécaniques telles que la viscosité et l'élasticité d'un fluide complexe. Ces propriétés sont déterminées en étudiant et mesurant les déformations ainsi que la résistance ou l'écoulement d'un milieu sous l'application d'une force ou une contrainte.

En biologie, la rhéologie est principalement utilisée en biologie cellulaire. Elle permet par exemple d'étudier la motilité [203], la prolifération [204], et la différentiation cellulaires [205]. La technique permet, par exemple, de déterminer d'un point de vue mécanique une différence entre les cellules saines et cancéreuses, ce qui ouvre une voie d'application de diagnostic tumoral [206].

Il existe de nombreuses façons de faire de la rhéologie sur les cellules. Les mesures sont faites par exemple, par aspiration avec une micropipette [207], par cytométrie de torsion mécanique et par AFM [208][209]. L'AFM permet de faire de la nano indentation pour sonder et caractériser un matériau à l'échelle submicrométrique [210].

Ces techniques fournissent une information capitale sur la mécanique des fluides ou des cellules. Elles permettent principalement de remonter au module d'élasticité (ou module de Young).

La microrhéologie est une branche de la rhéologie utilisant des sondes à l'échelle micrométrique dont le volume étudié peut se réduire jusqu'au picolitre. Elle a été développée dans le but de se rapprocher le plus des conditions cellulaires lors des mesures.

Nous allons donc présenter dans ce chapitre les résultats obtenus en Cryo-EM montrant la structure du complexe CTR38:ADN. Nous présenterons ensuite dans ce chapitre les détails de la technique de microrhéologie utilisée et les résultats obtenus.

2. Analyse du complexe CTR38/ADN

Pour cette étude, les mêmes échantillons que ceux du chapitre III ont été utilisés, à savoir le CTR38 (le CTR11 n'interagit pas avec l'ADN) et un duplex de type $(dA:dT)_{60}$.

Les résultats obtenus en Cryo-EM nous montrent que les fibres amyloïdes du CTR38 seules et celles du CTR38 complexées à l'ADN sont bien différentes. Ainsi, comme montré sur la figure 50, on peut voir que le complexe forme une structure avec une apparence clairement hélicoïdale, contrairement au peptide CTR38 auto-assemblé seul, en absence d'ADN. Il faut noter que cette structure hélicoïdale est complexe et provient probablement du super-enroulement de deux fibres (voir modèle sur figure 49). Par ailleurs, les fibres ont tendance à être parallèles entre elles, ce qui peut suggérer un système complexe et rigide. C'est ce que nous allons tenter de confirmer dans ce chapitre par microrhéologie.

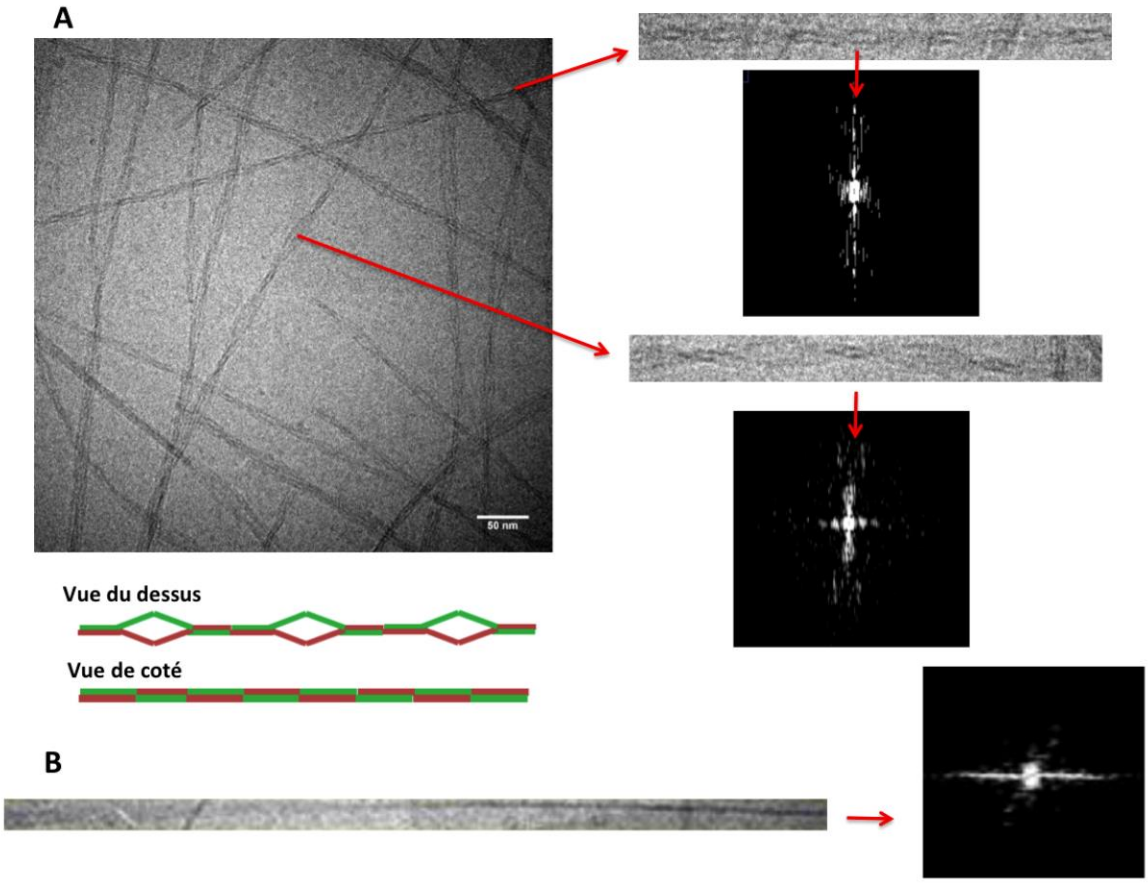


Figure 49 – Analyse en Cryo-EM du complexe CTR38/ADN et comparaison avec les fibres de CTR38 seul. On remarque que le complexe CTR38/ADN forme une structure hélicoïdale (A) contrairement au peptide auto-assemblé seul (B). Ces résultats sont confirmés par l'analyse par transformée de Fourier (cadres noirs).

Il faut noter qu'il est difficile de dire si l'ADN est présent dans ces fibres sur cette seule analyse, même si on le soupçonne. Pour cette raison, l'analyse en AFM-IR serait clairement un atout, comme expliqué dans les chapitres II et III. Cette analyse est prévue dès que le « nouveau » AFM-IR, permettant l'analyse des sucres, sera installé au laboratoire dans quelques semaines.

3. Microrhéologie : la spectroscopie magnétique rotationnelle

Par la suite, nous avons voulu mieux comprendre l'impact de la formation de ce complexe sur les propriétés du système. Nous avons en particulier souhaité étudier les propriétés de viscosité de la solution, en utilisant une technique de rhéologie. Comme dans toute technique de rhéologie, on mesure la résistance d'un liquide, en réponse à une force appliquée.

La technique originale de microrhéologie présentée dans ce chapitre a été développée par J.F Berret au laboratoire Matière et Système Complexe de l'université Paris-Diderot [211] [212][213]. Nous avons utilisé des micro-aiguilles magnétiques composées de maghémite, un oxyde de fer (Fe_2O_3). Elles ont une taille moyenne de 1-2 microns de diamètre et une dizaine voire une centaine de microns de longueur. La susceptibilité magnétique $\Delta\chi$ des aiguilles utilisées a été déterminée en utilisant des mélanges eau-glycérol de viscosité connus. Nous avons trouvé $\Delta\chi = 0.054$. Les micro-aiguilles sont injectées dans l'échantillon afin d'étudier son mouvement de rotation lors de l'application d'un champ tournant et contrôlé.

Le comportement des micro-aiguilles magnétiques dépend de plusieurs paramètres :

- La susceptibilité magnétique, notée $\Delta\chi$.
- La fréquence du champ appliqué, notée ω.
- Les dimensions des aiguilles : longueur, notée L et diamètre, noté D.
- Les propriétés rhéologiques du milieu : le coefficient de viscosité, noté η, et le module d'élasticité, noté G.

Les paramètres inconnus sont bien entendu les propriétés rhéologiques du milieu. Leur détermination se fait par l'analyse du mouvement de l'aiguille en prenant en compte l'ensemble des paramètres du système.

Projetée sur un repère à deux dimensions, la position de l'aiguille à l'angle θ en fonction du temps est étudiée (figure 50). Le mouvement diffère selon la nature du milieu (fluide newtonien, fluide de maxwell, gel Kelvin-voigts, gel de Hooke) où se trouve la micro-aiguille. En absence ou à très faible résistance du milieu, la micro-aiguille tourne à une fréquence angulaire imposée par la fréquence du champ tournant.

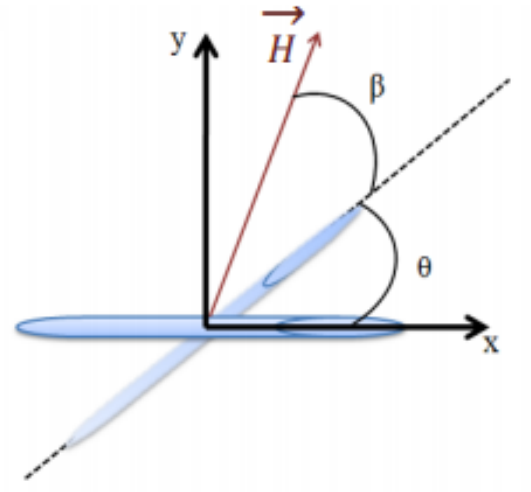


Figure 50 – Représentation d'une micro-aiguille en rotation dans un repère.
θ représente l'angle formé par l'aiguille à deux instants t. L'angle β est formé entre l'aiguille et le champ tournant $\vec{\boldsymbol{H}}$. Adapté de [211].

Le but de la technique consiste à déterminer la fréquence angulaire moyenne, notée Ω, et l'angle « de retour » ou moyenne des amplitudes des oscillations, noté θ_B, pour différentes aiguilles.

Dans le cas d'un fluide viscoélastique ou modèle de « fluide de Maxwell », l'angle θ en fonction du temps s'exprime par :

$$\theta = \omega t - \beta$$

Il existe une fréquence critique, notée ω_c, et exprimée par :

$$\omega_c = \frac{3}{8}\frac{\mu_0 \Delta\chi}{\eta} g\left(\frac{L}{D}\right)\frac{D^2}{L^2}H_0^2$$

Avec H_0 l'excitation magnétique, μ_0 la perméabilité magnétique du vide et $g\left(\frac{L}{D}\right)$ une fonction telle que :

$$g\left(\frac{L}{D}\right) = \ln\left(\frac{L}{D}\right) - 0{,}662 + \frac{0{,}917D}{L} - \frac{0{,}05D^2}{L^2}$$

Cette fonction prend en compte les proportions de l'aiguille approximée à un cylindre.

Pour chaque aiguille dans chaque milieu, on peut déterminer la fréquence ω_c qui est délimitée par deux régimes de rotation de l'aiguille (figure 51):

- **Le régime synchrone :** L'aiguille est synchrone avec le champ et l'angle θ croit linéairement avec le temps. Dans ce régime $\omega < \omega_c$ et $\Omega(\omega) = \omega$.
- **Le régime asynchrone :** A de fortes fréquences de champ magnétique, l'aiguille en rotation se désynchronise et subit des inversions de rotation de l'angle (θ_0) en réponse aux forces mécaniques du fluide. Dans ce régime $\omega > \omega_c$ et $\Omega(\omega) = \omega - \sqrt{\omega^2 - \omega_c^2}$.

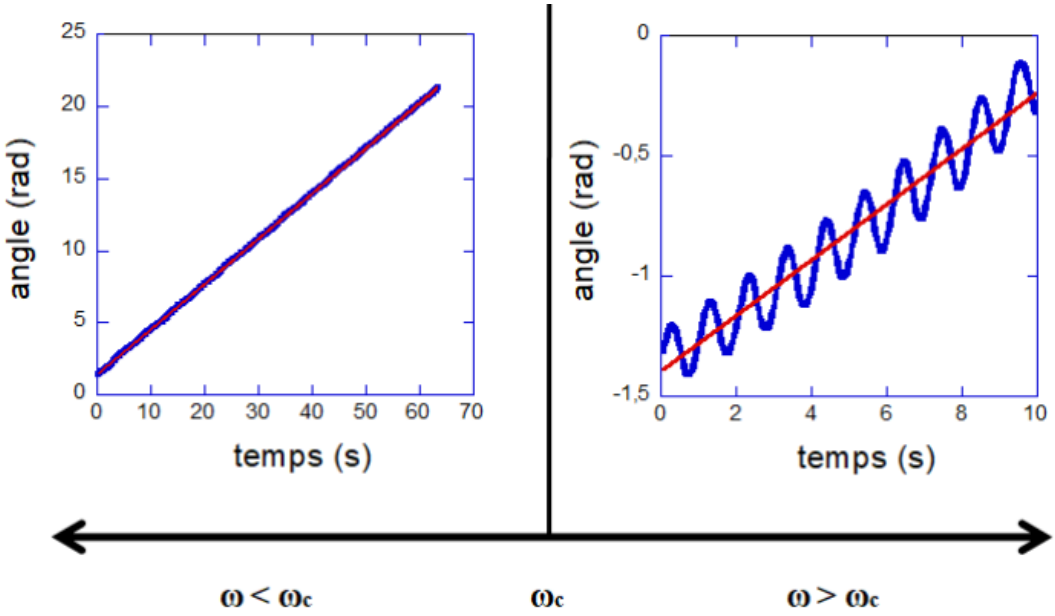


Figure 51 – Types de réponse possibles à la rotation d'une aiguille dans un fluide viscoélastique (modèle de Maxwell). A gauche est représenté le cas d'un régime synchrone où l'aiguille tourne à une fréquence angulaire égale à celle du champ. A droite est représenté le cas d'un régime asynchrone où l'aiguille subit une force mécanique du milieu saccadant sa rotation d'un angle θ_B de manière régulière. En rouge est représenté le fit linéaire moyen, de pente Ω. Données tirées des expériences menées sur CTR et de l'ADN, un duplex $(dAdT)_{60}$ (identique à celui utilisé en SRCD).

La mesure de Ω(ω) permet de remonter à la fréquence critique ω_c nécessaire à la détermination de la viscosité du milieu.

En régime asynchrone, on peut déterminer le module d'élasticité du milieu, notée *G*, par la mesure de l'amplitude des oscillations θ_B en fonction de ω, donnée par la relation :

$$\lim_{\omega \to \infty} \theta_B = \frac{3}{4} \frac{\mu_0 \Delta\chi}{G} g\left(\frac{L}{D}\right) \frac{D^2}{L^2} H_0^2$$

Dans le cas d'un gel viscoélastique ou modèle de « fluide de « Kelvin-Voigt», les micro-aiguilles oscillent de manière plus prononcée, avec Ω(ω) proche de zéro. Dans ce modèle, l'amplitude θ_B est définie par :

$$\theta_B(\omega) = \frac{3}{4} \frac{\mu_0 \Delta\chi}{G'(\omega)} g\left(\frac{L}{D}\right) \frac{D^2}{L^2} H_0^2$$

Avec *G'(ω)* correspondant au module de cisaillement.

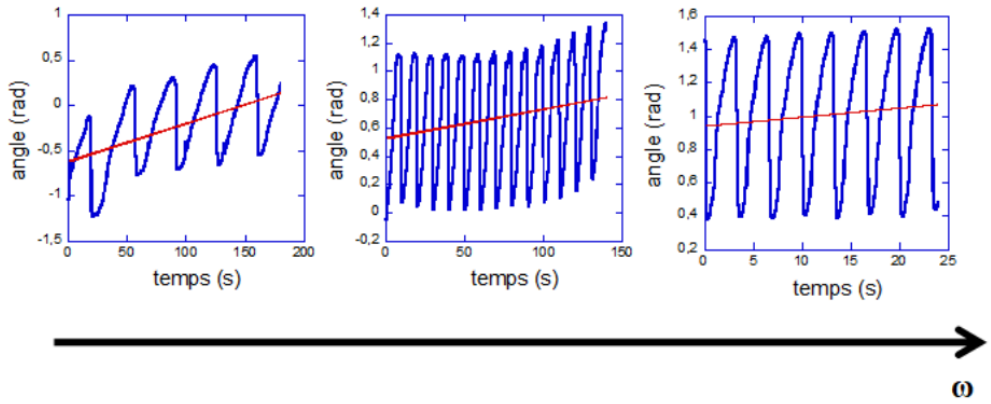


Figure 52 – Type de réponse au mouvement d'une aiguille dans un gel viscoélastique (modèle de Kelvin Voigt). Les amplitudes d'oscillation sont plus grandes et la valeur de Ω(ω) est proche de zéro. Le milieu a une plus grande contribution de l'élasticité dans ses propriétés mécaniques. Données tirées des expériences menées sur CTR+$(dAdT)_{60}$.

Cette technique de micro rhéologie, par procédé de spectroscopie magnétique rotationnelle, a été utilisée pour étudier et déterminer les propriétés mécaniques sous-jacentes de l'interaction entre le CTR-Hfq et l'ADN.

4. Montage de l'expérience

L'acquisition pour l'analyse du comportement des aiguilles est réalisée par microscopie à champ clair, sur un microscope IX73 (Olympus) avec un objectif ×20. Les échantillons dans

lesquels les micro-aiguilles ont été injectées à une quantité optimale sont déposés sur lamelles de verre scellées par un système adhésif GeneFrame (Abgene/Advanced Biotech). Elles sont déposées sur un dispositif de deux paires de bobines conduisant un courant sinusoïdal déphasé l'un de l'autre de $\pi/2$ et induisant de facto un champ magnétique tournant. Ce dispositif est relié à un générateur permettant d'imposer une gamme de fréquence variablede 5 à 5000 mHz et un champ de 0 à 20 teslas. Le système d'acquisition d'images est composé d'une caméra EXi Blue CCD (QImagine) couplée au logiciel MetaVue (Universal Imaging Inc.). Les images sont ensuite traitées grâce à un plugin dédié sur ImageJ. Les acquisitions, par ce montage, ont été faite au laboratoire Matière et Système Complexe.

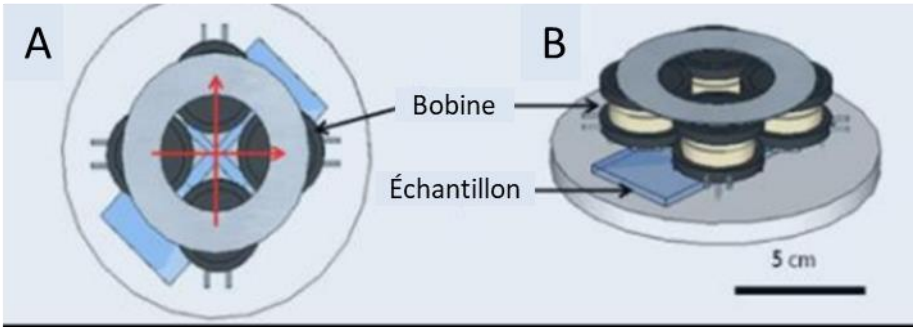


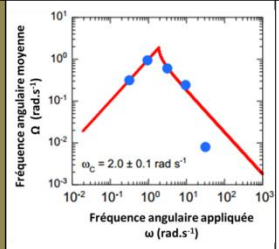
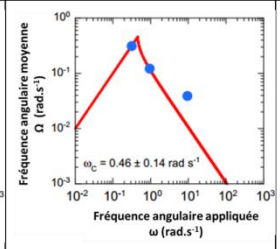
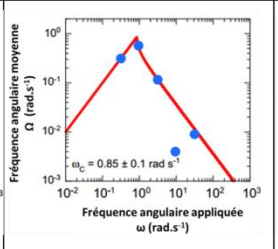
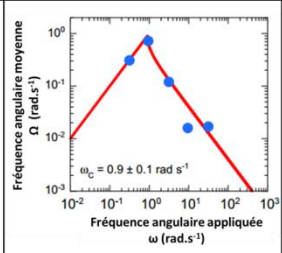
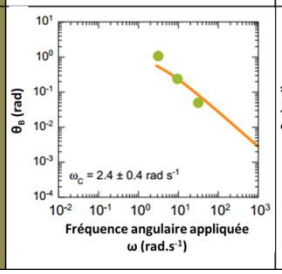
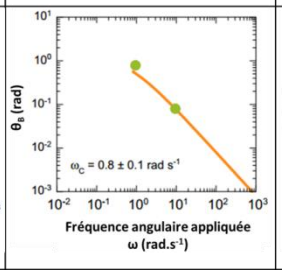
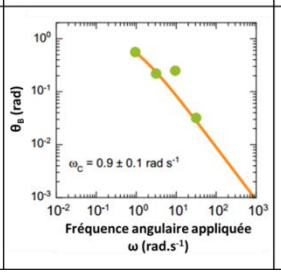
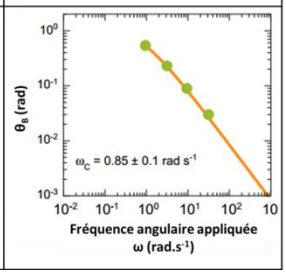
Figure 53 – Schéma du dispositif de bobines utilisé pour la microrhéologie. (A) Vue du dessus. (B) Vue de ¾ en plongée. L'échantillon sur la lamelle est placé au centre des bobines afin d'avoir une répartition isotropique du champ magnétique sur l'échantillon. Mofidiée de [211].

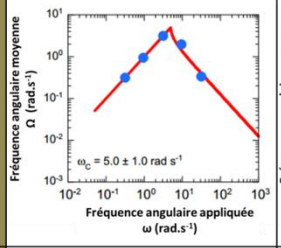
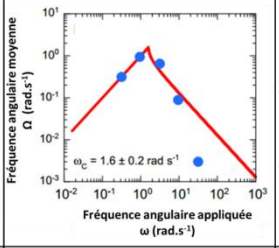
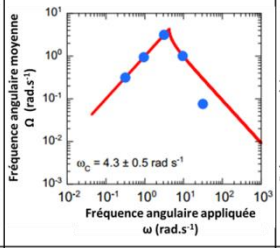
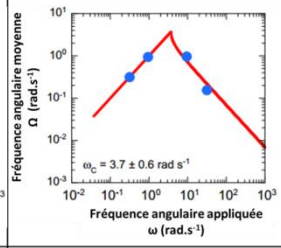
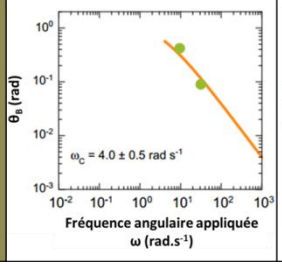
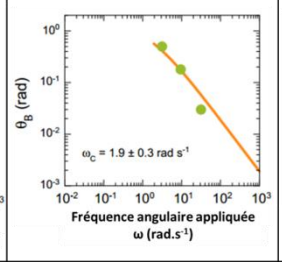
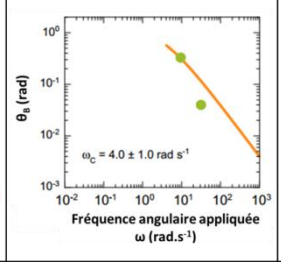
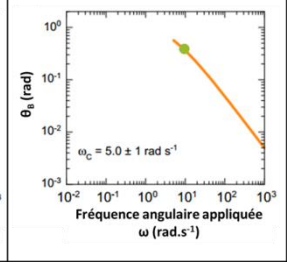
5. Résultats et discussion

Pour déterminer et comparer les propriétés rhéologiques d'une solution de CTR mélangée à de l'ADN, l'expérience de microrhéologie a été menée sur trois échantillons : le CTR38 de Hfq à 7 $mg.ml^{-1}$, le double brins d'ADN $(dAdT)_{60}$ à 7,3 mM et le CTR38 + $(dAdT)_{60}$ (respectivement à 7 $mg.ml^{-1}$ et 7,3 mM), solutions préparées à 3 semaines avant les expériences pour permettre un bon assemblage du complexe (voir chapitre IV). Pour chaque échantillon, les micro-aiguilles magnétiques ont été injectées et des films de plusieurs aiguilles de tailles différentes ont été enregistrés à des champs tournants, de fréquences allant de 15 à 5000 mHz. Le comportement rotationnel des aiguilles a été analysé en traçant les θ en fonction du temps.

5.1. Double brin d'ADN (dAdT)60 seul

Un ensemble de huit micro-aiguilles a été étudié durant l'expérience. Celles-ci ayant chacunes des formes différentes. Les ω_c ont été déterminées et présentées dans le tableau suivant :

Aiguille	1	2	3	4
L (µm)	51.3	72.37	112.83	83.74
D (µm)	1.8	2.71	2.4	1.53
Ω(ω)				
$\theta_B(\omega)$				

Aiguille	5	6	7	8
L (µm)	27,46	45.51	52.48	33.03
D (µm)	1.84	2.27	2.14	1.85
Ω(ω)				
$\theta_B(\omega)$				

En analysant, la fréquence angulaire moyenne en fonction de la fréquence angulaire appliquée, ces dernières étant toutes deux divisées par les fréquences critiques mesurées, on obtient une courbe de points en accord avec un ajustement de prédiction. Cette prédiction est calculée pour un fluide purement visqueux dont les équations sont les suivantes :

- Pour $\omega/\omega_c \leq 1$, on a : $\Omega(\omega) = \frac{\omega}{\omega_c}$

- Pour $\omega/\omega_c \geq 1$, on a : : $\Omega(\omega) = \frac{\omega}{\omega_c} - \sqrt{(\omega/\omega_C{}^2 - 1}$

De même, en analysant l'amplitude d'oscillations en fonction de la fréquence réduite (ω/ω_c), il existe un bon accord entre le modèle de Newton et les données de l'échantillon.

Les ajustements, tracés rouges, et les données obtenues, points bleus sont présentés figure 54.

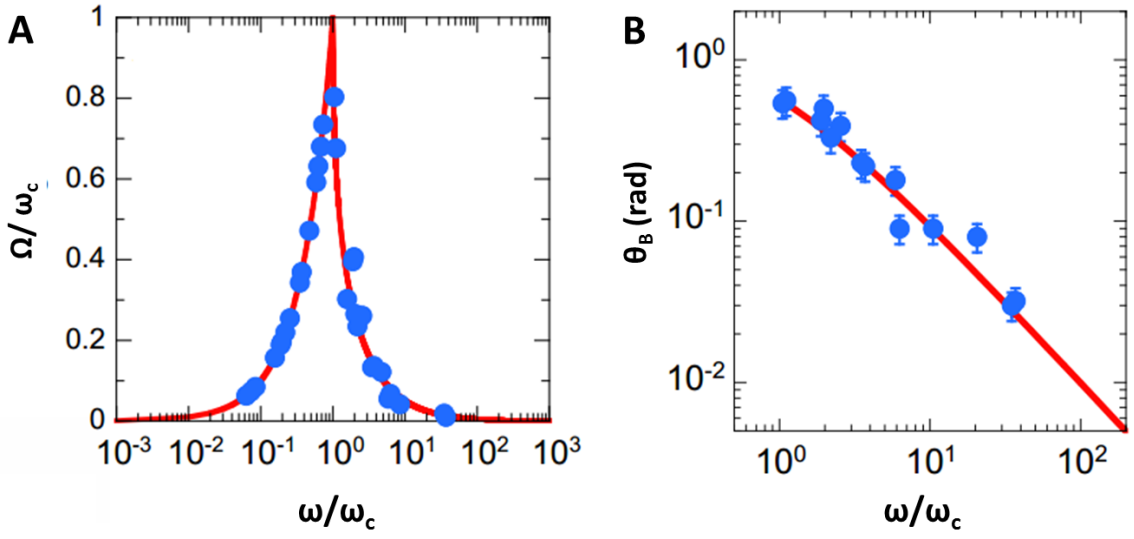


Figure 54 – Analyses rhéologiques de l'échantillon d'ADN (dAdT)$_{60}$. (A) Analyse de fréquence angulaire moyenne en fonction de la fréquence appliquée, toutes deux divisées par les fréquences critiques mesurées. Les points bleus correspondent aux valeurs expérimentales obtenues et la courbe rouge est un ajustement calculé pour un fluide purement visqueux. (B) Analyse des amplitudes d'oscillation en fonction de la fréquence réduite. Le bon accord entre le modèle de Newton (courbe rouge) et les données expérimentales (pointillés bleus) montrent que l'échantillon ne présente aucun signe d'élasticité.

L’échantillon ne présente aucun signe d'élasticité, il se comporte comme un fluide de Maxwell. Ce résultat est en accord avec une viscosité mesurée qui est proche de celle d’un solvant. La viscosité mesurée est η = 1,62 ± 0,32 mPa s. La solution d’ADN seule est très diluée et ne change quasiment pas les propriétés d’écoulement de l’eau.

La viscosité est mesurée à partir du coefficient directeur de la droite obtenu en étudiant la fréquence critique en fonction des longueurs des aiguilles réduites L*. L’expression utilisée pour l’ajustement de courbe est la suivante : $\omega_c = \frac{3}{8}\frac{\mu_0 \Delta\chi H_0^2}{\eta L^*}$ (figure 55).

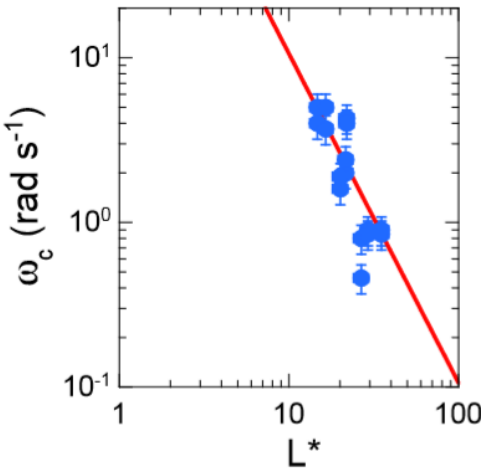
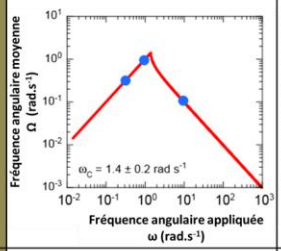
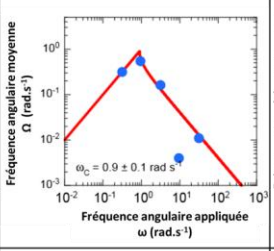
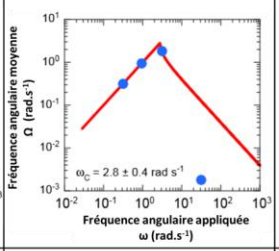
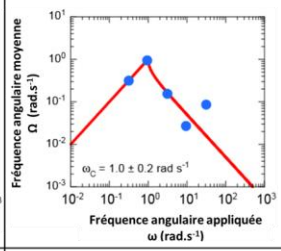
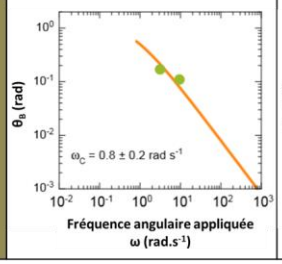
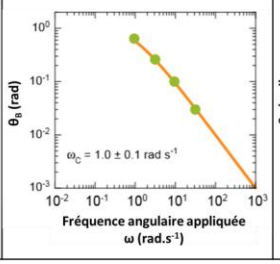
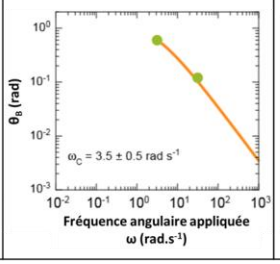
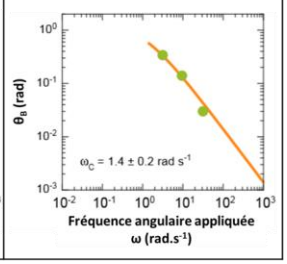


Figure 55 – Analyse de la fréquence critique en fonction de la longueur des aiguilles réduites L*. Cette analyse permet de déterminer la viscosité par l’intermédiaire du coefficient directeur de la droite.

5.2. Peptide CTR38 seul

Un ensemble de 4 micro-aiguilles a été étudié durant l’expérience ayant chacun des formes différentes. Les ω_c ont été déterminées et présentées dans le tableau suivant :

Aiguille	1	2	3	4
L (µm)	45,9	47	28	34
D (µm)	2	1,6	1,5	1,6
Ω(ω)				
$\theta_B(\omega)$				

Les résultats révèlent les mêmes profils que l'échantillon $(dAdT)_{60}$ lors de l'analyse de la fréquence angulaire moyenne divisée par la fréquence critique mesurées en fonction de la fréquence réduite. Il en va de même pour l'analyse de l'amplitude d'oscillations en fonction de la fréquence réduite (figure 56).

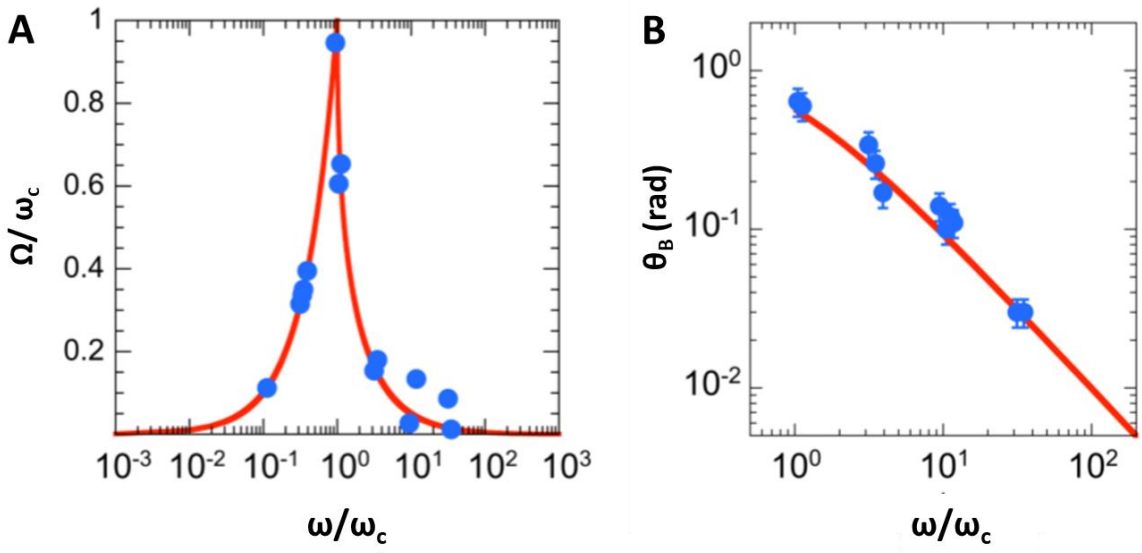


Figure 56 – Analyses rhéologiques de l'échantillon CTR38. (A) Analyse de fréquence angulaire moyenne en fonction de la fréquence appliquée, divisée par les fréquences critiques. Les points bleus correspondent aux valeurs obtenues expérimentales et la courbe rouge est un ajustement calculé pour un fluide purement visqueux. (B) Analyse des amplitudes d'oscillations en fonction de la fréquence

réduite. Le bon accord entre le modèle de Newton (courbe rouge) et les données expérimentales (pointillés bleus) montre que l'échantillon ne présente aucun signe d'élasticité.

La viscosité mesurée pour l'échantillon du CTR38 à 7 mg.ml^{-1} est égale à $\eta = 3{,}31 \pm 0{,}84$ mPa s, ce qui est supérieur à celle de l'eau. Mais il ne semble pas y avoir d'élasticité dans cet échantillon. À cette concentration de peptide, on s'attend aussi à mesurer la viscosité du solvent. Ici c'est légèrement plus élevé, ce qui pourrait suggérer que les CTR38 sont peu agrégés.

5.3. Complexe CTR38 + $(dAdT)_{60}$

Cinq micro-aiguilles ayant des formes différentes ont été étudiées. A la différence des deux premiers échantillons, le complexe CTR38+$(dAdT)_{60}$ se comporte comme un gel. Avec : $\Omega(\omega) \approx 0$ et θ_B dépend des formes des aiguilles et de ω. Les résultats de l'ensemble des aiguilles sont présentés dans la figure 57.

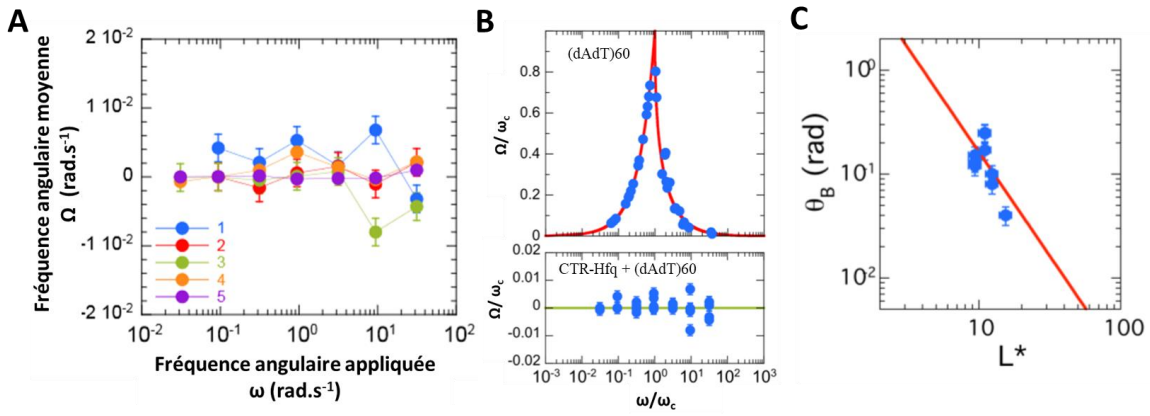


Figure 57 – Analyses rhéologiques de l'échantillon complexe CTR-Hfq + $(dAdT)_{60}$. (A) Dans le cas d'un gel, la fonction $\Omega(\omega) \approx 0$, c'est ce qu'on observe pour cinq aiguilles différentes. (B) Comparaison entre la solution d'ADN et celle du complexe ADN-peptide. Les deux comportements sont très différents : le graphique supérieur est caractéristique d'un liquide, tandis que le graphique inférieur est similaire à celui d'un gel. Dans ce dernier, il n'existe pas de fréquence critique car à toutes les fréquences, les aiguilles sont en régime asynchrone. (C) Analyse de l'amplitude des oscillations θ_B en fonction de la longueur des aiguilles réduite L*. Cette analyse permet de déterminer le module élastique par l'intermédiaire du coefficient directeur de la droite.

Le complexe présente la réponse d'un solide mou. A partir de l'expression θ_B en fonction de L*, on peut déduire une estimation du module élastique G' du matériau. Il est d'environ 0,2

Pa, indiquant un comportement de gel faible. Afin de trouver *G'*, expression utilisée pour l'ajustement est la suivante : $\theta_b = \frac{3}{4}\frac{\mu_0 \Delta\chi H_0^2}{G\, L^{*2}}$

La viscosité de cet échantillon est « infinie », et le mélange ADN + CTR38 peut être décrit comme un fluide à seuil. L'association entre ADN et CTR est forte, elle structure le fluide, sans doute par la formation d'un réseau connecté entre ces deux espèces.

Les différences entre ADN $(dAdT)_{60}$ seul, CTR38 seul et le complexe ADN + CTR38 sont remarquables. En effet, le complexe forme un gel. Cette observation permet de clarifier potentielle le rôle potentiel qu'a Hfq à interagir avec l'ADN et formerait avec lui un réseau de fibres. *In vivo,* un environnement similaire au sein du nucléoïde est possible. Cette viscosité affecte probablement la diffusion des protéines, telles que des facteurs de transcription ou des polymérases. Malheureusement, à ce jour il est difficile de faire cette analyse *in vivo* dans une bactérie car nous n'avons pas trouvé encore un moyen de faire pénétrer l'aiguille dans la cellule. Cependant un des moyens serait l'utilisation d'un « canon à gène », utilisé en biologie végétale.

Conclusion

La virulence bactérienne et la résistance aux antibiotiques sont des problèmes majeurs de santé publique. Il en va de même concernant les amyloïdes. Dans le cadre de cette thèse, on s'est intéressé à ces deux problèmes en étudiant la protéine Hfq, facteur de virulence bactérien. Hfq fait en effet partie des protéines formant des amyloïdes dits fonctionnels grâce à son domaine CTR amyloïde.

Ce travail de thèse a permis d'étudier cette région amyloïde en vue d'en comprendre sa structure et son rôle fonctionnel, mais aussi de chercher un moyen d'en inhiber la formation. Compte tenu du rôle de Hfq dans la vie bactérienne, cette inhibition pourrait en effet ouvrir la voie vers le développement de nouveaux antibiotiques.

La fonction majeure de ce domaine est de donner à Hfq une aptitude à s'auto-assembler *in vivo* et *in vitro*. Pour notre étude, nous avons travaillé essentiellement sur deux peptides dont les séquences sont comprises dans la région C-terminale de Hfq, mais aussi sur la protéine purifiée entière. On remarque que, plus la séquence amyloïde est longue, plus les cinétiques sont complexes et difficiles à comprendre, notamment pour la protéine Hfq entière. C'est pour cette raison que l'on a étudié des peptides issus du domaine CTR qui forment spontanément des fibres amyloïdes. Mon travail de thèse a permis de mieux comprendre les conditions physico-chimiques favorisant la formation de cette structure, notamment la concentration, la température et la force ionique. Cependant, la cinétique de formation des structures amyloïdes *in vitro* reste assez différente de celle obtenue en conditions physiologiques, ce qui nous a amené à poursuivre nos recherches sur un éventuel cofacteur cellulaire qui permettrait de favoriser l'agrégation.

Le domaine CTR de Hfq est capable d'interagir avec l'ADN et de le ponter, permettant la condensation de l'ADN bactérien. Ce travail a mis en évidence que l'ADN favorise la formation de fibres amyloïde de ce domaine, en particulier les doubles hélices d'ADN riches en adénosines. Ce mécanisme de condensation de l'ADN induit par les protéines amyloïdes était encore méconnu mais pourrait représenter un nouveau moyen de condenser l'ADN dans la cellule. Cette interaction pourrait avoir des conséquences bénéfiques, comme pour Hfq (compaction de l'ADN pour compenser l'absence de membrane) ou délétères comme cela pourrait être le cas des amyloïdes pathologiques. L'analyse de la relation existant entre Hfq et l'ADN s'avère donc primordiale afin de mieux comprendre la compaction de l'ADN par une amyloïde et ses conséquences fonctionnelles.

Enfin, une importante partie des travaux de cette thèse a consisté à développer une nouvelle voie pour l'analyse des nanostructures. Nous avons mis en place une technique d'imagerie corrélative inédite. Celle-ci permet d'étudier le même nano-objet isolé avec la microscopie électronique en transmission et la technique de nano-spectroscopie infrarouge (AFM-IR) en déposant l'échantillon sur un substrat de nitrure de silicium. Même s'il existe des limites pour l'imagerie chimique sur la surface corrélative, la technique d'imagerie corrélative permet de mieux analyser les nano-objets tels que les fibres amyloïde. Les améliorations autour de cette technique permettraient des analyses avancées, et nous donneraient des informations telles les caractéristiques morphologiques, les changements de conformation mais également la cartographie chimique du complexe amyloïde-ADN.

Bibliographie

[1] C. P. J. Maury, “The emerging concept of functional amyloid,” *J. Intern. Med.*, vol. 265, no. 3, pp. 329–334, 2009.

[2] A. Daskalov *et al.*, “Identification of a novel cell death-inducing domain reveals that fungal amyloid-controlled programmed cell death is related to necroptosis,” *Proc. Natl. Acad. Sci.*, vol. 113, no. 10, pp. 2720–2725, 2016.

[3] M. T. Franze de Fernandez, L. Eoyang, and T. August, “Factor Fraction required for the Synthesis of Bacteriophage Qβ-RNA,” *Nature*, vol. 219, pp. 588–590, 1968.

[4] A. Taghbalout, Q. Yang, and V. Arluison, “The *Escherichia coli* RNA processing and degradation machinery is compartmentalized within an organized cellular network,” *Biochem. J.*, vol. 458, no. 1, pp. 11–22, 2014.

[5] V. Arluison *et al.*, “Three-dimensional structures of fibrillar Sm proteins: Hfq and other Sm-like proteins,” *J. Mol. Biol.*, vol. 356, no. 1, pp. 86–96, 2006.

[6] E. Fortas *et al.*, “New insight into the structure and function of Hfq C-terminus,” *Biosci. Rep.*, vol. 35, no. 2, pp. 1–9, 2015.

[7] D. G. Scofield and M. Lynch, “Evolutionary diversification of the Sm family of RNA-associated proteins,” *Mol. Biol. Evol.*, vol. 25, no. 11, pp. 2255–2267, 2008.

[8] X. Sun, I. Zhulin, and R. M. Wartell, “Predicted structure and phyletic distribution of the RNA-binding protein Hfq.,” *Nucleic Acids Res.*, vol. 30, no. 17, pp. 3662–71, 2002.

[9] S. Tharun, W. He, A. E. Mayes, P. Lennertz, J. D. Beggs, and R. Parker, “Yeast Sm-like proteins function in mRNA decapping and decay,” *Nature*, vol. 404, no. 6777, pp. 515–518, 2000.

[10] A. G. Seto, A. J. Zaug, S. G. Sobel, S. L. Wolin, and T. R. Cech, “Saccharomyces cerevisiae telomerase is an Sm small nuclear ribonucleoprotein particle,” *Nature*, vol. 401, no. 6749, pp. 177–180, 1999.

[11] J. M. Moll, M. Sobti, and B. C. Mabbutt, “The Lsm Proteins: Ring Architectures for RNA Capture,” in *RNA Processing Prof. Paula Grabowski (Ed.)*, 2011, pp. 229–248.

[12] E. Sauer, “Structure and RNA-binding properties of the bacterial LSm protein Hfq,” *RNA Biol.*, vol. 10, no. 4, pp. 610–618, 2013.

[13] M. A. Schumacher, R. F. Pearson, T. M??ller, P. Valentin-Hansen, and R. G. Brennan, “Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: A bacterial Sm-like protein,” *EMBO J.*, vol. 21, no. 13, pp. 3546–3556, 2002.

[14] C. Sauter, J. Basquin, and D. Suck, “Sm-like proteins in Eubacteria: The crystal structure of the Hfq protein from Escherichia coli,” *Nucleic Acids Res.*, vol. 31, no. 14, pp. 4091–4098, 2003.

[15] A. Nikulin *et al.*, “Structure of Pseudomonas aeruginosa Hfq protein,” *Acta*

Crystallogr. Sect. D Biol. Crystallogr., vol. 61, no. 2, pp. 141–146, 2005.

[16] D. Schilling and U. Gerischer, "The Acinetobacter baylyi hfq gene encodes a large protein with an unusual C terminus," *J. Bacteriol.*, vol. 191, no. 17, pp. 5553–5562, 2009.

[17] A. Sharma *et al.*, "The unusual glycine rich C-terminus of the *Acinetobacter baumannii* RNA chaperone Hfq plays an important role in bacterial physiology," *J. Biol. Chem.*, p. jbc.RA118.002921, 2018.

[18] M. Beich-Frandsen *et al.*, "Structural insights into the dynamics and function of the C-terminus of the E. coli RNA chaperone Hfq," *Nucleic Acids Res.*, vol. 39, no. 11, pp. 4900–4915, 2011.

[19] K. Jiang *et al.*, "Effects of Hfq on the conformation and compaction of DNA," *Nucleic Acids Res.*, vol. 43, no. 8, pp. 4332–4341, 2015.

[20] Z. R. Yang, R. Thomson, P. McNeil, and R. M. Esnouf, "RONN: The bio-basis function neural network technique applied to the detection of natively disordered regions in proteins," *Bioinformatics*, vol. 21, no. 16, pp. 3369–3376, 2005.

[21] J. Kyte and R. F. Doolittle, "A simple method for displaying the hydropathic character of a protein," *J. Mol. Biol.*, vol. 157, no. 1, pp. 105–132, 1982.

[22] J. R. Feliciano, A. M. Grilo, S. I. Guerreiro, S. A. Sousa, and J. H. Leitão, "Hfq: a multifaceted RNA chaperone involved in virulence," *Future Microbiol.*, vol. 11, no. 1, pp. 137–151, 2016.

[23] H. T. Tsui, H. E. Leung, and M. E. Winkler, "Characterization of broadly pleiotropic phenotypes caused by an hfq insertion mutation in Escherichia coli K-12," *Mol. Microbiol.*, vol. 13, no. 1, pp. 35–49, 1994.

[24] N. Zambrano, P. P. Guichard, Y. Bi, B. Cayrol, S. Marco, and V. Arluison, "Involvement of HFq protein in the post-transcriptional regulation of E. coli bacterial cytoskeleton and cell division proteins," *Cell Cycle*, vol. 8, no. 15, pp. 2470–2472, 2009.

[25] B. Cayrol *et al.*, "Riboregulation of the bacterial actin-homolog MreB by DsrA small noncoding RNA," *Integr. Biol. (United Kingdom)*, vol. 7, no. 1, pp. 128–141, 2015.

[26] B. VECEREK, I. MOLL, and U. BLASI, "Translational autocontrol of the Escherichia coli hfq RNA chaperone gene," *Rna*, vol. 11, no. 6, pp. 976–984, 2005.

[27] G. M. Cech, B. Pakuła, D. Kamrowska, G. Wegrzyn, V. Arluison, and A. Szalewska-Pałasz, "Hfq protein deficiency in Escherichia coli affects ColE1-like but not λ plasmid DNA replication," *Plasmid*, vol. 73, pp. 10–15, 2014.

[28] J. Vogel and B. F. Luisi, "Hfq and its constellation of RNA," *Nature Reviews Microbiology*, vol. 9, no. 8. pp. 578–589, 2011.

[29] V. Arluison, S. Hohng, R. Roy, O. Pellegrini, P. Régnier, and T. Ha, "Spectroscopic observation of RNA chaperone activities of Hfq in post-transcriptional regulation by a small non-coding RNA," *Nucleic Acids Res.*, vol. 35, no. 3, pp. 999–1006, 2007.

[30] A. Fender, J. Elf, K. Hampel, B. Zimmermann, and E. G. H. Wagner, "RNAs actively cycle on the Sm-like protein Hfq," *Genes Dev.*, vol. 24, no. 23, pp. 2621–2626, 2010.

[31] W. Hwang, V. Arluison, and S. Hohng, "Dynamic competition of DsrA and rpoS fragments for the proximal binding site of Hfq as a means for efficient annealing," *Nucleic Acids Res.*, vol. 39, no. 12, pp. 5131–5139, 2011.

[32] E. G. H. Wagner, "Cycling of RNAs on Hfq," *RNA Biology*, vol. 10, no. 4. Taylor & Francis, pp. 619–626, 27-Apr-2013.

[33] C. K. Vanderpool and S. Gottesman, "Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system," *Mol. Microbiol.*, vol. 54, no. 4, pp. 1076–1089, 2004.

[34] F. Mika *et al.*, "Targeting of csgD by the small regulatory RNA RprA links stationary phase, biofilm formation and cell envelope stress in Escherichia coli," *Mol. Microbiol.*, vol. 84, no. 1, pp. 51–65, 2012.

[35] K. I. Udekwu, F. Darfeuille, J. Vogel, J. Reimegård, E. Holmqvist, and E. G. H. Wagner, "Hfq-dependent regulation of OmpA synthesis is mediated by an antisense RNA," *Genes Dev.*, vol. 19, no. 19, pp. 2355–2366, 2005.

[36] E. Massé, C. K. Vanderpool, and S. Gottesman, "Effect of RyhB small RNA on global iron use in Escherichia coli," *J. Bacteriol.*, vol. 187, no. 20, pp. 6962–6971, 2005.

[37] H. Aiba, "Mechanism of RNA silencing by Hfq-binding small RNAs," *Curr. Opin. Microbiol.*, vol. 10, no. 2, pp. 134–139, 2007.

[38] T. Morita, K. Maki, and H. Aiba, "RNase E-based ribonucleoprotein complexes: Mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs," *Genes Dev.*, vol. 19, no. 18, pp. 2176–2186, 2005.

[39] N. Majdalani, C. Cunning, D. Sledjeski, T. Elliott, and S. Gottesman, "DsrA RNA regulates translation of RpoS message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription," *Proc. Natl. Acad. Sci.*, vol. 95, no. 21, pp. 12462–12467, 1998.

[40] N. Majdalani, D. Hernandez, and S. Gottesman, "Regulation and mode of action of the second small RNA activator of RpoS translation, RprA," *Mol. Microbiol.*, vol. 46, no. 3, pp. 813–826, 2002.

[41] I. Moll, T. Afonyushkin, O. Vytvytska, V. Kaberdin, and U. Bläsi, "Coincident Hfq binding and RNase E cleavage sites on mRNA and small regulatory RNAs," *RNA*, vol. 9, no. 11, pp. 1308–1314, 2003.

[42] M. Rabhi *et al.*, "The Sm-like RNA chaperone Hfq mediates transcription

antitermination at Rho-dependent terminators," *EMBO J.*, vol. 30, pp. 2805–2816, 2011.

[43] J. Le Derout, I. V Boni, P. Régnier, and E. Hajnsdorf, "Hfq affects mRNA levels independently of degradation," *BMC Mol. Biol.*, vol. 11, no. 17, 2010.

[44] M. Folichon *et al.*, "Fate of mRNA extremities generated by intrinsic termination: Detailed analysis of reactions catalyzed by ribonuclease II and poly(A) polymerase," *Biochimie*, vol. 87, no. 9–10, pp. 819–826, 2005.

[45] L. Argaman, M. Elgrably-Weiss, T. Hershko, J. Vogel, and S. Altuvia, "RelA protein stimulates the activity of RyhB small RNA by acting on RNA-binding protein Hfq," *Proc. Natl. Acad. Sci.*, vol. 109, no. 12, pp. 4621–4626, 2012.

[46] A. Smirnov, C. Wang, L. L. Drewry, and J. Vogel, "Molecular mechanism of mRNA repression in trans by a ProQ-dependent small RNA," *EMBO J.*, vol. 36, pp. 1029–1045, 2017.

[47] V. Arluison *et al.*, "The C-terminal domain of Escherichia coli Hfq increases the stability of the hexamer," *Eur. J. Biochem.*, vol. 271, no. 7, pp. 1258–1265, Mar. 2004.

[48] T. M. Link, P. Valentin-Hansen, and R. G. Brennan, "Structure of Escherichia coli Hfq bound to polyriboadenylate RNA," *Proc. Natl. Acad. Sci.*, vol. 106, no. 46, pp. 19292–19297, 2009.

[49] T. A. Geissmann and D. Touati, "Hfq, a new chaperoning role: Binding to messenger RNA determines access for small RNA regulator," *EMBO J.*, vol. 23, no. 2, pp. 396–405, 2004.

[50] D. J. Schu, A. Zhang, S. Gottesman, and G. Storz, "Alternative Hfq-sRNA interaction modes dictate alternative mRNA recognition," *EMBO J.*, vol. 34, no. 20, pp. 2557–2573, 2015.

[51] M. V. Sukhodolets and S. Garges, "Interaction of Escherichia coli RNA polymerase with the ribosomal protein S1 and the Sm-like ATPase Hfq," *Biochemistry*, vol. 42, no. 26, pp. 8022–8034, 2003.

[52] E. Hajnsdorf, P. Régnier, and P. Regnier, "Host factor Hfq of Escherichia coli stimulates elongation of poly(A) tails by poly(A) polymerase I.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 97, no. 4, pp. 1501–1505, 2000.

[53] A. Malabirade *et al.*, "Compaction and condensation of DNA mediated by the C-terminal domain of Hfq," *Nucleic Acids Res.*, vol. 45, no. 12, pp. 7299–7308, 2017.

[54] G. M. Cech *et al.*, "The Escherichia Coli Hfq Protein: An Unattended DNA-Transactions Regulator," *Front. Mol. Biosci.*, vol. 3, no. 36, 2016.

[55] C. J. Dorman, *Nucleoid-Associated Proteins and Bacterial Physiology*, 1st ed., vol. 67, no. C. Elsevier Inc., 2009.

[56] T. A. Azam, A. Iwata, A. Nishimura, T. a L. I. Azam, and S. Ueda, “Growth Phase-Dependent Variation in Protein Composition of the Escherichia coli Nucleoid Growth Phase-Dependent Variation in Protein Composition of the Escherichia coli Nucleoid,” *J. Bac*, vol. 181, no. 20, p. 6361, 1999.

[57] T. A. Azam and A. Ishihama, “Twelve species of the nucleoid-associated protein from Escherichia coli. Sequence recognition specificity and DNA binding affinity,” *J. Biol. Chem.*, vol. 274, no. 46, pp. 33105–33113, 1999.

[58] T. A. Azam, S. Hiraga, and A. Ishihama, “Two types of localization of the DNA-binding proteins within the Escherichia coli nucleoid,” *Genes to Cells*, vol. 5, no. 8, pp. 613–626, 2000.

[59] A. Grove, “Functional evolution of bacterial histone-like HU proteins,” *Curr. Issues Mol. Biol.*, vol. 13, no. 1, pp. 1–12, 2011.

[60] C. J. Dorman, “H-NS: A universal regulator for a dynamic genome,” *Nat. Rev. Microbiol.*, vol. 2, no. 5, pp. 391–400, 2004.

[61] S. T. Arold, P. G. Leonard, G. N. Parkinson, and J. E. Ladbury, “H-NS forms a superhelical protein scaffold for DNA condensation.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 107, no. 36, pp. 15728–32, Sep. 2010.

[62] P. A. Wiggins, R. T. Dame, M. C. Noom, and G. J. L. Wuite, “Protein-mediated molecular bridging: A key mechanism in biopolymer organization,” *Biophys. J.*, vol. 97, no. 7, pp. 1997–2003, 2009.

[63] B. R. G. Gordon *et al.*, “Structural basis for recognition of AT-rich DNA by unrelated xenogeneic silencing proteins,” *Proc. Natl. Acad. Sci.*, vol. 108, no. 26, pp. 10690–10695, 2011.

[64] X. Shi and G. N. Bennett, “Plasmids bearing hfq and the hns-like gene stpA complement hns mutants in modulating arginine decarboxylase gene expression in Escherichia coli,” *Journal of Bacteriology*, vol. 176, no. 21. pp. 6769–6775, 1994.

[65] R. A. Lease, M. E. Cusick, and M. Belfort, “Riboregulation in Escherichia coli: DsrA RNA acts by RNA:RNA interactions at multiple loci,” *Proc. Natl. Acad. Sci.*, vol. 95, no. 21, pp. 12456–12461, 1998.

[66] A. Badrinarayanan, T. B. Le, and M. T. Laub, “Bacterial Chromosome Organization and Segregation,” *Annu. Rev. Cell Dev. Biol.*, vol. 31, no. 1, pp. 171–199, 2015.

[67] E. Diestra, B. Cayrol, V. Arluison, and C. Risco, “Cellular electron microscopy imaging reveals the localization of the hfq protein close to the bacterial membrane,” *PLoS One*, vol. 4, no. 12, p. e8301, 2009.

[68] M. Guillier and S. Gottesman, “Remodelling of the Escherichia coli outer membrane by two small regulatory RNAs,” *Mol. Microbiol.*, vol. 59, no. 1, pp. 231–247, 2006.

[69] V. Arluison and A. Taghbalout, “Cellular localization of RNA degradation and

processing components in Escherichia coli.," in *Methods Mol Biol*, p. 1259: 87-101, 2015.

[70] D. Partouche *et al.*, "Techniques to analyse sRNA protein cofactor self-assembly in vitro," in *Bacterial Regulatory RNA: Methods and Protocols, Methods in Molecular Biology*, vol. 1737, no. 1980, 2018, pp. 321–340.

[71] A. Malabirade *et al.*, "Membrane association of the bacterial riboregulator Hfq and functional perspectives," *Sci. Rep.*, vol. 7, no. 1, pp. 1–12, 2017.

[72] A. S. Olsen, J. Møller-Jensen, R. G. Brennan, and P. Valentin-Hansen, "C-Terminally Truncated Derivatives of Escherichia coli Hfq Are Proficient in Riboregulation," *J. Mol. Biol.*, vol. 404, no. 2, pp. 173–182, 2010.

[73] A. Santiago-Frangos, K. Kavita, D. J. Schu, S. Gottesman, and S. A. Woodson, "C-terminal domain of the RNA chaperone Hfq drives sRNA competition and release of target RNA," *Proc. Natl. Acad. Sci.*, vol. 113, no. 41, pp. E6089–E6096, 2016.

[74] V. N. Uversky, "Intrinsically disordered proteins from A to Z," *Int. J. Biochem. Cell Biol.*, vol. 43, no. 8, pp. 1090–1103, 2011.

[75] T. Zhang, E. Faraggi, Z. Li, and Y. Zhou, "Intrinsically Semi-disordered State and Its Role in Induced Folding and Protein Aggregation," *Cell Biochem. Biophys.*, vol. 67, no. 3, pp. 1193–1205, 2013.

[76] S. Fukuchi, K. Hosoda, K. Homma, T. Gojobori, and K. Nishikawa, "Binary classification of protein molecules into intrinsically disordered and ordered segments," *BMC Struct. Biol.*, vol. 11, no. 1, p. 29, 2011.

[77] G. Grateau, J. Verine, M. Delpech, and M. Ries, "Les amyloses, un modèle de maladie du repliement des protéines," *Medecine/Sciences*, vol. 21, no. 6–7. pp. 627–633, 2005.

[78] C. M. Dobson, "Protein folding and misfolding," *Nature*, vol. 426, no. 6968, pp. 884–890, 2003.

[79] F. Chiti and C. M. Dobson, "Protein Misfolding, Amyloid Formation, and Human Disease: A Summary of Progress Over the Last Decade," *Annu. Rev. Biochem.*, vol. 86, no. May, pp. 27–68, 2017.

[80] T. P. J. Knowles, M. Vendruscolo, and C. M. Dobson, "The amyloid state and its association with protein misfolding diseases," *Nature Reviews Molecular Cell Biology*, vol. 15, no. 6. pp. 384–396, 2014.

[81] A. T. Petkova *et al.*, "A structural model for Alzheimer's β-amyloid fibrils based on experimental constraints from solid state NMR.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 99, no. 26, pp. 16742–7, 2002.

[82] A. T. Petkova, W.-M. Yau, and R. Tycko, "Experimental Constraints on Quaternary Structure in Alzheimer's β-Amyloid Fibrils," *Biochemistry*, vol. 45, no. 2, pp. 498–512, 2006.

[83] M. Fändrich and C. M. Dobson, “The behaviour of polyamino acids reveals an inverse side chain effect in amyloid structure formation,” *EMBO J.*, vol. 21, no. 21, pp. 5682–5690, 2002.

[84] F. Ferrone, “[17] Analysis of protein aggregation kinetics,” *Methods Enzymol.*, vol. 309, no. 1, pp. 256–274, 1999.

[85] T. R. Serio *et al.*, “Nucleated conformational conversion and the replication of conformational information by a prion determinant,” *Science (80-.).*, vol. 289, no. 5483, pp. 1317–1321, 2000.

[86] P. Arosio, T. P. J. Knowles, and S. Linse, “On the lag phase in amyloid fibril formation,” *Phys. Chem. Chem. Phys.*, vol. 17, no. 12, pp. 7606–7618, 2015.

[87] S. I. A. Cohen *et al.*, “Proliferation of amyloid-β42 aggregates occurs through a secondary nucleation mechanism,” *Proc. Natl. Acad. Sci.*, vol. 110, no. 24, pp. 9758–9763, Jun. 2013.

[88] F. S. Ruggeri, G. Longo, S. Faggiano, E. Lipiec, A. Pastore, and G. Dietler, “Infrared nanospectroscopy characterization of oligomeric and fibrillar aggregates during amyloid formation,” *Nat. Commun.*, vol. 6, pp. 1–9, 2015.

[89] J. F. Smith, T. P. J. Knowles, C. M. Dobson, C. E. MacPhee, and M. E. Welland, “Characterization of the nanoscale properties of individual amyloid fibrils,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 103, no. 43, pp. 15806–15811, Nov. 2006.

[90] M. Sunde and C. Blake, “The Structure of Amyloid Fibrils by Electron Microscopy and X-Ray Diffraction,” in *Advences in protein chemistry*, vol. 50 ,pp. 123–159, 1997.

[91] E. . Eanes and G. G. Glenner, “X-ray diffraction studies on amyloid filaments,” *J. Histochem. Cytochem.*, vol. 16, no. 11, pp. 673–677, 1968.

[92] G. Zandomeneghi, M. R. H. Krebs, M. G. McCammon, and M. Fändrich, “FTIR reveals structural differences between native β-sheet proteins and amyloid fibrils,” *Protein Sci.*, vol. 13, no. 12, pp. 3314–3321, 2004.

[93] W. E. Klunk, R. F. Jacob, and R. P. Mason, “Quantifying amyloid by congo red spectral shift assay,” *Methods Enzymol.*, vol. 309, no. 1974, pp. 285–305, 1999.

[94] H. Levine, “Thioflavine t interaction with amyloid βsheet structures,” *Amyloid*, vol. 2, no. 1, pp. 1–6, 1995.

[95] M. R. Nilsson, “Techniques to study amyloid fibril formation in vitro,” *Methods*, vol. 34, no. 1, pp. 151–160, 2004.

[96] R. Sabate, F. Rousseau, J. Schymkowitz, C. Batlle, and S. Ventura, “Amyloids or prions? That is the question,” *Prion*, vol. 9, no. 3, pp. 200–206, 2015.

[97] J. Shorter and S. Lindquist, “Prions as adaptive conduits of memory and inheritance,” *Nat. Rev. Genet.*, vol. 6, no. 6, pp. 435–450, 2005.

[98] V. V. Kushnirov, A. B. Vishnevskaya, I. M. Alexandrov, and M. D. Ter-Avanesyan, “Prion and nonprion amyloids: a comparison inspired by the yeast Sup35 protein.,” *Prion*, vol. 1, no. 3, pp. 179–184, 2007.

[99] H. J. Dyson and P. E. Wright, “Intrinsically unstructured proteins and their functions,” *Nat. Rev. Mol. Cell Biol.*, vol. 6, no. 3, pp. 197–208, 2005.

[100] K. Namba, “Roles of partly unfolded conformations in macromolecular self-assembly,” *Genes to Cells*, vol. 6, no. 1, pp. 1–12, 2001.

[101] D. M. Fowler, A. V. Koulov, W. E. Balch, and J. W. Kelly, “Functional amyloid - from bacteria to humans,” *Trends Biochem. Sci.*, vol. 32, no. 5, pp. 217–224, 2007.

[102] C. P. J. Maury, “Amyloid and the origin of life: self-replicating catalytic amyloids as prebiotic informational and protometabolic entities,” *Cell. Mol. Life Sci.*, vol. 75, no. 9, pp. 1499–1507, 2018.

[103] J. Greenwald, W. Kwiatkowski, and R. Riek, “Peptide Amyloids in the Origin of Life,” *Journal of Molecular Biology*, Jun-2018.

[104] K. Matsuzaki and C. Horikiri, “Interactions of amyloid β-peptide (1-40) with ganglioside-containing membranes,” *Biochemistry*, vol. 38, no. 13, pp. 4137–4142, 1999.

[105] B. R. Lentz, “Membrane ‘fluidity’ as detected by diphenylhexatriene probes,” *Chem. Phys. Lipids*, vol. 50, no. 3–4, pp. 171–190, 1989.

[106] R. Bomba, W. Kwiatkowski, R. Riek, and J. Greenwald, “Cooperative induction of ordered peptide and fatty acid aggregates,” *bioRxiv*, p. 323030, May 2018.

[107] S. Braun, C. Humphreys, E. Fraser, A. Brancale, M. Bochtler, and T. C. Dale, “Amyloid-Associated Nucleic Acid Hybridisation,” *PLoS One*, vol. 6, no. 5, p. e19125, May 2011.

[108] N. Jain *et al.*, “Inhibition of curli assembly and Escherichia coli biofilm formation by the human systemic amyloid precursor transthyretin,” *Proc. Natl. Acad. Sci.*, vol. 114, no. 46, pp. 12184–12189, 2017.

[109] J. D. Taylor and S. J. Matthews, “New insight into the molecular control of bacterial functional amyloids,” *Front. Cell. Infect. Microbiol.*, vol. 5, no. April, pp. 1–7, 2015.

[110] A. Ostrowski, A. Mehert, A. Prescott, T. B. Kiley, and N. R. Stanley-Wall, “YuaB functions synergistically with the exopolysaccharide and TasA amyloid fibers to allow biofilm formation by Bacillus subtilis,” *J. Bacteriol.*, vol. 193, no. 18, pp. 4821–4831, 2011.

[111] J. Oh *et al.*, “Amyloidogenesis of type III-dependent harpins from plant pathogenic bacteria,” *J. Biol. Chem.*, vol. 282, no. 18, pp. 13601–13609, May 2007.

[112] P. Aguilera *et al.*, “Identification of key amino acid residues modulating intracellular

and in vitro Microcin E492 amyloid formation," *Front. Microbiol.*, vol. 7, no. 35, 2016.

[113] I. Pallarès and S. Ventura, "The Transcription Terminator Rho : A First Bacterial Prion," *Trends Microbiol.*, vol. 25, no. 6, pp. 434–437, 2017.

[114] E. Chorell *et al.*, "Bacterial chaperones CsgE and CsgC differentially modulate human α-Synuclein amyloid formation via transient contacts," *PLoS One*, vol. 10, no. 10, p. e0140194, Oct. 2015.

[115] A. Frydman-Marom, R. Shaltiel-Karyo, S. Moshe, and E. Gazit, "The generic amyloid formation inhibition effect of a designed small aromatic β-breaking peptide," *Amyloid*, vol. 18, no. 3, pp. 119–127, 2011.

[116] E. Gazit, "A possible role for π-stacking in the self-assembly of amyloid fibrils," *FASEB J.*, vol. 16, no. 1, pp. 77–83, 2002.

[117] S. Brahmachari, A. Paul, D. Segal, and E. Gazit, "Inhibition of amyloid oligomerization into different supramolecular architectures by small molecules: Mechanistic insights and design rules," *Future Med. Chem.*, vol. 9, no. 8, pp. 797–810, 2017.

[118] K. Fosgerau and T. Hoffmann, "Peptide therapeutics: Current status and future directions," *Drug Discov. Today*, vol. 20, no. 1, pp. 122–128, 2015.

[119] C. Soto, E. M. Sigurdsson, L. Morelli, R. A. Kumar, E. M. Castaño, and B. Frangione, "β-sheet breaker peptides inhibit fibrillogenesis in a rat brain model of amyloidosis: Implications for Alzheimer's therapy," *Nat. Med.*, vol. 4, no. 7, pp. 822–826, 1998.

[120] T. Ziehm *et al.*, "Increase of Positive Net Charge and Conformational Rigidity Enhances the Efficacy of d -Enantiomeric Peptides Designed to Eliminate Cytotoxic Aβ Species," *ACS Chem. Neurosci.*, vol. 7, no. 8, pp. 1088–1096, 2016.

[121] A. A. Reinke and J. E. Gestwicki, "Structure-activity relationships of amyloid beta-aggregation inhibitors based on curcumin: Influence of linker length and flexibility," *Chem. Biol. Drug Des.*, vol. 70, no. 3, pp. 206–215, 2007.

[122] K. Ono, K. Hasegawa, H. Naiki, and M. Yamada, "Curcumin Has Potent Anti-Amyloidogenic Effects for Alzheimer's β-Amyloid Fibrils In Vitro," *J. Neurosci. Res.*, vol. 75, no. 6, pp. 742–750, 2004.

[123] J. Bieschke *et al.*, "EGCG remodels mature α-synuclein and amyloid-β fibrils and reduces cellular toxicity," *Proc. Natl. Acad. Sci.*, vol. 107, no. 17, pp. 7710–7715, 2010.

[124] H. A. Lashuel, D. M. Hartley, D. Balakhaneh, A. Aggarwal, S. Teichberg, and D. J. E. Callaway, "New class of inhibitors of amyloid-β fibril formation: Implications for the mechanism of pathogenesis in Alzheimer's disease," *J. Biol. Chem.*, vol. 277, no. 45, pp. 42881–42890, 2002.

[125] J. A. McLaurin, R. Golomb, A. Jurewicz, J. P. Antel, and P. E. Fraser, “Inositol stereoisomers stabilize an oligomeric aggregate of alzheimer amyloid β peptide and inhibit Aβ-induced toxicity,” *J. Biol. Chem.*, vol. 275, no. 24, pp. 18495–18502, 2000.

[126] A. Sood *et al.*, “Disassembly of preformed amyloid beta fibrils by small organofluorine molecules,” *Bioorganic Med. Chem. Lett.*, vol. 21, no. 7, pp. 2044–2047, 2011.

[127] M. Török, M. Abid, S. C. Mhadgut, and B. Török, “Organofluorine inhibitors of amyloid fibrillogenesis,” *Biochemistry*, vol. 45, no. 16, pp. 5377–5383, 2006.

[128] H. Gong *et al.*, “Effects of several quinones on insulin aggregation,” *Sci. Rep.*, vol. 4, no. 5648, 2014.

[129] A. Barth, “Infrared spectroscopy of proteins,” *Biochim. Biophys. Acta - Bioenerg.*, vol. 1767, no. 9, pp. 1073–1101, 2007.

[130] W. D. Perkins, “Fourier transform infrared spectroscopy. Part III. Advantages of FT-IR,” *J. Chem. Educ.*, vol. 64, no. 11, pp. 296–305, 1987.

[131] W. D. Perkins, “Fourier transform-infrared spectroscopy: Part l. Instrumentation,” *J. Chem. Educ.*, vol. 63, no. 1, p. A5, 1986.

[132] L. Servant, G. Le Bourdon, and T. Buffeteau, “Comprendre la spectroscopie infrarouge : principes et mise en œuvre,” *Photoniques*, vol. 53, pp. 68–73, 2011.

[133] shimadzu France www.shimadzu.fr, “Livre Blanc Spectroscopie infrarouge Un guide pour tout savoir sur la spectroscopie InfraRouge,” 2015. [Online]. Available: https://www.shimadzu.fr/sites/default/files/shimadzu_lb_ftir_2015-bd.pdf. [Accessed: 27-Jun-2018].

[134] P. Dumas, G. L. Carr, and G. P. Williams, “Enhancing the lateral resolution in infrared microspectrometry by using synchrotron radiation: applications and perspectives,” *Analusis*, vol. 28, no. 1, pp. 68–68, 2000.

[135] P. Lasch and D. Naumann, “Spatial resolution in infrared microspectroscopic imaging of tissues,” vol. 1758, pp. 814–829, 2006.

[136] E. C. Mattson *et al.*, “Toward optimal spatial and spectral quality in widefield infrared spectromicroscopy of IR labelled single cells,” *Analyst*, vol. 138, no. 19, p. 5610, 2013.

[137] M. J. Nasse *et al.*, “High-resolution Fourier-transform infrared chemical imaging with multiple synchrotron beams,” *Nat. Methods*, vol. 8, no. 5, pp. 413–416, 2011.

[138] M. Minsky, “Memoir on inventing the confocal scanning microscope,” *Scanning*, vol. 10, no. 4, pp. 128–138, 1988.

[139] G. L. Carr, “Resolution limits for infrared microspectroscopy explored with synchrotron radiation,” *Rev. Sci. Instrum.*, vol. 72, no. 3, pp. 1613–1619, 2001.

[140] G. Binnig and H. Rohrer, “Scanning tunneling microscopy,” *Surf. Sci.*, vol. 126, no. 126, pp. 236–244, 1982.

[141] Binnig Gerd and Rohrer Heinrich, “The Scanning Tunneling Microscope,” *Sci. Am.*, vol. 253, no. April, pp. 50–58, 1985.

[142] G. Binnig and C. F. Quate, “Atomic Force Microscope,” *Phys. Rev. Lett.*, vol. 56, no. 9, pp. 930–933, 1986.

[143] Y. Martin, C. C. Williams, and H. K. Wickramasinghe, “Atomic force microscope-force mapping and profiling on a sub 100-Å scale,” *J. Appl. Phys.*, vol. 61, no. 10, pp. 4723–4729, 1987.

[144] T. R. Albrecht, P. Grütter, D. Horne, and D. Rugar, “Frequency modulation detection using high-Q cantilevers for enhanced force microscope sensitivity,” *J. Appl. Phys.*, vol. 69, no. 2, pp. 668–673, 1991.

[145] L. Bozec, A. Hammiche, M. J. Tobin, J. M. Chalmers, N. J. Everall, and H. m. Pollock, “Near-field photothermal Fourier transform infrared spectroscopy using synchrotron radiation,” *Meas. Sci. Technol.*, vol. 13, no. 8, pp. 1217–1222, 2002.

[146] A. Hammiche, H. M. Pollock, M. Song, and D. J. Hourston, “Photothermal FT-IR Spectroscopy: A Step Towards FT-IR Microscopy at a Resolution Better Than the Diffraction Limit,” *Meas. Sci. Technol.*, vol. 53, no. 7, pp. 810–815, 1999.

[147] A. Hammiche *et al.*, “Mid-infrared microspectroscopy of difficult samples using near-field photothermal microspectroscopy,” *Spectroscopy*, vol. 19, no. 2, pp. 20–30, 2004.

[148] A. Dazzi and C. B. Prater, “AFM-IR: Technology and applications in nanoscale infrared spectroscopy and chemical imaging,” *Chem. Rev.*, vol. 117, no. 7, pp. 5146–5173, 2017.

[149] A. Dazzi, F. Glotin, and R. Carminati, “Theory of infrared nanospectroscopy by photothermal induced resonance,” *Artic. J. Appl. Phys.*, 2010.

[150] F. Lu and M. A. Belkin, “Infrared absorption nano-spectroscopy using sample photoexpansion induced by tunable quantum cascade lasers,” *Opt. Express*, vol. 19, no. 21, pp. 19942–19947, 2011.

[151] F. Lu, M. Jin, and M. A. Belkin, “Tip-enhanced infrared nanospectroscopy via molecular expansion force detection,” *Nat. Photonics*, vol. 8, no. 4, pp. 307–312, 2014.

[152] A. Dazzi, R. Prazeres, F. Glotin, and J. M. Ortega, “Local infrared microspectroscopy with subwavelength spatial resolution with an atomic force microscope tip used as a photothermal sensor,” *Opt. Lett.*, vol. 30, no. 18, p. 2388, 2005.

[153] J. L. Arrondo, a Muga, J. Castresana, and F. M. Goñi, “Quantitative studies of the structure of proteins in solution by Fourier-transform infrared spectroscopy.,” *Prog. Biophys. Mol. Biol.*, vol. 59, no. 1, pp. 23–56, 1993.

[154] J. A. Hering, P. R. Innocent, and P. I. Haris, “Towards developing a protein infrared spectra databank (PISD) for proteomics research,” *Proteomics*, vol. 4, no. 8, pp. 2310–2319, 2004.

[155] P. B. Tooke, “Fourier self-deconvolution in IR spectroscopy,” *Trends Anal. Chem.*, vol. 7, no. 4, pp. 130–136, 1988.

[156] D. M. Byler and H. Susi, “Examination of the secondary structure of proteins by deconvolved FTIR spectra,” *Biopolymers*, vol. 25, no. 3, pp. 469–487, 1986.

[157] J. L. R. Arrondo and F. M. Goñi, “Structure and dynamics of membrane proteins as studied by infrared spectroscopy,” *Progress in Biophysics and Molecular Biology*, vol. 72, no. 4. pp. 367–405, 1999.

[158] R. Sarroukh, E. Goormaghtigh, J. M. Ruysschaert, and V. Raussens, “ATR-FTIR: A ‘rejuvenated’ tool to investigate amyloid proteins,” *Biochim. Biophys. Acta - Biomembr.*, vol. 1828, no. 10, pp. 2328–2338, 2013.

[159] D. Partouche *et al.*, “Techniques to analyze sRNA protein cofactor self-assembly in vitro,” in *Methods in Molecular Biology*, vol. 1737, 2018, pp. 321–340.

[160] V. Militello, C. Casarino, A. Emanuele, A. Giostra, F. Pullara, and M. Leone, “Aggregation kinetics of bovine serum albumin studied by FTIR spectroscopy and light scattering,” *Biophys. Chem.*, vol. 107, no. 2, pp. 175–187, 2004.

[161] S. Englander and N. R. Kallenbach, “Hydrogen exchange and structural dynamics of proteins and nucleic acids,” *Q. Rev. Biophys.*, vol. 16, no. 4, pp. 521–655, 2018.

[162] T. M. Raschke and S. Marqusee, “Hydrogen exchange studies of protein structure,” *Curr. Opin. Biotechnol.*, vol. 9, pp. 80–86, 1998.

[163] S. Cai and B. R. Singh, “A Distinct Utility of the Amide III Infrared Band for Secondary Structure Estimation of Aqueous Protein Solutions Using Partial Least Squares Methods,” *Biochemistry*, vol. 43, no. 9, pp. 2541–2549, 2004.

[164] A. Dessombz, D. Bazin, P. Dumas, C. Sandt, J. Sule-Suso, and M. Daudon, “Shedding Light on the Chemical Diversity of Ectopic Calcifications in Kidney Tissues: Diagnostic and Research Aspects,” *PLoS One*, vol. 6, no. 11, 2011.

[165] F. Le Naour *et al.*, “Quantitative assessment of liver steatosis on tissue section using infrared spectroscopy,” *Gastroenterology*, vol. 148, no. 2, pp. 295–297, 2015.

[166] M. Jackson and H. H. Mantsch, “Infrared Spectroscopy, Ex Vivo Tissue Analysis by,” *Encycl. Anal. Chem.*, pp. 1–26, 2006.

[167] C. Sandt, C. Nadaradjane, R. Richards, P. Dumas, and V. Sée, “Use of infrared microspectroscopy to elucidate a specific chemical signature associated with hypoxia levels found in glioblastoma,” *Analyst*, vol. 141, no. 3, pp. 870–883, 2016.

[168] R. Anthony Shaw *et al.*, “Infrared spectroscopy of biofluids in clinical chemistry and

medical diagnostics," in *Biomedical Vibrational Spectroscopy*, pp. 79–104, 2008.

[169] D. Helm, H. Labischinski, G. Schallehn, and D. Naumann, "Classification and identification of bacteria by Fourier-transform infrared spectroscopy," *J. Gen. Microbiol.*, no. 137, pp. 69–79, 1991.

[170] C. Sandt *et al.*, "FT-IR microspectroscopy for early identification of some clinically relevant pathogens," *J. Appl. Microbiol.*, vol. 101, no. 4, pp. 785–797, 2006.

[171] D. Naumann, "Infrared Spectroscopy in Microbiology," in *Encyclopedia of Analytical Chemistry*, 2000, pp. 102–131.

[172] D. Ami, L. Bonecchi, S. Calì, G. Orsini, G. Tonon, and S. M. Doglia, "FT-IR study of heterologous protein expression in recombinant Escherichia coli strains," *Biochim. Biophys. Acta - Gen. Subj.*, vol. 1624, no. 1–3, pp. 6–10, 2003.

[173] A. C. McGovern, R. Ernill, B. V Kara, D. B. Kell, and R. Goodacre, "Rapid analysis of the expression of heterologous proteins in Escherichia coli using pyrolysis mass spectrometry and Fourier transform infrared spectroscopy with chemometrics: application to alpha 2-interferon production.," *J. Biotechnol.*, vol. 72, no. 3, pp. 157–67, 1999.

[174] D. Ami, A. Natalello, P. Gatti-Lafranconi, M. Lotti, and S. M. Doglia, "Kinetics of inclusion body formation studied in intact cells by FT-IR spectroscopy," *FEBS Lett.*, vol. 579, no. 16, pp. 3433–3436, 2005.

[175] M. Santos, E. Gerbino, E. Tymczyszyn, and A. Gomez-Zavaglia, "Applications of Infrared and Raman Spectroscopies to Probiotic Investigation," *Foods*, vol. 4, no. 3, pp. 283–305, 2015.

[176] B. A. Wallace, "Protein characterisation by synchrotron radiation circular dichroism spectroscopy," *Q. Rev. Biophys.*, vol. 4, no. 42, pp. 317–370, 2009.

[177] R. M. Pagni, "Circular Dichroism and Linear Dichroism (Rodger, Alison; Norden, Bengt)," *J. Chem. Educ.*, vol. 75, no. 9, p. 1095, Sep. 1998.

[178] J. D. Hirst, K. Colella, and A. T. B. Gilbert, "Electronic Circular Dichroism of Proteins from First-Principles Calculations," pp. 11813–11819, 2003.

[179] S. M. M. M. Kelly and N. C. C. C. Price, "The use of circular dichroism in the investigation of protein structure and function.," *Curr. Protein Pept. Sci.*, vol. 1, no. 4, pp. 349–84, 2000.

[180] A. Micsonai *et al.*, "Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy," *Proc Natl Acad Sci U S A*, no. 112, pp. 3095–3103, 2015.

[181] V. I. Dodero, Z. B. Quirolo, and M. A. Sequeira, "Biomolecular studies by circular dichroism," *Front. Biosci.*, vol. 16, no. 1, p. 61, 2011.

[182] J. Kypr, I. Kejnovská, D. Renč Iuk, and M. Vorlíč, "Circular dichroism and conformational polymorphism of DNA," *Nucleic Acids Res.*, vol. 37, no. 6, pp. 1713–1725, 2009.

[183] J. Jaumot, R. Eritja, S. Navea, and R. Gargallo, "Classification of nucleic acids structures by means of the chemometric analysis of circular dichroism spectra," *Anal. Chim. Acta*, vol. 642, no. 1–2, pp. 117–126, 2009.

[184] D. Sirohi *et al.*, "The 3.8 Å resolution cryo-EM structure of Zika virus," *Science (80-.).*, vol. 352, no. 6284, pp. 467–470, 2016.

[185] L. Gremer *et al.*, "Fibril structure of amyloid- b (1 – 42) by cryo – electron microscopy," *Science (80-.).*, vol. 358, no. October, pp. 116–119, 2017.

[186] T. Bizien, D. Durand, P. Roblina, T. Aurélien, P. Vachette, and J. Pérez, "A Brief Survey of State-of-the-Art BioSAXS," *Protein Pept. Lett.*, vol. 23, no. 3, pp. 217–231, 2016.

[187] F. T. S. Chan, D. Pinotsi, G. S. K. Schierle, and C. F. Kaminski, "Structure-Specific Intrinsic Fluorescence of Protein Amyloids Used to Study their Kinetics of Aggregation," in *Bio-nanoimaging: Protein Misfolding & Aggregation*, Elsevier, , pp. 147–156, 2014.

[188] W. Chen *et al.*, "Fluorescence Self-Quenching from Reporter Dyes Informs on the Structural Properties of Amyloid Clusters Formed in Vitro and in Cells," *Nano Lett.*, vol. 17, pp. 143–149, 2017.

[189] T. Härd and C. Lendel, "Inhibition of amyloid formation," *J. Mol. Biol.*, vol. 421, no. 4–5, pp. 441–465, 2012.

[190] M. Biancalana and S. Koide, "Molecular mechanism of Thioflavin-T binding to amyloid fibrils," *Biochim. Biophys. Acta - Proteins Proteomics*, vol. 1804, no. 7, pp. 1405–1412, 2010.

[191] O. N. Murashko and S. Lin-Chao, "Escherichia coli responds to environmental changes using enolasic degradosomes and stabilized DicF sRNA to alter cellular morphology," 2017.

[192] K. A. Obregon, C. T. Hoch, and M. V. Sukhodolets, "Sm-like protein Hfq: Composition of the native complex, modifications, and interactions," *Biochim. Biophys. Acta - Proteins Proteomics*, vol. 1854, no. 8, pp. 950–966, 2015.

[193] T. C. Jackson and M. V Sukhodolets, "Functional analyses of putative PalS (Palindromic Self-recognition) motifs in bacterial Hfq," *Biochem. Biophys. Res. Commun.*, vol. 486, no. 4, pp. 1048–1054, 2017.

[194] P. Thévenaz and M. Unser, "User-Friendly Semiautomated Assembly of Accurate Image Mosaics in Microscopy," *Microsc. Res. Tech.*, vol. 70, pp. 135–146, 2007.

[195] R. Giraldo, "Defined DNA sequences promote the assembly of a bacterial protein into

distinct amyloid nanostructures.," *Proc. Natl. Acad. Sci. USA*, vol. 104, no. 44, pp. 17388–93, 2007.

[196] F. L. Palhano, J. Lee, N. P. Grimster, and J. W. Kelly, "Toward the molecular mechanism(s) by which EGCG treatment remodels mature amyloid fibrils.," *J. Am. Chem. Soc.*, vol. 135, no. 20, pp. 7503–10, May 2013.

[197] F. Meng *et al.*, "The Sulfated Triphenyl Methane Derivative Acid Fuchsin Is a Potent Inhibitor of Amyloid Formation by Human Islet Amyloid Polypeptide and Protects against the Toxic Effects of Amyloid Formation," *J. Mol. Biol.*, vol. 400, no. 3, pp. 555–566, Jul. 2010.

[198] B. Cheng *et al.*, "Coffee components inhibit amyloid formation of human islet amyloid polypeptide in vitro: Possible link between coffee consumption and diabetes mellitus," *J. Agric. Food Chem.*, vol. 59, no. 24, pp. 13147–13155, 2011.

[199] F. Yang *et al.*, "Curcumin inhibits formation of amyloid β oligomers and fibrils, binds plaques, and reduces amyloid in vivo," *J. Biol. Chem.*, vol. 280, no. 7, pp. 5892–5901, 2005.

[200] B. Cheng *et al.*, "Silibinin inhibits the toxic aggregation of human islet amyloid polypeptide," *Biochem. Biophys. Res. Commun.*, vol. 419, no. 3, pp. 495–499, 2012.

[201] A. T. Petkova, R. D. Leapman, Z. Guo, W.-M. Yau, M. P. Mattson, and R. Tycko, "Self-Propagating, Molecular-Level Polymorphism in Alzheimer's b-Amyloid Fibrils," *Cold Spring Harb. Monogr. Ser*, vol. 408, no. 5, p. 4842, 2000.

[202] D. O. Serra, F. Mika, A. M. Richter, and R. Hengge, "The green tea polyphenol EGCG inhibits E. coli biofilm formation by impairing amyloid curli fibre assembly and downregulating the biofilm regulator CsgD via the σE-dependent sRNA RybB," *Mol. Microbiol.*, vol. 101, no. 1, pp. 136–151, 2016.

[203] A. Diz-Muñoz, D. A. Fletcher, and O. D. Weiner, "Use the force: Membrane tension as an organizer of cell shape and motility," *Trends in Cell Biology*, vol. 23, no. 2. NIH Public Access, pp. 47–53, Feb-2013.

[204] M. G. Haugh, C. M. Murphy, R. C. McKiernan, C. Altenbuchner, and F. J. O'Brien, "Crosslinking and Mechanical Properties Significantly Influence Cell Attachment, Proliferation, and Migration Within Collagen Glycosaminoglycan Scaffolds," *Tissue Eng. Part A*, vol. 17, no. 9–10, pp. 1201–1208, 2011.

[205] O. Chaudhuri and D. J. Mooney, "Stem-cell differentiation: Anchoring cell-fate cues," *Nat. Mater.*, vol. 11, no. 7, pp. 568–569, 2012.

[206] M. Lekka *et al.*, "Cancer cell detection in tissue sections using AFM," *Arch. Biochem. Biophys.*, vol. 518, no. 2, pp. 151–156, 2012.

[207] T. Ohashi, M. Hagiwara, D. L. Bader, and M. M. Knight, "Intracellular mechanics and mechanotransduction associated with chondrocyte deformation during pipette aspiration.," *Biorheology*, vol. 43, no. 3–4, pp. 201–14, 2006.

[208] T. G. Kuznetsova, M. N. Starodubtseva, N. I. Yegorenkov, S. A. Chizhik, and R. I. Zhdanov, “Atomic force microscopy probing of cell elasticity,” *Micron*, vol. 38, no. 8, pp. 824–833, 2007.

[209] E. C. Faria *et al.*, “Measurement of elastic properties of prostate cancer cells using AFM,” *Analyst*, vol. 133, no. 11, pp. 1498–1500, 2008.

[210] Y. Ding, G. K. Xu, and G. F. Wang, “On the determination of elastic moduli of cells by AFM based indentation,” *Sci. Rep.*, vol. 7, p. 45575, Apr. 2017.

[211] J. F. Berret, “Local viscoelasticity of living cells measured by rotational magnetic spectroscopy,” *Nat. Commun.*, vol. 7, no. 10134, Jan. 2016.

[212] L. Chevry, N. K. Sampathkumar, A. Cebers, and J. F. Berret, “Magnetic wire-based sensors for the microrheology of complex fluids,” *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.*, vol. 88, no. 6, pp. 1–8, 2013.

[213] F. Loosli *et al.*, “Wire-Active Microrheology to Differentiate Viscoelastic Liquids from Soft Solids,” *ChemPhysChem*, vol. 17, no. 24, pp. 4134–4143, 2016.

Annexe A : Article 1

Techniques to analyze sRNA protein cofactor self-assembly in vitro

Chapitre de livre de la série « *Methods in Molecular Biology* » des éditions *Springer* en 2018 (DOI: 10.1007/978-1-4939-7634-8_187634-8_18)

Chapter 18

Techniques to Analyze sRNA Protein Cofactor Self-Assembly In Vitro

David Partouche, Antoine Malabirade, Thomas Bizien, Marisela Velez, Sylvain Trépout, Sergio Marco, Valeria Militello, Christophe Sandt, Frank Wien, and Véronique Arluison

Abstract

Post-transcriptional control of gene expression by small regulatory noncoding RNA (sRNA) needs protein accomplices to occur. Past research mainly focused on the RNA chaperone Hfq as cofactor. Nevertheless, recent studies indicated that other proteins might be involved in sRNA-based regulations. As some of these proteins have been shown to self-assemble, we describe in this chapter protocols to analyze the nano-assemblies formed. Precisely, we focus our analysis on *Escherichia coli* Hfq as a model, but the protocols presented here can be applied to analyze any polymer of proteins. This chapter thus provides a guideline to develop commonly used approaches to detect prokaryotic protein self-assembly, with a special focus on the detection of amyloidogenic polymers.

Key words Protein self-assembly, Noncoding RNA cofactor, Functional amyloid

Abbreviations

AFM	Atomic force microscopy
CF	Curve fitting
FTIR	Fourier transform infrared spectroscopy
SAXS	Small angle X-ray scattering
SRCD	Synchrotron radiation circular dichroism
TEM	Transmission electron microscopy
ThT	Thioflavin T

1 Introduction

In prokaryotes, the components of RNA metabolism self-assemble into complex structures, resulting in functional compartmentalization within the cell [1–3]. The RNA maturation components

Véronique Arluison and Claudio Valverde (eds.), *Bacterial Regulatory RNA: Methods and Protocols*, Methods in Molecular Biology, vol. 1737, https://doi.org/10.1007/978-1-4939-7634-8_18, © Springer Science+Business Media, LLC 2018

include proteins involved in post-transcriptional genetic regulation such as ribonucleases (RNAses) or RNA chaperone, but also small regulatory noncoding RNAs (sRNA) [2, 4, 5]. The best-characterized protein involved in such pathways using regulatory sRNA is probably the RNA chaperone Hfq [6]. Nevertheless, some others have also been identified such as ProQ and the archaeal Sm proteins [7, 8]. Our recent studies demonstrated that specific sequences in the *E. coli* Hfq C-terminus region (CTR) are able to polymerize, resulting in the formation of amyloid fibers [9]. Indeed, increasing evidence shows that amyloids are not only the result of protein misfolding associated with neurodegenerative diseases, but also found in cells for useful reasons [10]. In the later case, amyloids contribute to the physiology of the cell and are referred as *"functional"* [10]. Indeed, these functional amyloids have been widely reported in different organisms, including bacteria. In prokaryotes, diverse functions have been described for these amyloids, such as a role in biofilm development [11], formation of curli and pili [12], or creation of pores in the host membrane [13]. In this chapter we intend to provide a guideline to develop approaches that can be used to detect the fibers of sRNA protein cofactors, precisely molecular imaging, such as atomic force microscopy and transmission electron microscopy, infrared spectroscopy, synchrotron radiation circular dichroism, and small angle X-ray scattering. Our chapter will focus on Hfq and its amyloid nature as an example, but it could be applied to any self-assembling prokaryotic proteins.

2 Materials

2.1 Preparation of Samples

1. Synthetic peptide: 20 mg/mL in deionized RNA-grade water (*see* **Note 1**).
2. Purified Hfq/Sm proteins, prepared as described in the Chapter 15 by Stanek and Mura. Depending on buffer used and protein concentration reached during the purification process, the same limits apply as those described in **Note 1**.
3. UV spectrophotometer suitable for the precise determination of protein concentration measurements (micro-volumes).

2.2 Thioflavin Staining

1. ThT stock solution: 0.8 mg/mL in 10 mM phosphate buffer pH 7 containing 150 mM NaCl.
2. Fluorescence microscope equipped with a digital camera and emission/excitation filter sets to permit visualization of ThT staining.

2.3 AFM Imaging

1. AFM microscope with liquid cell.
2. AFM tips with different force constants: 2–75 N/m and resonant frequencies in the range of 100–320 kHz for imaging in

air or in the range of 0.01–0.8 N/m and resonant frequencies in the range of 70–100 kHz for imaging in solution.

3. AFM-grade mica.
4. AFM-free image analysis software such as WsXM (http://www.wsxmsolutions.com/) and Gwyddion (http://gwyddion.net/).

2.4 TEM Imaging

1. Cryo-transmission electron microscope equipped with a slow scan CCD camera.
2. Carbon-coated electron microscopy copper grid, CF200-Cu (200 mesh) or CF300-Cu (300 mesh), EMS.
3. Carbon-coated electron microscopy copper grid, ref. 01881-F (200 mesh) or 01883-F (300 mesh), Ted Pella.
4. Uranyl acetate solution: 2% (w/v).
5. Uranyl-less solution containing Gadolinium salts, Delta Microscopies.
6. Plasma cleaner.
7. Leica EM-CPC plunge-freezing device for cryo-fixation.

2.5 FTIR Spectroscopy

1. Interferometer-based spectrophotometer coupled with an IR source.
2. For transmission measurements, liquid cell equipped with CaF_2 windows and a 6 μm-thick polytetrafluoroethylene (Teflon) spacer (for example, Omni-Cell transmission Cell from Specac Company or Demountable liquid Cell from Piketech).
3. For reflection measurements, attenuated total reflection (ATR) sampling setup.

2.6 SRCD Spectroscopy

1. SRCD endstation such as that of DISCO beamline at SOLEIL synchrotron [14, 15].
2. Manually loaded circular demountable CaF_2 cells of 30 μm path length with a loading volume of 2 μL [16].
3. CDtool software for data acquisition and treatment, including averaging, smoothing, subtraction of buffer baselines from sample-spectra, calibration ((+)–camphor-10-sulfonic acid CSA), and normalization [17].
4. BestSel open access software (http://bestsel.elte.hu/) [18].

2.7 Small Angle X-Ray Scattering (SAXS)

1. SAXS beamline such as SWING at SOLEIL.
2. Quartz capillaries of 1.5 mm diameter and 0.01 mm thickness.
3. Foxtrot software (http://www.synchrotron-soleil.fr/Recherche/LignesLumiere/SWING).

3 Techniques

3.1 Analysis of Hfq Self-Assembly by Molecular Imaging

Characterization of Hfq self-assembly can be performed by microscopy techniques, including fluorescence microscopy, transmission electron microscopy (TEM), and atomic force microscopy (AFM).

3.1.1 Thioflavin Staining

Thioflavin T (ThT) is a benzothiazole salt commonly used to visualize amyloids. When bound to aggregated β-sheets, such as those found in amyloids, the dye displays an enhanced red-shifted fluorescence emission spectrum [19]. Even if the dye fluorescence is not completely specific and may bind to other structures (such as double-stranded DNA), it is however a good indicator that amyloid structures are formed in the sample. Practically, for ThT staining:

1. Dilute protein samples to 0.2 mg/mL.
2. Prepare a stock solution of ThT. This solution can be filtered and stored in the dark for 1 week. Before use, the stock solution is diluted in 10 mM phosphate buffer pH 7 containing 150 mM NaCl (dilution 1/50 = working solution); this must be done extemporaneously.
3. For a spectroscopic assay, add a 10 μL aliquot of protein to 1 mL of working solution, incubate for 1 min, and measure the fluorescence intensity with excitation set at 440 nm and emission detected in the range 450–500 nm (*see* **Note 2**).
4. For slide staining and visualization under a microscope, place 10 μL aliquots of stained samples and image using a fluorescence microscope with a 60× oil-objective (NA = 1.4). The laser excitation can be set at 415 or 450 nm and emission is detected at 482 nm.

3.1.2 AFM Imaging

Atomic force microscopy (AFM) has become a common tool for the analysis of protein aggregation and fibrilization [20]. It provides height information that complements structural information from two-dimensional projection images obtained using electron microscopy (*see* Subheading 3.1.3). In essence, an AFM setup consists of a microtip attached to a cantilever held on a piezo scanner that moves along the surface of the sample with high precision over a defined distance within the micrometer range. The deflection of the cantilever follows the surface topography and is monitored following the position of a laser spot reflecting from its surface on a photodiode. The motion of the cantilever can also be used to interrogate the mechanical properties of the sample. AFM imaging thus provides information about biological objects dimensions at nanometer resolution [21]. As amyloid fibrils are usually in the μm range in length with a lateral dimension between 3 and 30 nm, AFM is thus perfectly adapted to probe fibrils structure.

1. Substrate preparation. In order to study fibrillar protein structures, it is essential to first adsorb them to a flat surface.

This step is crucial for obtaining high-quality data. The fibers formed in solution have to be at an adequate density and firmly attached to allow the tip to scan over them. One of the surfaces that is most frequently used for AFM studies is mica, although clean glass or annealed gold can also be used. Here we will concentrate on describing the adsorption of fibrils to mica surfaces. Mica is convenient because it is atomically flat and can be easily cleaved immediately before use to obtain a clean surface. AFM-grade mica can be easily obtained from several AFM product suppliers. The freshly cleaved mica is glued to the sample holder using, for example, double-sided tape before incubating the protein sample.

2. Protein adsorption. A protein solution is incubated on the mica surface to allow for fibril adsorption to the surface (*see* **Note 3**). The size, charge, and mechanical stability of the fibers determine the details of the preparation protocol. Basically this protocol consists of two steps: first, conditioning the surface where the fibers will be deposited and second, incubating the protein solution to allow for filament deposition. As the surface/protein interaction will be affected by the surface charge of the protein and surface, the pH of the solution has to be selected to optimize this interaction. Precisely, mica is constituted of sheets of octahedral hydroxyl-aluminum lying between two silicon tetrahedral layers. The most favorable cleavage plane is located in the layer of unbound K^+ that lies in the interlayer space between neighboring silicon layers. When immersed in water, the hydrated potassium ions can be dissociated from the mica surface, leaving negatively charged aluminum tetrahedra (AlO_2^-) in the first outer layer [22]. Thus, the solution pH and ionic salt concentration should be adjusted to allow for the protein to be positively charged to promote its interaction with the negatively charged mica surface.
3. Imaging in air. If the protein binding to the surface is not strong enough to allow imaging with the sample immersed in solution, the best alternative is to image in air. Incubate a small amount of diluted protein (typically 10 μL of 5–10 μg/mL fibrils) on the mica surface and allow to air dry at room temperature. All the fibers in the solution will now be adsorbed on the mica surface. Before imaging, the sample must be extensively washed with Milli-Q water to remove the salt from the buffer and then carefully air-dried again. AFM can operate in different modes, and the ones more conveniently used to image amyloid self-assemblies are contact mode and tapping mode, depending on the characteristics of the fibrils of interest [21]. Contact mode images are obtained when the tip is in contact with the sample, so it requires that the fibers are firmly attached to the surface and rigid enough to withstand the tip

pressure as it scans along the surface without deformation or sliding. In tapping mode, the cantilever is oscillated at its resonant frequency during the scan, so that it does not drag the sample along as it moves over the surface. Resonant frequencies of the soft cantilevers used for biological materials are, typically, within the 70–100 KHz range, depending on whether the microscope is operated in air or in solution, so that the interaction time between the tip and sample is reduced dramatically to only a few microseconds. Fibers can be imaged either in tapping or in contact mode, depending on the characteristics of the sample.

4. Imaging in solution. If the fibers are sufficiently long and their interaction is strong enough to allow for a stable adsorption to the surface, the fibers will remain firmly attached during imaging in solution. After allowing 10–60 min of incubation, the sample is extensively rinsed with buffer to remove excess protein. The imaging buffer is selected to optimize the interaction between the tip and the sample and could differ from the one selected for fiber adsorption to the mica. The pH or salt concentration can be adjusted again so that the tip–protein interaction is minimized to reduce damaging the fibers during imaging. Operating in solution offers the possibility of imaging fiber growth. In such a case, imaging should be done in the presence of small amounts of protein in the solution to avoid interfering with the imaging process. Images of the same area taken at different times can be used to follow filament growth kinetics.
5. AFM imaging procedure. Once the sample is properly prepared, the AFM is operated selecting the most appropriate tip for the imaging mode. If imaging in air, stiffer tips with force constants in the range of 2–75 N/m and resonant frequencies in the range of 100–320 kHz can be more convenient to reduce the snap-to-contact region. For imaging in solution, softer tips, with force constants in the range of 0.01–0.8 N/m and resonant frequencies in the range of 70–100kHz can be used. In brief, AFM operation includes the following steps: placing the cantilever in the holder and then both to the scanner; aligning the laser to impinge the cantilever tip; aligning the laser reflection on the center of the photodiode; placing the sample in the sample holder; if operating in tapping mode, the cantilever resonance frequency has to be selected; approaching the tip and sample. Once the tip is touching the sample, the scanning parameters, gains, scanning speed, and setpoint, are adjusted to obtain reproducible trace and retrace images.
6. AFM topographic analysis. AFM obtains direct topographical images of the sample surface. Each AFM provider includes software to analyze the images, but there are also free specialized

software available to process scanning probe microscopy images (WsXM or Gwyddion).

Figure 1a–f shows Hfq peptide fibers imaged in air or in solution.

3.1.3 TEM Imaging

A quick way to check the presence and structural characteristics of fibers (length, diameter, interweaving), as well as the homogeneity of the sample, is negative staining and observation by transmission electron microscopy.

Protocol for Negative Staining

1. Sample preparation. Deposit a 5 μL drop of Hfq sample at 1 mg/mL in 50 mM phosphate buffer pH 7 (i.e., stock solution diluted in water) on a glow-discharged carbon-coated electron microscopy copper grid (200 mesh square grid, EMS). In order to optimize the adsorption of the sample to the grid it is recommended, but not strictly required, to realize a glow-discharge (gas ionization process giving the carbon film hydrophilic).

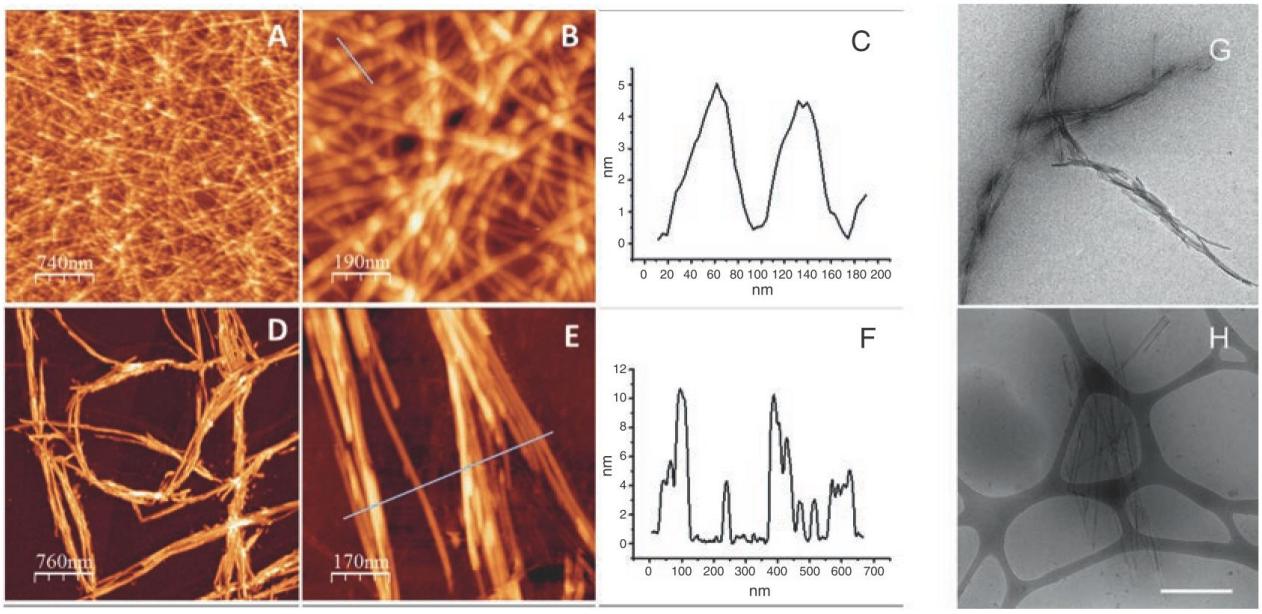


Fig. 1 Visualization of Hfq self-assembly by molecular microscopy. Transmission electron microscopy (TEM) and atomic force microscopy (AFM) are efficient tools to investigate protein nanostructures. (**a**–**f**) AFM images of Hfq peptide fibers in air and under liquid. Upper panels (**a**, **b**, **c**) are fibers imaged in air. A drop of the protein solution was allowed to dry on the mica before imaging in contact mode. Lower panels (**c**, **d**, **f**) show filaments observed under buffer at pH 4 in tapping mode, after incubating the fibers for 5 min on mica at pH 4 at a concentration of 0.5 mg/mL. Panels C and F show the height profile under the blue lines shown, respectively, in panels **b** and **e**. (**g**–**h**) In TEM experiments, 5 μL of Hfq solution is deposited on a glow-discharged electron microscopy grid covered by a continuous carbon film for negative staining (**g**) or a holey carbon film for cryo-TEM (**f**). (**g**) The sample is dried and subsequently stained with a solution of either uranyl acetate or uranyl-less (filling the protein cavities with heavy atoms). (**h**) The excess of sample is blotted and subsequently frozen in ethane at liquid nitrogen temperature (embedding the sample in thin amorphous ice layers). Both methods show filaments that seem more aggregated in negative staining. *Scale bar:* 500 nm

2. Sample adsorption. After 5 min of interaction, blot out the excess sample using a Whatman filter paper (1–3 min is usually enough to have efficient adsorption of the sample onto the grid).
3. Contrasting of the sample. The adsorbed sample can then be put in contact with the contrasting agent. To perform negative staining, 5 µL of uranyl acetate solution (2%) is applied onto the grid-containing sample. After 30 s incubation, the excess of uranyl acetate is blotted out and then the grids can be kept in a dry dark dust-free environment until observation at the electron microscope. Gadolinium salt (uranyl-less) is also used for negative staining. It is used instead of uranyl acetate because it is safer. However, it may create artifacts such as crystals. Store it at cold temperature and use it at room temperature to avoid salt deposition as much as possible.
4. The grids can be stored at room temperature for further observation on any standard transmission electron microscope.
5. Sample observation. The electron microscopy grid is then mounted onto a room temperature-equilibrated holder and subsequently introduced into an electron microscope (JEOL 2200FS, JEOL, Tokyo, Japan).
6. Image acquisition. 2 k by 2 k images are acquired using a Gatan Ultrascan 894 US1000 slow scan CCD Camera (Gatan, Pleasanton, CA, USA) at 40,000× nominal magnification (corresponding pixel size was 0.32 nm).

Figure 1g shows Hfq peptide fibers imaged by TEM after negative staining.

While it is true that negative staining is the tool of choice for quick characterization, it implies the dehydration of the sample, which can lead to artifactual aggregations of fibers or to structural modifications. Therefore, when high-resolution ultrastructural studies are required, samples should be observed in hydrated conditions, which is possible by cryo-transmission electron microscopy (cryo-TEM). Cryo-TEM is based on fast-freezing of the electron microscopy grid containing a thin layer (100–200 nm or less) of the adsorbed sample by quick immersion in liquid ethane cooled-down at liquid nitrogen temperature. This can be achieved by using a lab-made or commercial apparatus. Some commercial apparatuses can be fully automatized so that repetitiveness can be easily achieved when freezing conditions have been set up.

Protocol for Cryo-Electron Transmission Microscopy (Cryo-TEM)

1. Sample preparation. Deposit a 5 µL drop of Hfq suspension on a glow-discharged lacey carbon film grid (300 mesh).
2. Sample freezing. Blot the suspension through the grid for about 1 s using a Whatman filter paper to create thin suspended films in the carbon holes of the grid. Using a Leica EM-CPC, plunge

the grid into liquid ethane cooled-down with liquid nitrogen. Prior to sample freezing, the grid is kept in a liquid nitrogen storage unit until observation at the electron microscope.

3. Transfer to the electron microscope. The grid with the frozen hydrated sample is transferred to a cryo-TEM for observation. During this transfer process, and during observation, the sample should remain under liquid nitrogen to prevent water crystallization, which impedes the observation of the sample. Since the observed sample is fully hydrated, the obtained images reflect the near native structure of proteins and fibers. A simple but detailed description of negative staining and cryo-TEM techniques is described in [23]. Grids are transferred into a JEOL JEM 2200FS cryo-electron microscope equipped with a Ω-energy filter by using a Gatan 914 cryo-transfer system.
4. Image acquisition. Hfq peptide cryo-images can be acquired using a 2 k by 2 k Gatan Ultrascan 894 US1000 slow scan CCD Camera (Gatan, Pleasanton, CA, USA) at 40,000× nominal magnification (corresponding pixel size was 0.32 nm) with an energy window of 20 eV.
Figure 1h shows Hfq peptide fibers imaged by cryoTEM.

3.2 Analysis of Hfq Amyloid-like Self-Assembly Secondary and Super-Secondary Structure

3.2.1 Infrared (FTIR) Spectroscopy

Fourier transform infrared (FTIR) spectroscopy can be used to analyze protein secondary and super-secondary structures such as amyloid. The peptide bond is an amide group (CONH) that gives rise to seven active absorption bands in the infrared. This method can thus give information on secondary structure (three bands carry information on the secondary structure, but the Amide I band is the most used, see below) and tertiary structure (through the Amide II band). Indeed, the Amide II spectral region studied in the presence of D_2O solution (i.e., Amide II′, *see* **Note 4**) is a marker of the partial opening of the protein and thus of the tertiary structure changes, during which the hydrogens remained within the protein core undergo H-D exchange [24]. The Amide I is most often used for secondary structure elucidation since it has the smallest contribution from peptide side chain absorptions. The Amide I band absorption is mainly due to the C=O stretching of the amide group vibration but this vibration is strongly delocalized over the amide C–N and C–H bonds giving rise to a strong coupling: the trans dipole coupling. The exact frequency of the absorption peak is thus sensitive to the angle of the amide bond, which is dependent on the local structure of the protein domain. α-Helices, β-sheets, β-turns, random coils, and 3–10 helices are each restricted to a different set of angles as seen in a Ramachandran plot, giving rise to amide peaks located at different frequencies. This gives rise to the sensitivity of infrared spectroscopy to the secondary structure

of proteins. Here, we thus focus on the infrared zone in the region of the Amide I.

1. Hfq fibril samples (from 1 to 20 mg/mL) are loaded into liquid cell equipped with CaF_2 windows and a 6 μm thick spacer. This configuration allows studying a sample in solution for several hours (8 on average), and therefore allows for following the aggregation kinetics.
2. FTIR absorption measurements are performed between 1000 and 4000 cm^{-1} through the use of an interferometer-based spectrophotometer coupled with an IR source and a Triglycine sulfate (DTGS) detector. The best way to have good transmission spectra is by performing a measurement in which 256 scans are collected and at least 5 independent measurements are averaged. Spectra at a nominal resolution of 2 cm^{-1} are baseline corrected and background subtracted before analysis. A spectrum of the buffer must be collected using the same experimental parameters for water subtraction. Buffer subtraction is performed by an iterative procedure to remove the signal originating from the water. The association band of water between 1800 and 2400 cm^{-1} is used to compute the subtraction factor. Most commercial software packages propose an integrated subtraction routine.
3. Alternatively, an attenuated total reflection (ATR) sampling setup can be used. With an ATR setup, the infrared beam light is reflected inside an infrared-transparent, high refractive index crystal. The crystal can be single bounce or have a multiple bounce geometry to increase sensitivity. An evanescent wave is created at the crystal–solution interface, which extends into the sample with a penetration depth typically around 1 μm (depending on the wavelength and on the crystal's refractive index). A background spectrum without sample is first recorded. A total of 64 scans are usually averaged. Then, ~ 5–20 μL of fibril sample (1–20 mg/mL) is deposited directly on the crystal and covered with a cap to reduce evaporation. A spectrum of the buffer must also be recorded in the same conditions for latter subtraction.
4. The analysis consists of qualitative evaluation of secondary structure through the combination of 2nd-derivative, Fourier self-deconvolution (FSD), and curve-fitting (CF) analysis [25]. FSD and the 2nd-derivative can be used to find the number and exact positions of the overlapping peaks composing the Amide I band. CF involves fitting a series of model peaks to the experimental data. The best function for peak shape in liquid samples is usually a convolution of Lorentzian and Gaussian bands, the Voigt profile. Overlapping peaks composing the Amide I band are fitted with Voigt profiles with a bandwidth

varying between 15 and 25 cm^{-1}. The goodness-of-fit is generally evaluated by comparing the rest with the noise. The CF, being an iterative algorithm, is pursued until the rest is less than 5 times the noise. The integration of the area of the component bands used to generate the final curve fit can provide an estimate of the relative percentages of secondary structure present. Both qualitative and quantitative analyses can be achieved with dedicated software. Most software packages such as OMNIC™ (Thermo Fisher Scientific) and Opus (Bruker) provide curve-fitting algorithms.

5. The assignment of a peak at a given wavenumber to a given secondary structure should be done according to Table 1. Briefly, a peak near 1645 cm^{-1} is indicative of random coil, 1655 cm^{-1} of α-helix, and 1630 cm^{-1} and 1690 cm^{-1} of β-sheet (*see* **Note 5**). Stronger hydrogen bonding results in a shift to lower wavenumbers and, therefore, amyloid fibrils often have β-sheet peaks below 1630 cm^{-1} [26].

Figure 2a shows Hfq peptide fibers analyzed by FTIR spectroscopy.

3.2.2 Synchrotron Radiation Circular Dichroism SRCD

Circular Dichroism (CD) is a sensitive absorption spectroscopy technique for studying biological samples such as proteins and polynucleotides (DNA and RNA) as well as sugars. Circular left and right polarized light is differentially absorbed, e.g., by proteins due to the excitation of the n-π and π-π* electronic transition of the peptide bonds. Additionally the aromatic side chains produce characteristic absorptions between 210 and 275 nm. Standardized CD spectra can be deconvoluted and used for protein secondary

Table 1
Empirical assignments for Amide I infrared (IR) bands characteristic of proteins [25]

Wavenumber (cm^{-1})	Assignment
1615–1625	Intermolecular β-sheet (in particular vibrations of strongly bound intermolecular aggregated β-strands, including amyloids)
1630–1640	Intramolecular native β-sheet
1640–1650	Disordered random coil
1650–1660	α-Helix
1660–1695	β-turn
1675–1695	Intermolecular β-sheet (in particular antiparallel aggregated β-sheets)

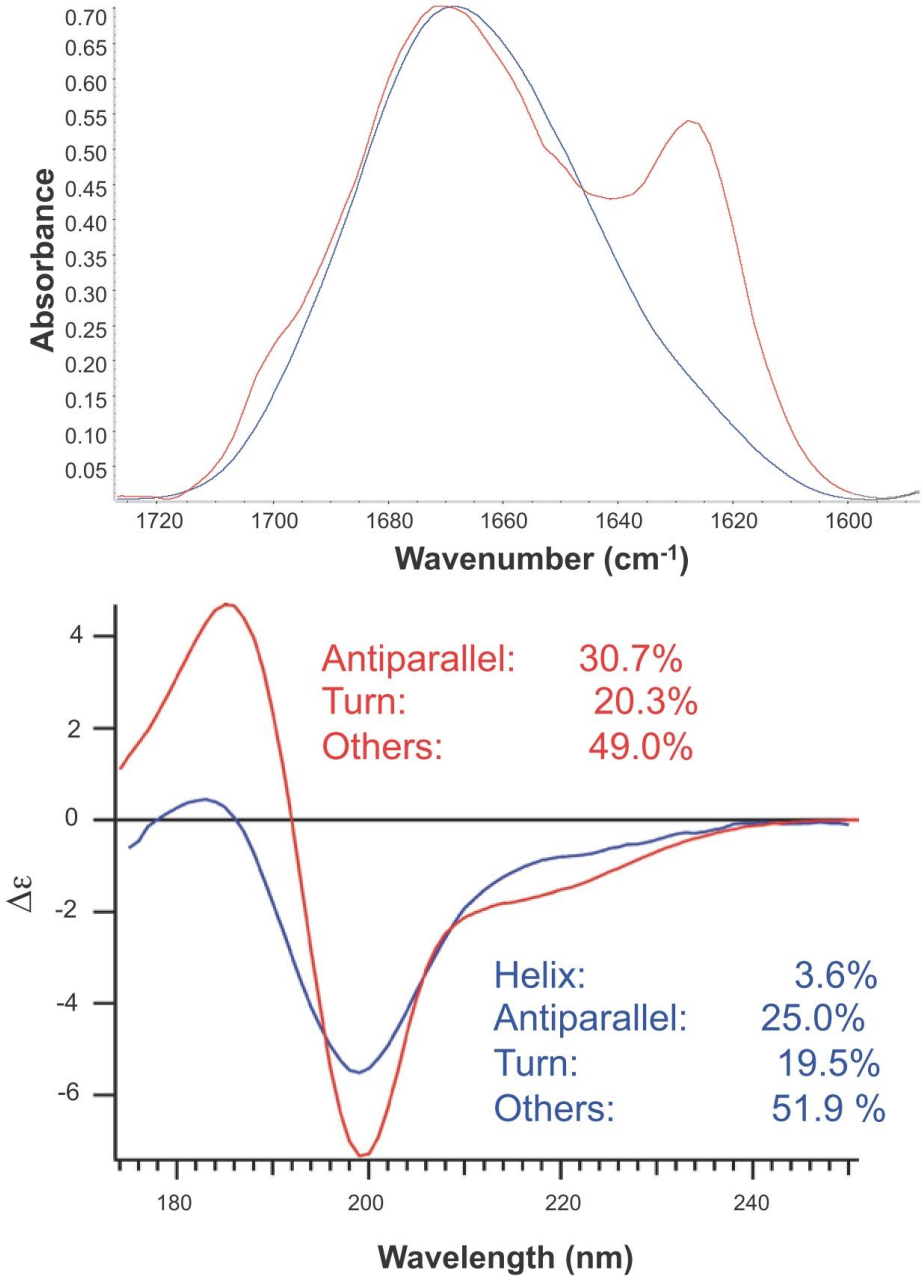


Fig. 2 Secondary and supersecondary structure determination. (**a**) FTIR spectra of Hfq peptide immediately after diluting in water (blue) and after incubation for 1 day at 20 °C (red). The peaks shown in (**a**) are due to intermolecular aggregated β-sheet at 1620 cm^{-1} and to the presence of native β-turn at 1670 cm^{-1} and intermolecular aggregated β-sheet at 1690 cm^{-1}, respectively. In addition, in a lot of proteins studied, the contemporary increase of the components at about 1620 and 1680 cm^{-1} indicates the building of antiparallel β-aggregated structures; the low frequency β-sheet moiety at approximately 1620 cm^{-1} are attributed to multistranded intermolecular β-sheet and resembles the typical fingerprint of the cross-β motif found in amyloid-like fibrils [37, 38]. (**b**) SRCD spectra of Hfq peptide at 20 mg/mL in water (blue) or after incubation at 20 °C for a few days at 100 mg/mL to ensure the formation of fibrils (red). In this case, sonication procedure to fragment fibrils into shorter forms or their monomeric form was not necessary. Secondary structure determination by BESTSEL (inset) revealed a structural increase of 5% in β-sheet content for the fibrils as opposed to the monomeric Hfq peptide

structure determination. Synchrotron radiation circular dichroism (SRCD) allows for the extension of the spectral range down to 168 nm in aqueous solution and 120 nm in hydrated films, with very good signal-to-noise ratios (typically ~20 mdeg/0.2 mdeg). The extension of the spectral range down to 168 nm has improved the information content obtainable from a SRCD spectrum. Especially for the weak CD signals of proteins with high β-sheet or disordered content, SRCD has improved the spectral analysis substantially [27]. Numerous algorithms exist for the estimation of the secondary structure composition from CD spectra. For the special case of β-sheet-rich proteins, with a broad range of protein folds, a recent publicly accessible algorithm BestSel [18] allows the determination of secondary structure contents including parallel and antiparallel β-sheets and potential fold recognition. The novel reference dataset with spectra that significantly differ from present reference sets extends the information content for secondary structure determination.

1. Sample preparation. Homogenized Hfq peptide fibrils were prepared by diluted peptide in water at 20 mg/mL to 100 mg/mL. Observation of fibrils can be made directly after few days (*see* **Note 1**) as described in **step 5**. Nevertheless, in some cases long and associated fibrils need to be fragmented to shorter and individual pieces to be analyzed, as described below (2–4).
2. Seeds can be prepared by a sonication procedure [28]. The sonication strength and time had to be carefully chosen. At low force, homogenization of the fibril solution will just decrease the viscosity by breaking the fibrils into shorter pieces. A microtip sonicator set up in the cold room needs to be regulated to minimal strength (force, amplitude), with a 50% duty cycle (e.g., 1 s pulses interrupted by 1 s silence). Sample volume should be at least 100 μL in microtubes. Sonication shall be carried out with the tubes containing sample cooled in ice-cold solution (for better contact), with the tip plunged right to the bottom of the tube without touching the tube wall. For homogenization of peptides and small proteins four cycles of five pulses with a 10 s rest between cycles produces reliable and reproducible results. During the rest period, mix well the solution containing the peptide by lifting and plunging the tip up and down. Droplet formation on the tube wall should be avoided. Gloves, mouth mask, and glasses are recommended for safety.
3. SRCD spectra are taken before and after sonication allowing the spectral distinction between fibril seeds and amyloid fibrils.
4. Ultracentrifugation in Beckmann TLA-100 rotor run at 53500-96500 Relative Centrifugal Field (RCF) × *g* for 15–30 min may be used to distinguish eventual monomers floating in the supernatant and the fibrils accumulating in the

pellet. The baseline spectrum for the fibril seeds should be taken from the supernatant. Once the sonication and centrifugation are finished, keep samples at 4 °C.

5. Sample concentration, buffer composition, and pathlength choices for SRCD data acquisition should be handled with care. In general, concentrations should be for α-helix rich proteins 1–2 g/L, for β-sheet rich proteins 3–4 g/L, and for unordered structures above 5 g/L in a 20 μm pathlength cell. Loading volumes are 2 microliters in special 20 μm CaF_2 cells [16] (Fig. 3). The choice of buffer is a function of how much chloride will be considered essential (10–50 mM NaCl allows obtaining spectra down to 185 nm). Chloride is readily replaced by fluoride allowing twice the initial chloride concentration. Ideally phosphate buffer is used being more temperature stable than Tris-Cl and being deep-UV transparent.

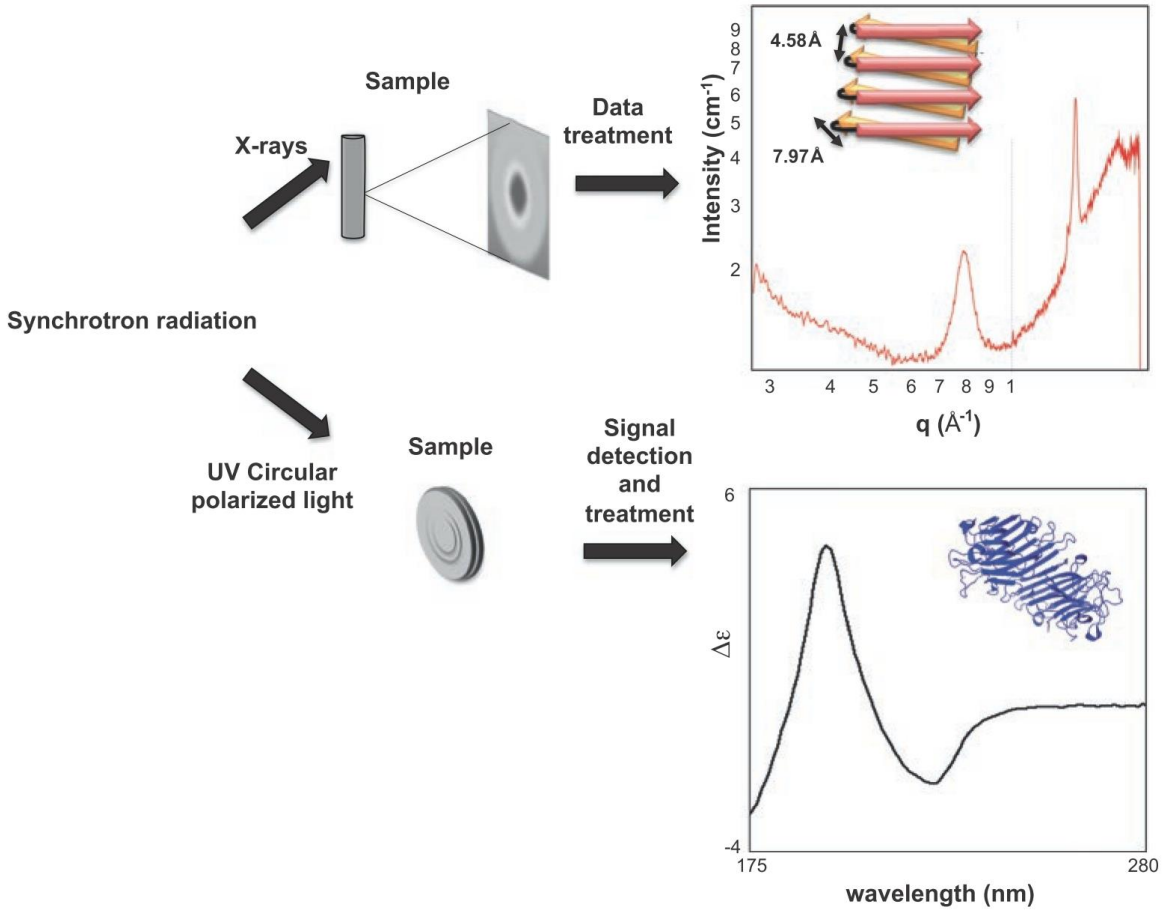


Fig. 3 Use of synchrotron radiation for the analysis of amyloids self-assemblies. Top: Basic sketch of data acquisition and treatment at a SAXS beamline. The sample scattering pattern recorded by a 2D detector is radially averaged to and background is subtracted to give the following SAXS curve. The CTR peptide of the Hfq protein exhibits typical diffraction peaks corresponding to a cross-β structure sketch in overlay. Bottom: Data acquisition and treatment at a SRCD beamline. A rich **β**-strand containing protein is shown as an example (Concanavalin A, PDB ID 4PF5 [39])

6. All spectra are normalized to the mean residue weight ellipticity (Θ_{MRW}, deg.cm^2.dmole^{-1}).
7. For secondary structure determination with BestSel, normalized root-mean-square deviation (NRMSD) indicates the most accurate fit for each spectrum; values of <0.15 are considered significant.
   Figure 2b shows Hfq peptide analyzed by SRCD spectroscopy.

3.3 Characterization of Cross-β Structure

Small angle X-ray scattering (SAXS) is a simple and powerful technique that directly provides structural information at scales ranging typically from 1 nm to 1 μm [29–31]. The physical principle is the same as X-ray crystallography; it is based on elastic scattering of an incoming X-ray wave by the electrons of the sample. The resulting "signal" is a scattering pattern most frequently collected using a bi-dimensional detector positioned behind the sample and perpendicular to the incoming beam. Practically, SAXS experiment can bring information for either peptides in solutions or assembled materials made of peptides. In this section, we will focus on how to prepare and analyze amyloid peptides and particularly Hfq cross-β structure.

3.3.1 Sample Preparation

1. Different solutions of peptide are prepared (25–100 μL), ranging from 0.5% w/v to the highest concentration by 2.5% steps by dissolving the peptide into an appropriate buffer. The buffer must be the same for the previously described experiments (FTIR, SRCD…) in order to compare the results; buffer containing phosphorous should be avoided and low salt concentration is better.
2. When all the solutions are ready, directly fill quartz capillaries of 1.5 mm diameter and 0.01 mm thickness. Then, capillaries are sealed using wax or nail polish. The capillaries containing the peptides are left at room temperature until cross-β strands are obtained (this is usually determined by other methods).
3. If the cross-β strand formation kinetics are unknown, FTIR will be a better method than SAXS to determine the kinetics of assembly.

3.3.2 SAXS Setup and Recording

The SAXS experiment can bring information for both peptides in solutions or assembled materials made of peptides. The intensity collected is proportional to the number of objects within the sample, *N*. It directly results from their size, shape, and internal structure, given by the Fourier Transformation modulus of their electronic density, showed here for an isotropic sample:

$$I(q) = r_e^2 \sum_{p=1}^{N} \left\langle \left| \iiint_{V_p} \left(\rho_e(\vec{r}) - \rho_{0)} \right) e^{-i\vec{q}\cdot\vec{r}} d^3r \right|^2 \right\rangle \Omega$$

where the triple integral is performed within the volume V_p of each particle and the average performed over all possible orientations Ω. r_e is the classical electron radius, $\rho_e(\vec{r})$ is the electronic density at position $\vec{r}$ within the particle, ρ_0 is the average electronic density of the surrounding buffer, $q = \frac{4\pi \sin\theta}{\lambda}$ is the momentum transfer, θ is the scattering angle, typically below a few degrees, and λ is the X-ray wavelength. The rotational average severely limits the structural information content of the scattering pattern for a non-assembled material obtained in a SAXS experiment. When isolated macromolecules assemble into regular structures, new correlations appear in the electronic density of the whole structure, giving rise to specific modulations in the scattering intensity. When motifs are repeated along a defined axis with a repetition distance d, the scattered photon waves undergo strong constructive interferences at q values given by $q = n\frac{\pi}{d}$, where n is an integer (Bragg law). The phenomenon is called diffraction, but is not different in nature from solution scattering. When identical units, e.g., a peptide or a protein, are regularly spaced, with periodicities typically on the order of or larger than 1 nm, diffraction is observed in the Small angle range (SAXD) [32].

1. Cross-β structure is investigated here. The sample-to-detector distance needs to be optimized. The typical signals resulting from β-sheet structure correspond to inter-distances of around 10 and 4.7 Å; the available q-range should thus be from 0.3 to 2 $Å^{-1}$.
2. Capillaries containing the samples and buffer (control) are mounted onto capillary holder system and positioned into the beam path.
3. Depending on the sample homogeneity, several positions within the sample capillaries can be chosen. At least 3 positions are enough: one at the top of the sample, one in the middle, and the last at the bottom.
4. One important step here is the detection of radiation damage. When exposed to X-rays, organic molecules tend to form free radicals, leading to aggregation of the sample onto the capillary wall. Radiation damage does not occur for all samples. This phenomenon is observable after data treatment by an increase of the curve intensity at small q. In order to avoid this problem, samples should be exposed to X-rays and then let to rest for a few seconds before a new exposition at the same spot.
5. As samples on the same capillary holder can vary greatly, the exposure time must be adjusted, in order to get enough signal and to avoid saturation on the detector (depending of the beamline detector).

6. Due to radiation damage, acquisition time should not be higher than 3 s with at least 5 s of pause. It is better to make 10 acquisitions of 3 s than 1 of 30 s.
7. Record the scattering pattern with short and identifiable sample and buffer names.

3.3.3 Data Treatment

1. When the entire scattering patterns are recorded as 2D images, radial integration is performed in order to obtain SAXS curves for each position inside all the samples and buffers. Such operation can be done using the Foxtrot software or equivalent.
2. For curves corresponding to the same position within the same sample, averaging is suggested in order to obtain better statistics. This must be done carefully in order to avoid averaging data with radiation damage.
3. As samples are analyzed within the buffer, the buffer signal needs to be subtracted. Data subtraction is provided by most of the SAXS data reduction software.
4. When the resulting final curve (averaging + subtraction of buffer) is obtained, diffraction peaks corresponding to cross-β sheets should be observed. The example of Hfq C-terminus peptide is shown on Fig. 3 (top), where the anisotropic reflections are indicative of a partially aligned fiber, perpendicular to the X-ray beam. The reflections at d = 4.58 Å and d = 7.97 Å correspond, respectively, to the inter-strand and inter-sheet spacing.

4 Concluding Remarks

The self-assembly of biological macromolecules constitutes a key process in all living organisms. We recently reported that sRNA cofactor Hfq belongs to the family of functional amyloids [9]. The various experimental approaches described in this chapter aim at further investigating the self-assembling properties of sRNA cofactors. This could indeed represent a versatile means to regulate sRNA-related processes in vivo. Furthermore, this would also enable the formation of mixed synthetic sRNA:protein synthetic self-assemblies, with future perspectives for nanotechnologies.

5 Notes

1. The main obstacle for the examination of peptides and proteins that polymerize, and in particular those that form amyloids, is sample preparation. Indeed, the buffer used for the synthesis/purification processes and its composition (counter-ions, mainly chloride ions or trifluoroacetic acid TFA^- ions used for

peptide chemical synthesis) may affect fibrillization [33]. For this reason, batch-to-batch variability of synthetic peptides is frequently reported, resulting in a poor reproducibility of experiments in term of kinetics of self-assembling. Proper storage conditions are also important. Typically, peptides are most stable when stored lyophilized at −20 °C. Nevertheless, even under this condition, hydration occurs. This may affect the presence of pre-formed aggregates in samples. Therefore, in order to ensure comparison between various experiments, we do not recommend storing peptides and proteins for a long time and rather suggest ordering or preparing a fresh batch and performing experiments in a short timeframe.

2. Use of ThT staining is a widely used method and relatively well accepted as an indicator of the presence of amyloid fibrils. Nevertheless, it has to be taken with precaution as ThT can also be an inhibitor of fibrillization [34].
3. A frequently used surface modification protocol for AFM involves fusing unilamelar lipid vesicles on glass or mica to form a supported lipid bilayer that can then be exposed to the protein fibrils. Indeed, interaction of amyloids with membrane is commonly reported and has been observed for Hfq [35, 36].
4. D_2O solutions may be used in samples to avoid the spectral overlaps between the Amide I band and strong absorption band of water at 1640 cm^{-1}. Note that Amide I wavenumbers are lowered in D_20 environment (5–10 cm^{-1}). In this case Amide I/II bands are referred to as Amide I′/II′.
5. The interpretation of an amyloid FTIR spectrum should usually begin with the examination of the primary sequence of the protein/peptide. Indeed, side chains such as asparagine and glutamine, which are very common in amyloid-forming proteins, have IR vibrations that overlap in the Amide I band.

Acknowledgments

This work was supported by Université Paris Diderot, CNRS, CEA, and Synchrotron SOLEIL. We gratefully acknowledge help to MV from the French Embassy for their program for scientific and university cooperation. We are indebted to F. Gobeaux (CEA Saclay, Gif-sur-Yvette, France) and A. Deniset-Besseau (LCP, Université Paris-Sud, France) for many fruitful discussions. We thank Kimberly Stanek (University of Virginia) for her careful and critical reading of our manuscript.

References

1. Taghbalout A, Yang Q, Arluison V (2014) The Escherichia coli RNA processing and degradation machinery is compartmentalized within an organized cellular network. Biochem J 458: 11–22
2. Arluison V, Taghbalout A (2015) Cellular localization of RNA degradation and processing components in Escherichia coli. Methods Mol Biol 1259:87–101
3. Lavelle C, Busi F, Arluison V (2015) Multiple approaches for the investigation of bacterial small regulatory RNAs self-assembly. Methods Mol Biol 1297:21–42
4. Cayrol B, Geinguenaud F, Lacoste J, Busi F, Le Derout J, Pietrement O, Le Cam E, Regnier P, Lavelle C, Arluison V (2009) Auto-assembly of E. coli DsrA small noncoding RNA: molecular characteristics and functional consequences. RNA Biol 6:434–445
5. Busi F, Cayrol B, Lavelle C, LeDerout J, Pietrement O, Le Cam E, Geinguenaud F, Lacoste J, Regnier P, Arluison V (2009) Auto-assembly as a new regulatory mechanism of noncoding RNA. Cell Cycle 8:952–954
6. Vogel J, Luisi BF (2011) Hfq and its constellation of RNA. Nat Rev Microbiol 9:578–589
7. Fischer S, Benz J, Spath B, Maier LK, Straub J, Granzow M, Raabe M, Urlaub H, Hoffmann J, Brutschy B, Allers T, Soppa J, Marchfelder A (2010) The archaeal Lsm protein binds to small RNAs. J Biol Chem 285:34429–34438
8. Smirnov A, Wang C, Drewry LL, Vogel J (2017) Molecular mechanism of mRNA repression in trans by a ProQ-dependent small RNA. EMBO J 36:1029–1045
9. Fortas E, Piccirilli F, Malabirade A, Militello V, Trepout S, Marco S, Taghbalout A, Arluison V (2015) New insight into the structure and function of Hfq C-terminus. Biosci Rep 35:e00190
10. Maury CP (2009) The emerging concept of functional amyloid. J Intern Med 265:329–334
11. Ostrowski A, Mehert A, Prescott A, Kiley TB, Stanley-Wall NR (2011) YuaB functions synergistically with the exopolysaccharide and TasA amyloid fibers to allow biofilm formation by Bacillus subtilis. J Bacteriol 193:4821–4831
12. Zhou Y, Blanco LP, Smith DR, Chapman MR (2012) Bacterial amyloids. Methods Mol Biol 849:303–320
13. Aguilera P, Marcoleta A, Lobos-Ruiz P, Arranz R, Valpuesta JM, Monasterio O, Lagos R (2016) Identification of key amino acid residues modulating intracellular and in vitro microcin E492 amyloid formation. Front Microbiol 7:35
14. Refregiers M, Wien F, Ta HP, Premvardhan L, Bac S, Jamme F, Rouam V, Lagarde B, Polack F, Giorgetta JL, Ricaud JP, Bordessoule M, Giuliani A (2012) DISCO synchrotron-radiation circular-dichroism endstation at SOLEIL. J Synchrotron Radiat 19:831–835
15. Giuliani A, Jamme F, Rouam V, Wien F, Giorgetta JL, Lagarde B, Chubar O, Bac S, Yao I, Rey S, Herbeaux C, Marlats JL, Zerbib D, Polack F, Refregiers M (2009) DISCO: a low-energy multipurpose beamline at synchrotron SOLEIL. J Synchrotron Radiat 16:835–841
16. Wien F, Wallace BA (2005) Calcium fluoride micro cells for synchrotron radiation circular dichroism spectroscopy. Appl Spectrosc 59:1109–1113
17. Lees JG, Smith BR, Wien F, Miles AJ, Wallace BA (2004) CDtool-an integrated software package for circular dichroism spectroscopic data processing, analysis, and archiving. Anal Biochem 332:285–289
18. Micsonai A, Wien F, Kernya L, Lee YH, Goto Y, Refregiers M, Kardos J (2015) Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. Proc Natl Acad Sci U S A 112:E3095–E3103
19. Reinke AA, Gestwicki JE (2011) Insight into amyloid structure using chemical probes. Chem Biol Drug Des 77:399–411
20. Rhiannon GC, Christopher TG, Nicolas HV (2012) Characterization of fiber-forming peptides and proteins by means of atomic force microscopy. Curr Protein Pept Sci 13:232–257
21. Dufrene YF, Ando T, Garcia R, Alsteens D, Martinez-Martin D, Engel A, Gerber C, Muller DJ (2017) Imaging modes of atomic force microscopy for application in molecular and cell biology. Nat Nanotechnol 12:295–307
22. Liberelle B, Banquy X, Giasson S (2008) Stability of silanols and grafted alkylsilane monolayers on plasma-activated mica surfaces. Langmuir 24:3280–3288
23. Thompson RF, Walker M, Siebert CA, Muench SP, Ranson NA (2016) An introduction to sample preparation and imaging by cryo-electron microscopy for structural biology. Methods 100:3–15
24. Militello V, Casarino C, Emanuele A, Giostra A, Pullara F, Leone M (2004) Aggregation kinetics of bovine serum albumin studied by FTIR spectroscopy and light scattering. Biophys Chem 107:175–187

25. Byler DM, Susi H (1986) Examination of the secondary structure of proteins by deconvolved FTIR spectra. Biopolymers 25:469–487
26. Zandomeneghi G, Krebs MR, McCammon MG, Fandrich M (2004) FTIR reveals structural differences between native beta-sheet proteins and amyloid fibrils. Protein Sci 13:3314–3321
27. Wallace BA (2009) Protein characterisation by synchrotron radiation circular dichroism spectroscopy. Q Rev Biophys 42:317–370
28. Kardos J, Micsonai A, Pal-Gabor H, Petrik E, Graf L, Kovacs J, Lee YH, Naiki H, Goto Y (2011) Reversible heat-induced dissociation of beta2-microglobulin amyloid fibrils. Biochemistry 50:3211–3220
29. Guinier A (1955) Small-angle scattering of x-rays, Structure of matter series. Wiley, New York
30. Feigin LAS, Svergun DI, Taylor GW (1987) Structure analysis by small-angle X-ray and neutron scattering. Plenum Press, NY, USA
31. Feigin O, Kratky O (1982) Small angle x-ray scattering. Academic Press, New York/London
32. Guinier A (1994) X-ray diffraction: in crystals, imperfect crystals, and amorphous bodies. Courier Corporation, North Chelmsford
33. Kaneko I, Tutumi S (1997) Matters arising: conformations of β-amyloid in solution. J Neurochem 68:438–439
34. Stains CI, Mondal K, Ghosh I (2007) Molecules that target beta-amyloid. ChemMedChem 2:1674–1692
35. Herrera AI, Tomich JM, Prakash O (2016) Membrane interacting peptides: a review. Curr Protein Pept Sci 17:827–841
36. Malabirade A, Morgado-Brajones J, Marquez I, Seguin J, Trepout S, Wien F, Marco S, Velez M, Arluison V (2017) Membrane association of the bacterial riboregulator Hfq and functional perspectives. Sci Rep 7(1):10724
37. Vetri V, D'Amico M, Fodera V, Leone M, Ponzoni A, Sberveglieri G, Militello V (2011) Bovine Serum Albumin protofibril-like aggregates formation: solo but not simple mechanism. Arch Biochem Biophys 508:13–24
38. Vetri V, Militello V (2005) Thermal induced conformational changes involved in the aggregation pathways of beta-lactoglobulin. Biophys Chem 113:83–91
39. Francois-Heude M, Mendez-Ardoy A, Cendret V, Lafite P, Daniellou R, Ortiz Mellet C, Garcia Fernandez JM, Moreau V, Djedaini-Pilard F (2015) Synthesis of high-mannose oligosaccharide analogues through click chemistry: true functional mimics of their natural counterparts against lectins? Chemistry 21:1978–1991

Annexe B : Article 2

Correlative infrared nanospectroscopy and transmission electron microscopy to investigate nanometric amyloid fibrils: Prospects and challenges

Soumis à la revue *Journal of Microscopy* en aout 2018.

Correlative infrared nanospectroscopy and transmission electron microscopy to investigate nanometric amyloid fibrils: Prospects and challenges.

David Partouche[†‡*], Jérémie Mathurin[¶*], Antoine Malabirade[‡], Sergio Marco[§II], Christophe Sandt[†], Véronique Arluison[‡#], Ariane Deniset-Besseau[¶*@] and Sylvain Trépout[§II*@]

[†] Synchrotron SOLEIL, L'Orme des Merisiers Saint Aubin, 91410 Gif-sur-Yvette, France

[‡] Laboratoire Léon Brillouin LLB, CEA, CNRS UMR12, Université Paris Saclay, CEA Saclay, 91191 Gif-sur-Yvette, France.

[§] INSERM, U1196, Université Paris Sud, Université Paris-Saclay, F-91405 Orsay, France

[II] Institut Curie, PSL Research University, CNRS, UMR 91873348, F-91405 Orsay, France

[¶] Laboratoire de Chimie Physique, CNRS, Univ. Paris-Sud, Université Paris-Saclay, 91405, Orsay, France.

[#] Université Paris Diderot-Paris 7, Sorbonne Paris Cité, 75013 Paris, FRANCE

* Equivalent contribution

Keywords: Fibrils, Amyloid, Correlative measurements, Infrared nanospectroscopy, transmission electron microscopy.

[@]Correspondence to:

Ariane Deniset-Besseau

ariane.deniset@u-psud.fr,

Laboratoire de Chimie-Physique

Université Paris-Sud

Bâtiment 350 1er étage

91405 Orsay, France

Phone : +33 1 69 15 32 74

Sylvain Trepout

sylvain.trepout@inserm.fr

Institut Curie / INSERM U1196

Campus Universitaire d'Orsay, Bat 112

91405 Orsay cedex FRANCE

Phone : +33 1 69 86 30 81

Fax : +33 1 69 07 53 27

Summary

The propagation of a structural information through conformational changes in host-encoded amyloid proteins are at the root of many neurodegenerative disorders. While important breakthroughs have been made in the field, fundamental issues like the 3D-structures of the fibrils implied in some of those disorders are still to be elucidate. To better characterise those nanometric fibrils, a broad range of techniques is currently available. Nevertheless none of them is able to perform direct chemical characterisation of single protein fibrils. In this work, we propose to investigate the structure of a model amyloid-like protein using a correlative approach. The complementary techniques used are TEM and a newly developed infrared nanospectroscopy technique called AFM-IR. We propose to introduce and discuss the strategy that we have implemented as well as the protocol, challenges and difficulties encountered during this study to characterize amyloid assemblies nearly at the single–molecule level.

Introduction

Amyloids are nanostructures originating from the aggregation of multiple copies of peptide or protein monomers that eventually form fibrils. Often described as being the cause for the development of various diseases, amyloid fibrils are of major significance in the public health domain. To better characterise these fibrils, a broad range of techniques is currently available for the detection and visualisation of amyloid nanostructures. These include antibodies for amyloid identification *in vivo* or *in vitro*[1], light microscopy using specific fluorescent dies such as Thioflavin T and S, infrared (IR) spectroscopy, synchrotron radiation circular dichroism, small angle X-ray scattering for their structural characterisation and molecular imaging at the nanoscale using either atomic force microscopy (AFM) or transmission electron microscopy (TEM)[2]. Alone, none of them is able to perform direct chemical characterisation of single protein fibrils. To deeply characterise these samples, it is essential to correlate or cross-correlate data acquired with complementary techniques on the same object. As vibrational spectroscopy is a powerful tool to directly probe the secondary structure and conformational changes of proteins with a limited sample preparation, correlative analysis combining newly developed vibrational nanospectroscopy techniques and transmission electron microscopy[3] will permit structural characterization of amyloid assemblies nearly at the single–molecule level. In this work, a correlative approach between TEM and an infrared nanospectroscopy technique is proposed for investigating the structure of amyloid-like proteins. The strategy, protocol, challenges and difficulties encountered in this approach are introduced and discussed herein.

Highly sensitive analysis can be performed using the AFM-IR technique. This approach combines IR spectroscopy measurement and AFM[4] and thus benefits from the advantages of both techniques, allowing morphological and chemical identification at the nanometre scale. Nevertheless, AFM measurement fails in its poor spatial (XY axis) resolution

compared to TEM. A correlative analysis with TEM can solve the resolution issue and give a better structural and chemical description of the object[2]. The sample analysed herein is the c-terminal (CTR) region of a bacterial protein called Hfq. Hfq forms amyloid-like structures and is known to be a nucleic acid (NA) binding protein[5,8]. Hfq self-assembles into amyloid fibres due to the presence of a 38 amino acid residue sequence in its CTR region[6–8]. Only a patch of 11 amino acids over 38 are necessary for Hfq to self-assemble, but not to interact with NA. As this analysis could open perspectives in the future to correlatively image nucleoprotein complexes[9], the synthetic 38 amino acid long sequence (referred as CTR_{38} through the manuscript) rather than the shorter peptide with 11 amino acids was chosen for this work. With this sequence, in the absence of NA about 20% of the peptide may show an intermolecular β-sheet secondary structure characteristic of the amyloid moiety, as it possesses specific absorption bands in the mid IR region[10–13]: a conventional IR spectrum of CTR_{38} amyloid shows an amide I band (stretching of the C=O of the peptidic bond centred around 1650 cm^{-1}) with several contributions, and among them a component at 1626 cm^{-1} corresponding to intermolecular β-sheet that represents for this sample 19 ± 2% of the amide I peak area.

One of the keys for correlative analyses[14] is to find a physical substrate for holding the specimen during the analysis suitable for both TEM and AFM-IR measurements: *(i)* with an appropriate roughness for AFM, and *(ii)* transparent for IR and TEM analysis. Silicon nitride (Si_3N_4)-based supports are a good compromise and have been used in this work[3]. In the first part of our study, the relevance of these supports for the AFM-IR study of CTR_{38} fibres, which are so far among one of the smallest objects studied with this technique, are discussed. Thereafter, the correlative analysis of an isolated 4 nm-thick CTR_{38} fibre is presented.

Experimental results and discussion

Relevance of the Si_3N_4-based support

The Si_3N_4-based supports used in this work possess an asymmetrical pattern composed of 8 square windows (100 µm x 100 µm) and a rectangular window (100 µm x 350 µm) in the silicon wafer with a 50 nm thick free-standing Si_3N_4 film. Initially, these supports were specifically designed for TEM purposes, since the Si_3N_4 is transparent to electrons and thus presents a precise lateral shape compatible with the electron beam. These supports are homogeneous in thickness, and have a surface roughness of 1-2 nm, permitting the resolution of AFM topographic images of objects smaller than 10 nm. Finally, these supports have a low absorption in the IR range characteristic of amide I, which contains most of the information regarding secondary structure of the protein. Thus Si_3N_4 supports are good candidates for both AFM-IR and TEM experiments[15].

AFM-IR imaging on Si3N4 supports.

AFM-IR analysis of isolated CTR_{38} fibres was first performed on Si_3N_4 film with a silicon wafer beneath. These preliminary measurements ensured that the surface state and properties of the film did not affect data acquisition, and confirmed the sensitivity of the technique to measure isolated fibres with a thickness less than 10 nm. In AFM-IR, the IR absorption measurement is based on the detection of photo-thermal effects, induced after the IR radiation absorption[16,17], by the AFM cantilever tip in contact with the surface. As the resulting photothermal expansion involved in the process is very fast (hundred of nanoseconds), it acts as a shock and forces the cantilever to oscillate on its Eigenmodes. The amplitude of those oscillations are directly proportional to the light absorbed by the sample and give a direct measurement of the IR absorption under the tip, with a sensitivity close to 50 nm[18]. In term of resolution, as the AFM tip acts as the detector, the lateral (xy) resolution is dependent of the contact between the studied surface and the tip (few nanometres). To reach sensitivity below 10 nm along the z-axis, AFM-IR is carried out in enhanced resonance mode[19]. This enhanced resonance mode allows to increase the sensitivity by a factor Q_n

(quality factor) which is around 100 in our case. Furthermore, previous studies made on objects smaller than 10 nm with this mode[20] used gold-coated surfaces as well as gold-coated AFM tips, to make a gold-gold interaction. This interaction increases the optic field intensity under the tip and enhances the sensitivity of the technique[21]. Since gold is a heavy atom that will prevent optimal TEM imaging, in the present study only the tip possesses a gold coating. Nevertheless, working as antenna, gold tips alone can lead to an increase of local fields and are already used in other tip-based spectroscopic techniques such as Tip-Enhanced Raman Spectroscopy TERS[22].

We first needed to confirm that this local field enhancement was sufficient to analyse isolated CTR_{38} fibres with a thickness of 4 nm without any photodamages. The AFM-IR technique simultaneously acquires topographic images and an IR map at a given wavenumber of a region of interest or local spectra at a given point. Figure 1 presents AFM-IR analysis of CTR_{38} fibres. The topographic image (Figure 1A) shows two populations of objects: micrometric long fibres and small dots. These two populations were frequently observed with other amyloid structures[23] and revealed the co-existence of oligomeric forms with long fibrils that are formed by polymerisation of oligomers. If we compare the local spectra obtained on fibres (blue spectra, Figure 1B) with the one obtained on one dot (red spectrum, Figure 1B), we observe a characteristic absorption shoulder in the amide I band at ~ 1620 cm^{-1}, attributed to intermolecular β-sheet vibrations[10–13]. The ratio between the IR map at 1620 cm^{-1} and the IR map at 1660 cm^{-1} (Figure 1C) allows identifying β-sheet rich areas in the fibres. Heterogeneities are observed: some regions without fibre morphology show a high β-sheet absorption signal. However, this result is not surprising, as oligomers are also amyloid-like assemblies and as very often the fibrils change their conformation during and after polymerisation.

Note that, as previously described, an absorption band attributed to trifluoroacetic acid (TFA) used in the production of peptide is also detected. This band is centred around 1745 cm^{-1}, characteristic of TFA in interaction with non-biological amide[24] or with other protonated species[25]. Note that this TFA IR absorption band interfere with the shape of the amide I of fibres and oligomers and antagonizes the presence of a weak shoulder in the amide I region around 1700 cm^{-1}, which is another specific marker of the antiparallel ß-sheet structure.

Nevertheless, our preliminary analyses confirm that the enhancement obtained with a gold-coated AFM tip is sufficient to analyse 4 nm-thick objects without any photodamages. The next step was to reproduce the same experiments on the rectangular windows (in the silicon wafer) with a 50 nm thick free-standing Si_3N_4 film.

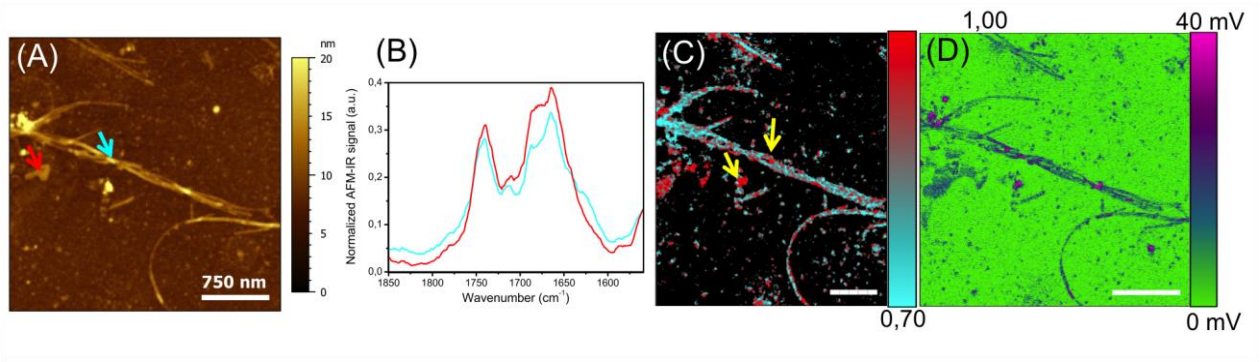


Figure 1: AFM-IR analysis of CTR_{38} Hfq fibre deposited on SI_3N_4 film above Si wafer. (A) AFM topography of 3µm x 3µm area containing a bundle of fibres and some oligomeric structures. (B) AFM-IR spectra taken on various part of the bundle of fibre (5 points / blue arrow) or in an oligomeric structure (red arrow). (C) Ratio between the 1620 cm^{-1} IR map and 1660 cm^{-1} IR map. Yellow arrows indicated area rich in B-sheet structure (red color) (D) IR mapping of the area at 1745 cm^{-1}.

Challenge of AFM-IR measurements on the free-standing Si3N4 film

Nevertheless, the experiments reveal some difficulties that lie directly in the detection method of the IR absorption used by the technique and that are linked to the use of 50 nm thick Si_3N_4 film. Indeed the film exhibits a small absorption that causes a rapid expansion of the free-standing Si_3N_4 film and its oscillation detectable during AFM-IR measurements. As previously explained, to reach nanometric sensitivity AFM-IR experiments are carried out in enhanced resonance mode. In this mode, the cantilever oscillations changed from a decaying transient mode to a continuous oscillation mode. Thus the repetition rate of the IR laser (see details in the Material and technical set-up section) is tuned to match on the contact resonance frequency of the cantilever. The resonance frequency depends on the cantilever spring constant and the sample stiffness. As a consequence, during AFM-IR measurements, it is necessary to track the peak frequency maximum of a given eigenmode. Any changes in term of stiffness of the support (Si_3N_4 film) may induce a shift of the AFM tip contact resonance frequency that may hamper the sensitivity of the AFM-IR method.

To evaluate the influence of the Si_3N_4 film, the global photo-thermal response of the wafer was tested for different areas: the area corresponding to the free-standing Si_3N_4 film; *(ii)* the Si_3N_4 film with the silicon wafer beneath and *(iii)* the border of the free Si_3N_4 film near the silicon wafer. As shown in Figure 2A, for a given power, a thermal dilation of the substrate is detected in each area. Nevertheless, the oscillations are more intense (~ 3 fold) in the case of the free-standing Si_3N_4 film with a shift of 25 kHz in the contact resonance frequency. Thus, even if its IR absorption in the range of the amide I band is low, it is absorbing enough to observe a specific oscillation of the Si_3N_4 film on or outside the silicon wafer. Therefore, under IR light, the film absorbs energy, retains the heat[26,27], and starts to quickly expand and oscillate. The oscillation intensity increases with the laser power but is independent of laser wavenumber contrary to photo-thermal expansion observed for the fibres. Interestingly, the difference of stiffness between the film and the CTR_{38} fibre induces a shift of the contact

resonance frequency (Figure 2B) allowing discrimination of the photo-thermal expansion of the fibre from the film oscillations. The intensities of both the film and sample responses are the same order of magnitude (Figure 2B) but, using lower IR laser power, it is possible to minimise the Si_3N_4 film response compared to the photo-thermal expansion of the sample. The decrease of the IR laser power was effective during AFM-IR analysis of fibres on the Si_3N_4 film with the silicon wafer beneath. But for the area corresponding to the free-standing Si_3N_4 film (used for correlative measurement), the film is less rigid, and the oscillation of the film is amplified. As a consequence, even with lower IR laser power, the effect of the thermal expansion of the film may hamper proper tracking of the tip contact resonance frequency. One possibility is to rely upon the field enhancement induced with the gold tip. However, this enhancement is highly dependent on the local nanostructure of the tip gold coating[29] especially in the case of the commercial tip used for this study.

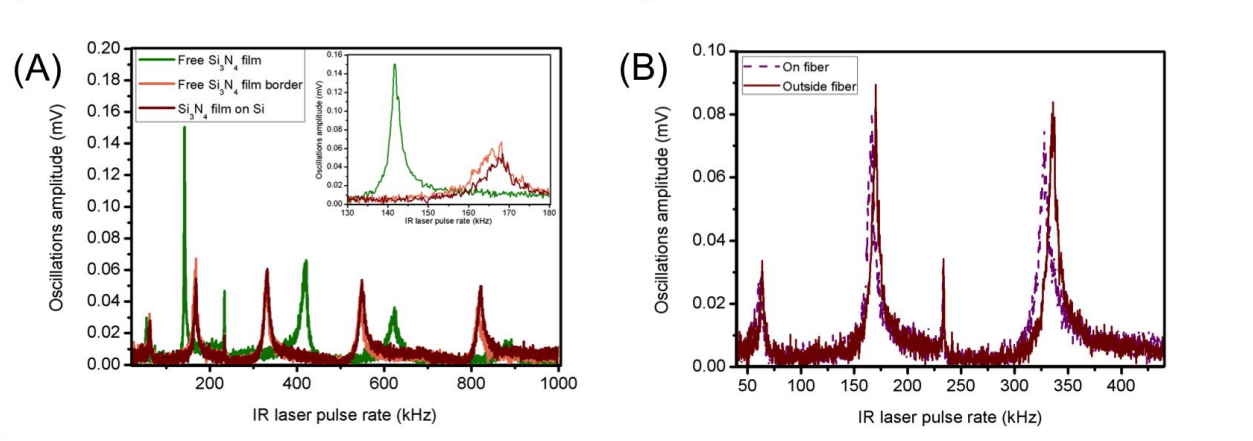


Figure 2: Study of the Si_3N_4 film bending. (A) Evolution of the oscillations of the cantilever in function of the position on the support (on the middle of the Si_3N_4 film, on the Si_3N_4 above the Si wafer or on free-Si_3N_4 at the edge of the Si wafer) – Insert: zoom on 130 kHz – 180 kHz range (B) Comparison of the frequencies of the cantilever oscillation in contact with the Si_3N_4 surface (oscillation of the film) or in contact with a Hfq CTR fibre (photo-thermal expansion).

Nevertheless, after several tests it was possible to select a tip with a sufficient near-field enhancement to detect the photo-thermal expansion of the sample, in such a way that measurements were not being affected by the film oscillations. Furthermore, constraining the liberty of the Si_3N_4 will improve the signal-to-noise ratio. As a consequence, finding the optimal thickness for the Si_3N_4 film where AFM-IR analysis is less affected by the thermal response of the film without blocking TEM measurement is also a possible way to improve correlative measurements. Electron microscopy grids of various Si_3N_4 films thicknesses are commercially available but their evaluation is beyond the scope of this study.

Despite those limitations, the analysis of 4 nm-thick CTR_{38} fibres on free-standing Si_3N_4 film was possible with a suitable gold tip and a rigorous following of the contact resonance frequency of the cantilever allowing the correlative approach AFM-IR-TEM.

TEM imaging

After careful comparison of the images obtained by AFM before and after TEM, it appears that TEM imaging damaged the amyloid fibres. The damage most probably originated from the fact that amyloid fibres were not stained using heavy atom salts as it is usually performed in electron microscopy. This omission of heavy salts was deliberate, since heavy atoms could perturb AFM and AFM-IR measurements. When performing correlative studies, the use of the identical specimen is of the essence in this regard, and the sample preparation protocol might not fit the requirements of all methods used in the correlative work. A trade-off on the sample preparation protocol has to be made, and we chose to skip the use of heavy-atom staining knowing that our sample would be more sensitive to electrons. By way of contrast, for TEM no difference in the sample measurement was observed if the fibres were first observed by AFM (data not shown). It is indeed possible to adjust the strength of the contact between the AFM tip and the sample in order to limit tip-sample interaction, thus limiting

damages to fibres. As a consequence, the area was first explored by AFM-IR and then by TEM.

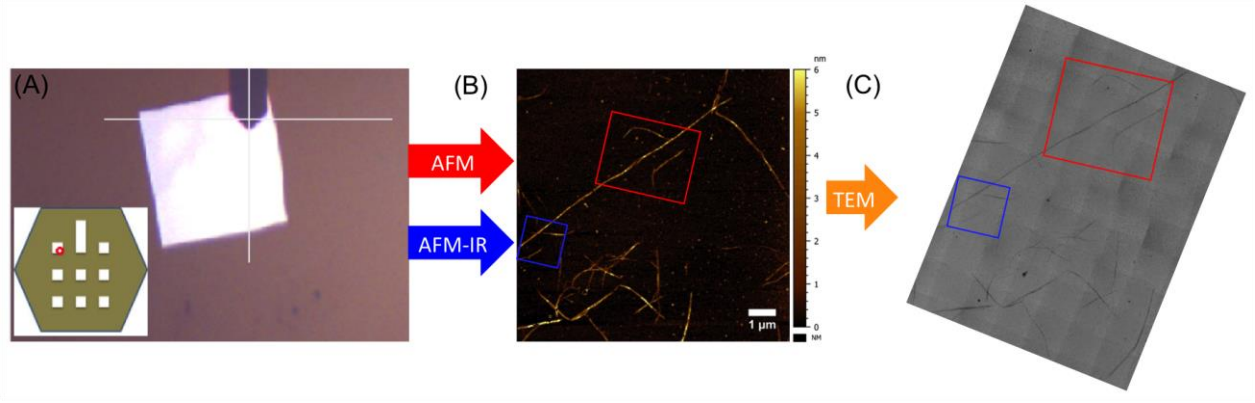


Figure 3: Strategy for correlative measurements. A) Light microscopy image of the 100x100µm top-left window with the AFM cantilever. Insert: scheme of the Si_3N_4 grid with the region targeted by AFM (red circle) for an easier correlation with TEM. To correlate TEM and AFM images, simple mathematical operations such as flipping and rotation were performed. B) 10 x 10 µm topographic AFM image to localize the fibres and estimate the distribution of individual CTR_{38} Hfq fibres. The detailed studies of both area indicated with a red and a blue square are presented in Figure 4 and 5. The sample is then study using TEM. C) TEM image of the same area observed by AFM. Globally the same fibres are detected despite some variations attributable to the imaging specificities of each method.

Correlative measurements

To perform the correlative analysis, the AFM cantilever was systematically positioned above the corners of the Si_3N_4 windows using light microscopy. Several corners were then scanned by AFM at low resolution to facilitate the recovery of the areas in TEM. AFM-IR spectra were collected on the topographically best-resolved fibres. After AFM and AFM-IR analysis, the grids were stored in a dry, dust-protected support and then transferred

to the TEM. Mosaic imaging at the TEM was performed to recover large fields of view and, thanks to the asymmetry of the grid, the areas previously observed in AFM and AFM-IR were easily recovered.

The correlative analysis between AFM-IR and TEM presented herein was carried out on the grid area highlighted by a blue square in Figure 3B. Using high magnification images, the fibres show various X-Y widths and Z-height thicknesses (Figure 4A). The dried fibre thickness is 3.8 nm ± 1.9 nm (n = 100) as measured by AFM (Figure 4C), a value in agreement with liquid AFM measurements made independently (Supplementary Figure S1). These measurements do not show a distribution of the thicknesses that would be expected if the fibres were stacked (in multiples of 3.8 nm). The measurements rather indicate a helicoidal structure of the fibres, where the thickness evolves along the fibre as it coils.

In TEM, amyloid fibres are detected as long thin dark filaments (up to few µm, Figure 3C). Due to the transparency of Si_3N_4 to electrons, it is possible to visualise single fibres, but also thicker filaments, as previously observed by AFM. Using higher magnification images, it is possible to discriminate the presence of two fibres coiling around each other (Figure 4A). These events resulted in variability of the fibre width, from 5 nm for the thinnest ones, up to 30 nm for the thickest ones (Figure 4D).

In the topographic images, the fibres appear larger than what is observed in TEM (Figure 4 A, B). As expected, after multiple measurements made on the fibres, we determined that AFM over-estimates the fibre width by about 14 nm (Figure 4D). This difference is explained by the convolution of the tip radius in AFM. As a consequence, t**he** fibre width measured by AFM is not fully consistent with the results obtained by TEM, an observation that strengthens the need for correlative imaging.

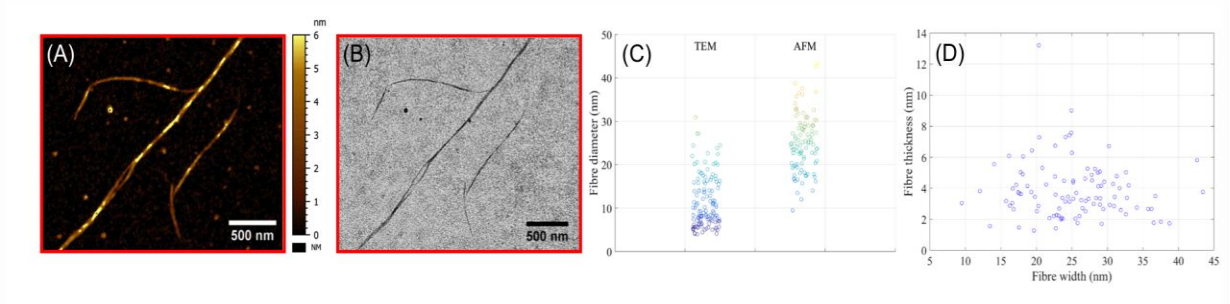


Figure 4: Structural characterisation of amyloids fibres deposited on a Si_3N_4 TEM window. These images are zoom-ins of the red area depicted Figure 3. A) AFM close-views on the fibres where detailed information can be obtained. B) Equivalent TEM image. C) Plot of the fibre width measured by TEM and AFM. The average width measured by AFM is 25.17 +/- 6.53 nm (n=100) whereas the average width measured by TEM is 11.20+/-5.68 nm (n=126). AFM measurements show an important increase of the fibre width. Width and thickness were measured on the same fibres observed in TEM and AFM on Si_3N_4 substrate. D) Plot showing the fibre thickness as a function of the width measured in AFM. The measurements do not agree with a stacking of the fibres but rather agree with a structure which height evolves along the its length.

On the same area of the specimen, we acquired local IR spectra to evaluate the local secondary structure of the fibre (Figure 5 A, B). Two different CTR_{38} fibres were analysed. In both spectra (green and blue curves, Figure 5C, D), four or five main contributions can be determined, attributed to intermolecular β-sheets, amyloid β-sheet, α-helix, β-turn, and antiparallel pleated sheet conformations[30]. Precisely, the amide I band of the amyloid fibre presents a maximum at 1666 cm^{-1}, a weak shoulder at 1703 cm^{-1} and a strong shoulder at 1626 cm^{-1}, a signature that is characteristic of amyloid structures (Figure 5D). The spectra obtained from the fibres are compared to a global IR absorption spectrum obtained in IR microspectroscopy from a CTR_{38} sample in bulk solution (dashed black line, Figure 5D).

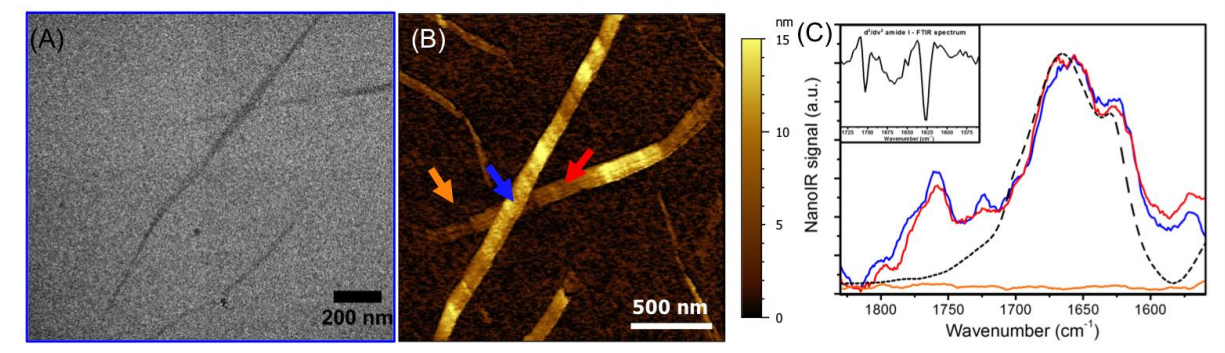


Figure 5: Correlative measurements: A) TEM image showing where AFMIR measurements were performed (blue zone depicted Figure 3). B) Local IR spectra (colored arrows on topographic image) overlaid with global FTIR spectrum (dashed black line). The two main absorption bands 1665 and 1625 cm^{-1} are easily identified with the second derivative signal of the IR local spectra (upper insert).

Comparing the local IR spectra and the global IR spectrum obtained, we can observe variations in the band ratio especially if we compare the ratio between 1626 cm^{-1} and 1665 cm^{-1}. At this scale, a part of the fibre probed may be slightly different in term of composition, inducing a change in the ratio between the random and the amyloid β-sheets bands. But some differences might also due to the support as FTIR measurements was performed using CaF_2 slide. Since decades, it's very well known that proteins adsorption may differ in term of density, conformations, and orientations, from one surface to another. Thus, surface chemistry plays a fundamental role in protein adsorption. Here, both CaF_2 and Si_3N_4 are known to be hydrophilic but with different charge surface properties that can induce some conformation variations.

Conclusion

In conclusion, we demonstrated for the first time the possibility to perform a correlative analysis with AFM-IR and TEM. This tandem approach allowed combining the ultrafine lateral resolution of TEM, the ultrafine axial resolution of AFM, and the chemical

and structural information brought by AFM-IR spectroscopy at the nanometre scale to give a detailed description of small biological fibres, demonstrated with the amyloid Hfq CTR_{38} fibres. In specific conditions it is possible to use an Si_3N_4 film support stage for these measurements. Although the interactions between the Si_3N_4 film and the AFM-IR tip can be limiting, it was possible to achieve IR spectroscopy of 4 nm-thick single amyloid fibres, which is close to the maximum sensitivity achievable for such a system. We suggest that it should still be possible to improve the sensitivity of the analysis by changing different parameters, for example increasing the thickness of the Si_3N_4 film or reducing the dimension of the Si_3N_4 windows. The AFM-IR measurements performed during this study also showed that using only enhancement from the gold-coated tip was sufficient to study nanometric fibres. These enhancements could pave the way to new supports for other correlative analyses.

Material and technical set-up

Amyloid fibres preparation

Amyloid fibres were prepared with the C-terminal region (CTR) of the Hfq protein (CTR_{38}, namely residues 64 to 102 of full length protein, SRPVSHHSNNAGGGTSSNYHHGSSAQNTSAQQDSEETE)[7]. Hfq CTR_{38} peptide was produced chemically (ProteoGenix, France). The lyophilized peptide was reconstituted in MilliQ water at 20 mg/mL^{-1} (stock solution). In average, self-assembly of Hfq-CTR into amyloid fibres occurs after a few weeks at 4°C. The process of self-assembly was confirmed by FTIR spectroscopy and TEM[2].

Fourier Transform Infrared (FTIR) spectroscopy

Infrared absorption measurements were performed in the region between 4000 and 400 cm^{-1} in order to control for the presence of amyloid absorbance signature. 1-2 µl of the CTR_{38} stock solution were deposited on a CaF_2 window, dried at room temperature and analysed

using a Fourier-transform infrared microscope (IN10, ThermoScientific) equipped with a conventional IR source, a 15X Schwarzschild objective, and a Mercury Cadmium Telluride (MCT) detector. A total of 256 scans were collected with a spectral resolution of 4 cm^{-1} on 100 µm x 100 µm spots of the dried solution. Spectra were background-subtracted and baseline-corrected (polynomial baseline correction) before analysis.

Deposition of the amyloid fibres on the Si3N4 support

A 2µl droplet of Hfq-CTR fibres diluted 1:1500 in water was deposited on Si_3N_4 windows (SN100-A50Q33/EM Resolutions Ltd: 100 micron thick frame with a 50 nm film on 9 windows; 8 windows 100 x 100 micron and 1 window 100 x 350 micron; see Figure 1). Before use, Si_3N_4 windows were treated using a plasma-cleaner (80% O_2, 20% Ar mixture at 0.6 mBar pressure for 20 seconds). Plasma allowed optimising the adsorption of the sample. Where needed, the samples were kept in a vacuum cabinet before analysis.

Transmission electron microscopy (TEM) analysis

The Si_3N_4 grids previously observed by AFM were mounted onto a room temperature holder and introduced into a JEOL JEM2200FS AFM instrument (JEOL, Japan) operated at an acceleration voltage of 200 kV. The images (2048 x 2048 pixels) were acquired on a Gatan UltraScan894 US1000 slow scan CCD camera (Gatan, USA). Mosaic images were collected using an in-house developed software to allow imaging of large areas, facilitating the recovery of the zone previously imaged by AFM. Negative staining was omitted, in order to keep the fibres in the same conditions as those used in AFM. The TEM images (in Figure 4) were obtained by the reconstruction of several individual images (at least 4 images) using the MosaicJ plugin on imageJ[31,32].

Atomic Force Microscopy (AFM) and Resonance-enhanced AFM-IR spectroscopy analysis

AFM topographic images and nanoscale infrared measurement were performed with a nanoIR2-s microscope (Anasys Instruments, USA), which combines an atomic force microscope with a tunable infrared quantum cascade laser (QCL source, 1550-1960cm^{-1} stage, MIRcat-QT, DAYLIGHT solutions; peak powers up to 1W; average powers up to 0.5W, and wavelength repeatability <0.1 cm^{-1}. In these experiments the sample was top-side illuminated. AFM-IR measurements were carried out directly on the Si_3N_4 grid at room temperature. We used a gold coated silicon AFM probe to avoid effects linked to the IR absorption of the silicon (HQ:CSC38/CR-AU, MIKROMASCH). This gold AFM probe has a force constant of 0.03$N.m^{-1}$ and a tip radius <35 nm (including coating). AFM topographies were acquired in contact mode, with a 0.5 Hz scan rate (500 x 500 pixels). Using this setup, two modes of acquisitions are possible. In one mode, the laser wavenumber is fixed and we acquire simultaneously the topographic image and IR absorption mapping at this wavenumber. In the other mode, the AFM tip is fixed at a given position to acquire a local IR spectrum by tuning the laser wavenumber. AFM images and IR maps were analysed by using Mountainsmap 7.3 software (Digital Surf, France). Spectra were collected with a sampling of 1 cm^{-1} within the range 1560-1850 cm^{-1} and averaged from at least 10 measurements. A Savitzky –Golay filter (second order, 15 pts) was then applied.

NOTES

The authors declare any competing financial interest(s).

ACKNOWLEDGEMENTS

This work was supported by CNRS, CEA, Synchrotron SOLEIL, University Paris-Saclay and Université Paris Diderot. We are indebted to Dr. Marisela Velez (UAM, Madrid, Sapin)

for her help in liquid AFM control measurements and for many fruitful discussions. We thank Dr. Richard Lease (The Ohio State University, OH, USA) and Dr. Valeria Militello (Universita degli Studi di Palermo, Palermo, Italy) for their careful and critical reading of our manuscript.

Bibliography

1. Moreno-Del Álamo, M., De La Espina, S. M. D., Fernández-Tresguerres, M. E. & Giraldo, R. Pre-amyloid oligomers of the proteotoxic RepA-WH1 prionoid assemble at the bacterial nucleoid. *Sci. Rep.* **5,** 1–12 (2015).
2. Partouche, D. *et al.* Techniques to Analyze sRNA Protein Cofactor Self-Assembly In Vitro. *Methods Mol. Biol.* **1737,** 321—340 (2018).
3. Hermelink, A. *et al.* Towards a correlative approach for characterising single virus particles by transmission electron microscopy and nanoscale Raman spectroscopy. *Analyst* **142,** 1342–1349 (2017).
4. Dazzi, A. & Prater, C. B. AFM-IR: Technology and Applications in Nanoscale Infrared Spectroscopy and Chemical Imaging. *Chem. Rev.* acs.chemrev.6b00448 (2016). doi:10.1021/acs.chemrev.6b00448
5. Cech, G. M. *et al.* The Escherichia Coli Hfq Protein: An Unattended DNA-Transactions Regulator. *Front. Mol. Biosci.* **3,** (2016).
6. Arluison, V. *et al.* Three-dimensional Structures of Fibrillar Sm Proteins: Hfq and Other Sm-like Proteins. *J. Mol. Biol.* **356,** 86–96 (2006).
7. Fortas, E. *et al.* New insight into the structure and function of Hfq C-terminus. *Biosci. Rep.* **35,** 1–9 (2015).
8. Malabirade, A. *et al.* Membrane association of the bacterial riboregulator Hfq and functional perspectives. *Sci. Rep.* **7,** 1–12 (2017).
9. Malabirade, A. *et al.* Compaction and condensation of DNA mediated by the C-terminal domain of Hfq. *Nucleic Acids Res.* **45,** 7299–7308 (2017).
10. Militello, V. *et al.* Aggregation kinetics of bovine serum albumin studied by FTIR spectroscopy and light scattering. *Biophys. Chem.* **107,** 175–187 (2004).
11. Zandomeneghi, G., Krebs, M. R. H., McCammon, M. G. & Fändrich, M. FTIR reveals structural differences between native beta-sheet proteins and amyloid fibrils. *Protein Sci.* **13,** 3314–21 (2004).
12. Hiramatsu, H. & Kitagawa, T. FT-IR approaches on amyloid fibril structure. *Biochim. Biophys. Acta - Proteins Proteomics* **1753,** 100–107 (2005).
13. Hoffner, G., André, W., Sandt, C. & Djian, P. Synchrotron-based infrared spectroscopy brings to light the structure of protein aggregates in neurodegenerative disease. *Rev. Anal. Chem.* **33,** 231–243 (2014).
14. Latour, G. *et al.* Correlative nonlinear optical microscopy and infrared nanoscopy reveals collagen degradation in altered parchments. *Sci. Rep.* **6,** 26344 (2016).

15. Ay, F. & Aydinli, A. Comparative investigation of hydrogen bonding in silicon based PECVD grown dielectrics for optical waveguides. *Opt. Mater. (Amst).* **26,** 33–46 (2004).
16. Dazzi, A., Prazeres, R., Glotin, F. & Ortega, J. M. Local infrared microspectroscopy with subwavelength spatial resolution with an atomic force microscope tip used as a photothermal sensor. *Opt. Lett.* **30,** 2388–2390 (2005).
17. Dazzi, A., Glotin, F. & Carminati, R. Theory of infrared nanospectroscopy by photothermal induced resonance. *J. Appl. Phys.* **107,** 124519 (2010).
18. Felts, J. R. *et al.* Atomic force microscope infrared spectroscopy on 15 nm scale polymer nanostructures. *Rev. Sci. Instrum.* **84,** (2013).
19. Lu, F. & Belkin, M. A. Infrared absorption nano-spectroscopy using sample photoexpansion induced by tunable quantum cascade lasers. *Opt. Express* **19,** 19942 (2011).
20. Ruggeri, F. S. *et al.* Nanoscale studies link amyloid maturity with polyglutamine diseases onset. *Sci. Rep.* **6,** 31155 (2016).
21. Lu, F., Jin, M. & Belkin, M. A. Tip-enhanced infrared nanospectroscopy via molecular expansion force detection. *Nat. Photonics* **8,** 307–312 (2014).
22. Xiong, M. H., Bao, Y., Yang, X. Z., Zhu, Y. H. & Wang, J. Delivery of antibiotics with polymeric particles. *Adv. Drug Deliv. Rev.* **78,** 63–76 (2014).
23. Ruggeri, F. S. *et al.* Infrared nanospectroscopy characterization of oligomeric and fibrillar aggregates during amyloid formation. *Nat. Commun.* **6,** 7831 (2015).
24. Purkina, A. V, Volchek, B. Z. & Bureau, C. \r + r2. 1051–1052 (1971).
25. Redington, R. L. & Lin, K. C. Infrared spectra of trifluoroacetic acid and trifluoroacetic anhydride. *Spectrochim. Acta Part A Mol. Spectrosc.* **27,** 2445–2460 (1971).
26. Ftouni, H. *et al.* Specific heat measurement of thin suspended SiN membrane from 8 K to 300 K using the 3ω-Völklein method. *Rev. Sci. Instrum.* **84,** (2013).
27. Ftouni, H. *et al.* Thermal conductivity of silicon nitride membranes is not sensitive to stress. *Phys. Rev. B - Condens. Matter Mater. Phys.* **92,** (2015).
28. Chen, G. & Hui, P. Thermal conductivities of evaporated gold films on silicon and glass. *Appl. Phys. Lett.* **74,** 2942–2944 (1999).
29. Asghari-Khiavi, M. *et al.* Exploring the origin of tip-enhanced Raman scattering; Preparation of efficient TERS probes with high yield. *J. Raman Spectrosc.* **43,** 173–180 (2012).
30. Barth, A. Infrared spectroscopy of proteins. *Biochim. Biophys. Acta - Bioenerg.* **1767,** 1073–1101 (2007).
31. Thévenaz, P. & Unser, M. User-friendly semiautomated assembly of accurate image mosaics in microscopy. *Microsc. Res. Tech.* **70,** 135–146
32. Abràmofff, M. D., Magalhães, P. J. & Ram, S. J. Image processing with ImageJ Part II. *Biophotonics Int.* **11,** 36–43 (2005).

Supplementary :

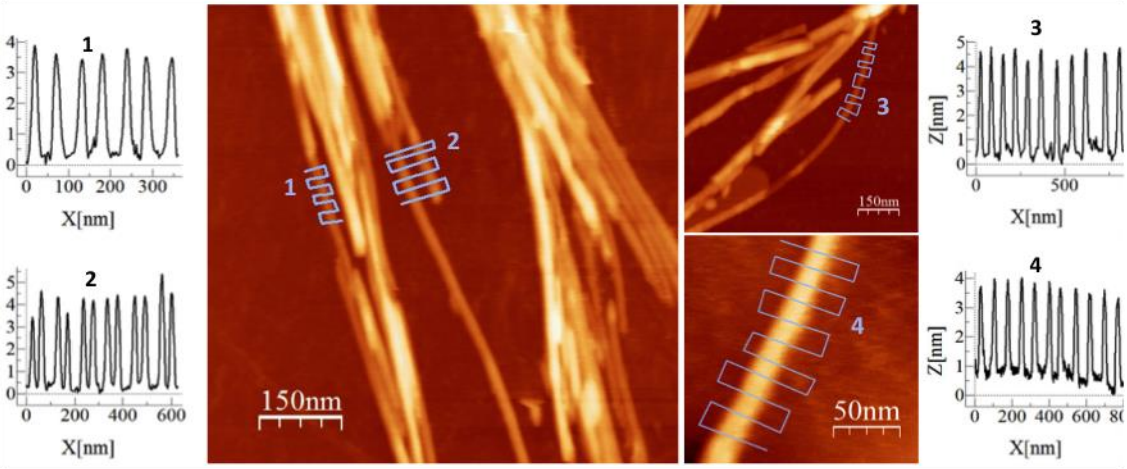


Figure S1: Topographic images of Hfq-CTR fibres in an acetate buffer (100mM, pH4) obtained with an Agilent technologies 5500 microscope (Santa Clara, CA, USA) operated in tapping mode at a resonance frequency of 15 kHz with 0.73 $N.m^{-1}$ spring constant cantilever (Olympus RC800PSA). Z-profiles 1, 2, 3 and 4 (grey lines in the pictures), show height of single fibres.

Annexe C : Article 3

Revised role for Hfq bacterial regulator on DNA topology

Article publié : *Scientific Reports*, 2018

(DOI : 10.1038/s41598-018-35060-9)

Revised role for Hfq bacterial regulator on DNA topology

Antoine Malabirade[1*], David Partouche[1,2*], Omar El Hamoui[2], Florian Turbant[1], Frédéric Geinguenaud[3], Pierre Recouvreux[4], Thomas Bizien[2], Florent Busi[5,6], Frank Wien[2] and Véronique Arluison[1,6]

[1] Laboratoire Léon Brillouin LLB, CEA, CNRS UMR12, Université Paris Saclay, CEA Saclay, 91191 Gif-sur-Yvette, France

[2] Synchrotron SOLEIL, L'Orme des Merisiers, Saint Aubin BP48, 91192, Gif-sur-Yvette, France

[3] UFR SMBH, Université Paris 13, Sorbonne Paris Cité, 93017 Bobigny, France

[4] Aix Marseille Université, CNRS, IBDM, Marseille, France

[5] Unité de Biologie Fonctionnelle et Adaptative, CNRS UMR8251, Université Paris Diderot, 75013 Paris, France

[6] Université Paris Diderot, Sorbonne Paris Cité, 75013 Paris France

[*] Equivalent contribution

Corresponding authors: Véronique Arluison, LLB CEA Saclay 91191 Gif-sur-Yvette France; Tel 33(0)169083282; veronique.arluison@univ-paris-diderot.fr

***Keywords**:* Sm protein; nucleoid-associated protein (NAP); DNA topology

Abstract

Hfq is a pleiotropic regulator that mediates several aspects of bacterial RNA metabolism. The protein notably regulates translation efficiency and RNA decay in Gram-negative bacteria, usually *via* its interaction with small regulatory RNA. Besides these RNA-related functions, Hfq has also been described as one of the nucleoid associated proteins shaping the bacterial chromosome. Therefore, Hfq appears as a versatile nucleic acid-binding protein, which functions are probably even more numerous than those initially suggested. For instance, *E. coli* Hfq, and more precisely its C-terminal region (CTR), has been shown to induce DNA compaction into a condensed form. In ***this paper, we establish that DNA induces Hfq-CTR amyloidogenesis, resulting in a change of DNA local conformation***. Furthermore, we clarify the effect of Hfq on DNA topology. Our results evidence that, even if the protein has a strong propensity to compact DNA thanks to its amyloid region, it does not affect overall DNA topology. We confirm however that *hfq* gene disruption influences plasmid supercoiling *in vivo*, indicating that the effect on DNA topology in former reports was indirect. Most likely, this effect is related to small regulatory sRNA-Hfq-based regulation of another protein that influences DNA supercoiling, possibly a nucleoid associated protein such as H-NS or Dps. Finally, we hypothesise that this indirect effect on DNA topology explains, at least partially, the previously reported effect of Hfq on plasmid replication efficiency.

INTRODUCTION

The hexameric Hfq is an abundant and phylogenetically conserved protein present in about half of bacteria and in some archaeal species. Consistent with its high cellular level, the protein coordinates multiple roles inside the bacterial cell. For instance, the initial description of Hfq was associated with its involvement in the activity of bacteriophage Qβ replicase (hence its name Host factor for bacteriophage Qβ[1]). The pleiotropic functions of Hfq were however highlighted later when the *hfq* gene was disrupted in *Escherichia coli*, decreasing bacterial growth rate, changing mutagenesis rate, increase in UV sensitivity, oxidant and osmo-sentivity, as well as a decreasing plasmid supercoiling[2,3]. Most of these *hfq*-null phenotypes are due to Hfq-dependant small noncoding RNA (sRNA) based regulations[4]. Indeed, Hfq is required to mediate sRNA stress-response[5]. This regulatory mechanism is based on the hybridisation of the sRNA to its target mRNA, therefore altering the mRNA translation and stability[6]. This effect can be either negative or positive. For instance the same sRNA can positively regulate a mRNA target, while it regulates negatively another[7-9].

Besides, Hfq also binds DNA and has been described as one of the *E. coli* nucleoid associated proteins (NAP)[10-13]. This was notably revealed *in vivo* by cellular localisation experiments demonstrating that about 20% of Hfq was complexed with DNA within the bacterial nucleoid[13-15].

Interestingly, Hfq is structurally related to the Sm eukaryotic family of proteins, which participates in nucleic acid-related processes, including splicing, telomeres replication, RNA decapping and decay[16,17]. Indeed, the amino-terminal region of Hfq (about 65 amino acid residues) folds similarly to Sm proteins. This region comprises a bent antiparallel β-sheet, capped by an N-terminal α-helix. The β-sheets from six monomers interact with each other to assemble in a toroidal structure[18,19]. Although the mechanism by which Hfq binds nucleic acids is not fully understood, it is now well-established that the inner pore on the proximal face of the torus (on which the α-helix is exposed) binds U-rich RNA, that A-rich sequences bind to the distal face, and that the lateral rim also binds A/U-rich sequences[20-22]. The distal face and the rim of the protein are involved in both DNA and RNA binding, while the proximal face seems to be involved in RNA fixation only[11,20]. Besides its Sm-like domain, Hfq C-terminal region (CTR) also plays a role in nucleic acid recognition[23,24]. We previously showed that Hfq and precisely its CTR bridges distant regions of plasmids[12,13] and highlighted that it induces a strong compaction of DNA[24]. Note that DNA compaction results

from a synergy of different mechanisms, including DNA-bridging, DNA-bending, DNA-supercoiling, self-assembly of nucleoid proteins, as well as phase separation[25,26]. Thus Hfq DNA-bridging may just be one contributor to nucleoid compaction. 3D-structures of various Hfqs have been resolved[21,22,27-30], all lack the CTR and the way this region folds remains unknown so far. It has however been demonstrated that Hfq is able to self-assemble and that this characteristic is dependent on the presence of CTR, an observation in accordance with the observation that the flexible CTR could facilitate protein:protein interactions[31,32]. Indeed, it has recently been shown that the CTR region of *E. coli* Hfq forms an amyloid-like structure[33-35], explaining why the protein is able to self-assemble *in vitro* and *in vivo*[12,31,33,36]. While the same region of Hfq is responsible of its propensity for self-assembly, DNA bridging and compaction[12,24], the relation between these properties remains unclear. Both the architectural bridging properties of nucleoid associated proteins (NAPs) and their ability to self-associate may have a role on bacterial DNA condensation[25].

So far, only a few studies have shed light on the role of Hfq in DNA metabolism[13], *e.g.* it has been shown to influence plasmid negative supercoiling *in vivo*[2], replication efficiency[37] and some studies have shown a role in transcriptional activity[38-40]. Some of the phenotypic effects due to the lack of Hfq may be attributed to defects in DNA-related processes; filamentation or slowed-down bacterial growth are for instance a sign of replication deregulation, sensitivity to mutagens could indicate unreliable DNA repair. Motivated by this unsolved role of Hfq in DNA-related processes, we have investigated how Hfq interaction with DNA affects the nucleic acid structure[24]. In this work, we focus our attention on two aspects of DNA:Hfq interaction, how binding to DNA induces Hfq self-assembly and how Hfq affects DNA structure and topology. Our results enable us to propose a mechanism underlying the function of Hfq in DNA packaging and its precise influence on plasmid supercoiling.

RESULTS

E. coli Hfq:DNA complex has been analysed previously *in vitro* by molecular imaging (transmission electron TEM and atomic force AFM microscopies)[12,23,24]. This allowed us to observe that Hfq tends to bridge two double-stranded distant DNA sections together, through its CTR amyloid-like region[12,23,24].

DNA induces amyloidogenesis of Hfq CTR

First, we focused our attention on the effect of DNA on Hfq self-assembly. Indeed, Hfq binding to DNA may be due to simple DNA binding of individual monomers or a cooperative self-assembly. A related question concerns the effect of DNA on amyloidogenesis of Hfq. Hfq-CTR self-assembly *in vitro* is quite a slow mechanism[33,34], and cellular co-factors probably help to accelerate this process. Such an effect has already been demonstrated for lipids[41]. In our analysis we used Synchrotron Radiation Circular Dichroism (SRCD), which allows extending the wavelength range down to 170 nm for the identification and distinction of amyloid proteins peaks[42]. Aggregation into β-sheets in an amyloidal structure implies a significant SRCD spectral change: negative band shift from ~ 200nm to ~ 210nm. As shown on figure 1, the red-shift observed upon DNA binding proves that the amyloid structure formation is induced by the interaction with DNA, while the same peptide in the absence of DNA remains unassembled even after 2 weeks.

We confirmed this result using Fourier Transform InfraRed (FTIR) spectroscopy (Fig. 2). FTIR is useful to identify the presence of amyloid fibrils. Indeed, a peak around 1645 cm^{-1} is indicative of random coil, 1655 cm^{-1} of α-helix, and 1630–1640 cm^{-1} of β-sheet[43]. Hydrogen-bonding in cross-β structure, the fingerprint of amyloids, induces a shift to lower wavenumbers so amyloid fibrils show a β-sheet absorption band below 1620 cm^{-1} [44]. Here D_2O buffered samples have been used to avoid the spectral overlaps between Amide I band and strong absorption band of water at 1640 cm^{-1} (in this case Amide I/II bands are referred as Amide I'/II'). The amide I' wavenumbers are then lowered in D_20 environment by 5-10 cm^{-1}. In Figure 2, we clearly observe a band at 1615 cm^{-1} indicative for the formation of amyloids in the presence of DNA.

Finally, the formation of the cross-β structure was also investigated by small angle X-ray scattering (SAXS) in the wide-angle region. SAXS curves are shown on figure 3. Weak β-sheet peaks are observed for Hfq-CTR in the presence of DNA (incubated 10 days, blue curve) and for Hfq-CTR alone incubated for 6 weeks at 20 mg/mL (to ensure the formation of amyloid fibrils in the absence of cofactor, black curve), while no β-sheet peak is observed for Hfq-CTR alone after 10 days (red curve). The peaks intensities on the blue curve are much lower than the reference (black curve) due to less incubation time resulting in less β-sheet formation. This definitely confirms that DNA promotes the formation of Hfq-CTR amyloid. Interestingly, peaks corresponding to the β-sheet structure are not the same with or

without DNA. A 0.21 Å shift for inter-sheet spacing (increase from 8.00 to 8.21 Å in the presence of DNA) and a 0.04 Å shift for inter-strand spacing (increase from 4.59 to 4.63 Å in the presence of DNA) towards lower q are observed while comparing both curves. This indicates a slight change in the structure of Hfq-CTR β-Sheet arrangement upon DNA addition. Taken together, these results firmly establish that Hfq amyloid formation is induced by the interaction with DNA.

Influence of Hfq on DNA topology

Nevertheless, the link with change in DNA structure and topology still remains unclear. It has been described previously that the linking number of plasmids isolated from cells carrying an *hfq* mutation is altered compared with DNA of WT strain[2]. This could be a direct consequence of the interaction of Hfq with DNA, or an indirect effect consecutively to a sRNA mediated regulation of other DNA-binding proteins, changing DNA topology.

In order to distinguish between these two possibilities, we studied the interaction of purified Hfq with plasmids *in vitro*. pHSG298 plasmid was incubated with purified Hfq and then treated with calf thymus topoisomerase I to relax unconstrained supercoils. Plasmid DNA was deproteinised and electrophoresed on native agarose gels (Figure 4). Under these conditions, the most supercoiled topoisomers migrate faster, and the relaxation induced by topoisomerase treatment slows down the migration. As shown in figure 4, in the absence of Hfq, plasmid DNA was completely relaxed by topoisomerase. The same result is observed when DNA is incubated with Hfq. This implies that Hfq does not constrain supercoils. This is in contrast to the effect observed for H-NS[45]. Thus, Hfq affects indirectly DNA topology reported previously *in vivo*, possibly due to the sRNA-based regulation of another protein affecting DNA supercoiling.

To confirm this result, we used magnetic tweezers to handle a single topologically constrained double stranded (ds) DNA molecule[46] and monitor its deformation upon Hfq binding in real time. Rotation of the magnets imposed a variable amount of supercoiling to a single 8.8 kbase pair (bp) dsDNA molecule. The extension of the molecule is measured for multiple values of the supercoiling σ (defined as the number of turns added or removed relative to the total number of turns in the relaxed naked molecule), giving access to the extension-versus-rotation response of the molecule in the presence or absence of Hfq (Fig. 5)[46]. Maximal extension of naked dsDNA (2.25 µm) is reached when no supercoiling is applied (σ = 0) (Fig. 5A red line). The apex is flanked by extensive shortening regimes, due

to formation of loops (plectonemes), either upon negative or positive supercoiling. Asymmetry in the response of naked DNA is due to denaturation of the double helix under negative supercoiling[46]. After this measurement, we injected a solution containing 60 nM Hfq without supercoiling ($\sigma = 0$). We then measured the extension-versus-rotation response of the molecule at a constant force of 0.25 pN (Fig. 5A, blue line) and observed a distinctive behaviour. The curve is still centered at $\sigma = 0$, indicating that no twist is applied to the DNA molecule through Hfq interaction. Hfq binding to DNA only slightly changes the maximal extension of the molecule under the same force (1.75 µm compared to 2.25 µm). This shortening could reflect a decrease in the persistence length of the filament[47] or could be due to the formation of kinks in the dsDNA backbone by Hfq. The extension-versus-rotation curve profile of Hfq-DNA filament also displays an apex flanked by plectonemic regimes. Strikingly, the width of the apex is unchanged, indicating that the torsional stiffness is not affected by Hfq binding. Slopes of the plectonemic regime are smaller, consistent with a decrease in persistence length[48]. We then forced Hfq-DNA filament extension to be close to 0 µm by holding it for several minutes at low force and no supercoiling ($F < 0.1$ pN and $\sigma = 0$). Consequently the extension-versus-rotation curve (Fig. 5A, green) displayed a broad and flat profile still centered at $\sigma = 0$ with a maximal extension shortened to 0.6 µm. This reflects the ability of Hfq to bridge adjacent dsDNA segments, preventing full extension under a moderated force. This sequential analysis shows that Hfq binds to dsDNA independently and prior bridging.

Finally, in order to further understand Hfq binding mode, DNA compaction was also investigated through length-versus-time measurements (Fig. 5B). Starting from the compact state at $\sigma = 0$ (center of the green line in Fig 5A), the force was increased from 0.1 pN to 2.2 pN. A continuous lengthening of the filament was observed up to a plateau at 2.2 µm. The force was then lowered to 1 pN, a force at which the extension of the molecule remained constant about 2 µm. The force was then further decreased to 0.6 pN, leading to the relaxation of the fiber towards a partially compact state with a length of 1 µm. Such a dynamic and reversible transition between compact and extended states can be interpreted as the breakage/formation of interactions between Hfq proteins linked to the dsDNA molecule. Hfq-Hfq binding is likely to be responsible for the stabilization of this compact state.

As SRCD also allows to analyse DNA local conformation[49], we investigated how full length Hfq affects $(dA:dT)_{59}$ conformations (Figure 6). A-tract was chosen as a model because Hfq

preferentially interacts with deoxyadenosine rich sequences[12]. As expected, the spectrum of the (dA:dT) duplex contains positive bands at ~ 216 nm and ~ 190 nm, with a shoulder at ~ 180 nm, and a negative band at ~ 205 nm (Figure 6). The unusual positive band at 190 nm is characteristic of A-tract conformation, which differs from the canonical B conformation[50]. In agreement with our previous report, the smaller bands at 180 and 190 nm in the presence of Hfq may represent base-pairing disruption and a local change in helix conformation[12]. Note that such a change in helix conformation may result in a change in DNA topology only if large regions of DNA are covered by Hfq, which is unlikely *in vivo*[13-15] (especially because Dps is the major DNA binding protein during stationary phase[51]).

In order to go further in the mechanisms of Hfq-mediated DNA conformational change and the possible role of its CTR, we analysed the effect of Hfq on plasmid supercoiling *in vivo*. As Hfq is more abundant during the stationary phase, we focused our analysis on late exponential and stationary phases[15,37]. These phases were also those analysed in previous reports[2]. Note that it is difficult to compare WT and Δ*hfq* strains in the same region of exponential growth phase as both strains grow in a very different way. Topology gels of pHSG298 plasmids extracted from different strains (WT, *hfq72*=ΔCTR and Δ*hfq*) are shown on Figure 7. Plasmid samples were electrophoresed on a native or chloroquine agarose gels to separate topoisomers. Under native electrophoresis conditions (Figure 7 left), the more supercoiled they were, the quicker topoisomers migration was. In contrast to native gels, under the presence of chloroquine (Figure 7 right), the most supercoiled topoisomers migrated slowly. From this analysis we can conclude that plasmids are usually more relaxed (= less negatively supercoiled) at 24h (stationary phase) than at 6h (late exponential phase). Nevertheless, we also observed that there is no difference between WT and *hfq72* strains, indicating that the strong compaction due to the CTR of Hfq does not affect plasmid topology. On the other hand, we confirm that plasmids are more relaxed in Δ*hfq* strain, as previously observed[2].

Hfq compaction influences plasmid replication efficiency

We then compared the effect on plasmid compaction and supercoiling to replication efficiency in the same phase of growth. As expected, our analysis confirms a greater number of plasmids in late exponential and stationary phases, and in the total absence of Hfq (Figure 8)[37]. Under these conditions, *i.e.* total absence of Hfq, plasmids are not Hfq-compacted and hence are more relaxed. A possibility is that replication is more efficient due to these

differences in DNA topology and compaction. In order to distinguish between these possibilities, we analysed the effect of Hfq-CTR deletion on replication efficiency. Total suppression of Hfq results in a difference for both supercoiling and compaction, while deletion of its CTR changes compaction only (no effect is observed on supercoiling, see above). As seen on figure 8, no significant difference is observed between WT and *hfq72* strains. Thus, our results indicate that the effect on replication efficiency is directly related to the change in DNA supercoiling, but independent of DNA compaction. If replication efficiency were dependent on compaction, deletion of Hfq-CTR in *hfq72* should also change replication efficiency.

Discussion

In this manuscript we establish for the first time that Hfq self-assembly on DNA, responsible of DNA bridging and compaction, is due to the amyloid-like structure of its CTR. Furthermore we show that DNA induces amyloidogenesis of this region of the protein. In contrast to what has been observed in another extensively studied bacterial protein, RepA[52], in the case of Hfq, DNA seems to be a constituent of the fibril. This conclusion is based on following observations: *(i)* peptide hydrogens are entrapped inside the complex with DNA and cannot exchange with deuterium (Fig. 2); *(ii)* DNA is found in the pellet upon centrifugation of fibrils prepared with DNA; *(iii)* this is confirmed with electrophoretic analysis where DNA is retained in the well in the presence of Hfq-CTR (see Malabirade *et al*[24]), consistent with a stable fiber–DNA interaction; *(iv)* as previously observed, the CTR peptide surrounds DNA molecule with a succession of covered and naked DNA regions[13,24], indicating the self-assembled CTR is bound to DNA; *(v)* increase of inter-sheet and inter-strand spacing in the presence of DNA (Fig. 3); *(vi)* DNA is partially protected and FTIR amyloid signal is reduced after DNAse I enzymatic digestion of the complex, strongly indicating that DNA is a component of the fiber. This contrasts with the case of RepA, where DNA was essentially in the soluble fraction after centrifugation, where the complex in native gels showed no retarded DNA in the well and where DNase I efficiently degraded most of DNA and does not alter the fibers[52]. Strikingly, RepA, has also been shown to bridge together two plasmid DNA molecules through an amyloid oligomeric structure, relevant to control plasmid copy number *in vivo*[53]. These two distinct proteins thus affect plasmid biology using a closely related mechanism (amyloidogenesis) but through direct or indirect pathways.

This new property opens perspectives about how Hfq, in addition to its role in RNA

metabolism, binds DNA and may play a role in organising the bacterial chromosome by inducing a mechanical modification of the nucleic acid. Taking into account that a *E. coli* cell contains approximately 10 000 Hfq hexamers[54]), ~20% of which are found in the nucleoid[10,15], the high amount of this protein gives a fairly good potential for DNA-bound domains. Hfq was previously shown to interact with DNA *via* its distal surface[11], which is dedicated to A-rich sequence binding[20]; thus, one can speculate that the protein first binds and nucleates on A-containing sequences *via* its distal face[12] (for instance in the A/T-rich region of some consensus sequences[11]) and then propagates to surrounding sequences. Consequently, Hfq may cover up quite large regions of DNA due to binding to the nucleic acid that generates a fiber-like pattern[12,24]. This is supported by the previous observation that Hfq self-assembles into filaments *in vitro*[31,33] and presumably *in vivo*, as suggested by the fiber-like structures seen in the nucleoid by using TEM cellular imaging[13,15].

In this manuscript, we demonstrate for the first time that DNA itself promotes Hfq self-assembly into an amyloid-like structure. Self-assembly on DNA may thus have important consequences on the nucleic acid structure. One can expect effects on DNA compaction and supercoiling, and both have been reported earlier[2,23,24]. In this manuscript, we clarify the precise effect of Hfq on DNA compaction and topology. If compaction has been clearly assigned to direct Hfq DNA binding *in vitro*, and precisely to its CTR[23,24], the effect of Hfq on supercoiling observed *in vivo* may be indirect. Indeed, Hfq regulates the expression of many proteins, including some affecting DNA topology[8,55] (Figure 9). Here we clearly establish that if Hfq compacts DNA thanks to its CTR amyloid region[23], and if its CTR induces bridging of double stranded DNA distant sequences[12], these processes do not result in a change in DNA supercoiling. Topological effects observed previously in Δ*hfq* strain are thus indirect: Hfq by itself does not change DNA twisting significantly and its impact on topology observed *in vivo* is likely due the riboregulation of another protein, such as H-NS or Dps [56,57] (figure 9). Note that the bridging property reported for Hfq is parallel to H-NS self-assembly and bridging capacity[58], the underlying mechanism of which was largely solved by single-molecule techniques[59-62]. H-NS uses a coiled-coil motif for its self-assembly[58], while Hfq uses an amyloid-like structure. Hfq thus enlarges the bridging family of nucleoid-associated proteins[63], even if in this case bridging results from the formation of an amyloid-like structure induced by DNA and does not result in an effect on overall DNA topology, in contrast with H-NS[45,55,64]. Note also that similarly to H-NS, the link between DNA-bridging and change in topology may be sequence-dependent[65]. Finally, both proteins

have very similar sequence specificity[11,66] and the way they could physically[67] or functionally[56] interact should be investigated further (Figure 9).

Although the relation between DNA bridging, supercoiling and genetic regulation is still the subject of intense speculation[68,69], our study provides new insights into the role of Hfq for replication efficiency[13,37]. The majority of plasmids replicate less efficiently in the relaxed form, as supercoiling is required to facilitate opening of the origin. However, in our case we used a ColE1-like plasmid[70], for which replication depends on transcription and formation of the RNAI-RNAII complex[71,72]. We hypothesise that the effect observed for Hfq on replication efficiency is mainly due to its (indirect) role in DNA supercoiling and precisely, in its role in primer (RNAII) transcription (Figure 9). Nevertheless, additional effects of Hfq on RNAI-RNAII annealing could also influence replication efficiency, besides its effect on primer transcription[73]. In addition, Hfq may also play a role in the segregation of plasmids, by reducing the linking number during elongation of DNA synthesis, similarly to DNA gyrase[74].

In conclusion, although work remains to be done to decipher Hfq precise mechanism at stake, our results definitely demonstrate Hfq, a riboregulator, is an important player in bacterial chromosome structure, directly or *via* sRNA based regulations (Figure 9). Furthermore, as pathologic amyloids have also gene-regulatory functions[75-77], our work additionally opens new perspectives on the interaction of these amyloids with DNA in eukaryotic cells.

MATERIALS AND METHODS

Expression and purification of Hfq proteins

Full length and truncated Hfq (residues 1-72) were expressed and prepared as described previously[24,36]. Hfq-CTR was chemically synthesised (Proteogenix, France)[33].

Preparation of the complexes for SRCD, FTIR and SAXS

Complexes between Hfq-CTR and $(dA:dT)_{59}$ duplex (Eurogentech) were prepared in water and used at a final concentration of 1.8 mM and 7.3 mM, respectively. The stoichiometry was 1 Hfq-CTR per 4 base pair. For full length HfqWT and $(dA:dT)_{59,}$ 10 µM of Hfq were mixed with 2.23 mM of $(dA:dT)_{59}$, giving a stoichiometry of 1 Hfq hexamer per 220 base pairs. A lower Hfq/bp ratio was chosen to minimize the signal of the protein sample in order

to focus on DNA structural change. Samples were analysed at specific times. Note that when the complex is formed in the presence of salts and buffers, the spectral bandwidth accessible was limited, reducing the spectral information content. Nevertheless, comparison in the UV spectral range of complexes in water and salt did not show any spectral differences. Furthermore addition of Na^+ at 50 and 250 mM did not allow to detect any effect on peptide self-assembly kinetics at the concentration used for our analysis, ruling out the possibility that the effect observed could be due to traces of salts. Besides, addition of salt may result in complex disassembly[78].

Synchrotron Radiation Circular Dichroism (SRDC)

For SRCD analysis, measurements and data collection were carried out on DISCO beam-line at the SOLEIL Synchrotron (proposal 20171061)[79]. After different incubation time, 2-4 µl of samples were loaded into circular demountable CaF_2 cells of 33 microns path length[80]. Three separated data collections with fresh sample preparations were carried out to ensure consistency and repeatability. Spectral acquisitions of 1 nm steps at 1.2 integration time, between 320 and 170 nm were performed in triplicate for the samples as well as for the baselines. (+)-camphor-10-sulfonic acid (CSA) was used to calibrate amplitudes and wavelength positions of the SRCD experiment. Data-analyses including averaging, baseline subtraction, smoothing, scaling and standardization were carried out with CDtool.

Fourier Transform Infrared spectroscopy (FTIR)

For FTIR analysis, the same solutions used for SRCD analysis were lyophilized and re-dissolved in D_2O (5 µL). FTIR spectra were recorded using a Bruker Tensor 27 spectrophotometer. For transmission experiments, samples were deposited between two ZnSe windows without spacer. 30 scans were usually accumulated under continuous dry air purge, with a resolution of 1 cm^{-1}. Data treatment was performed using the OPUS software (Bruker) and consisted of multiple point base line correction and spectral subtraction after normalization.

Small Angle X-ray Scattering (SAXS)

SAXS measurements were performed on the high brilliance SWING beam line at the Soleil synchrotron facility, with monochromator set at 12 KeV (proposal 20170388) using an EIGERX-4M detector at 0.560 m from the sample[81]; diffraction patterns were recorded for reciprocal spacing $q = 4\ \pi \sin(\theta)/\lambda$ varying between 0.02 and 1.72 $Å^{-1}$, that is, repetitive

distances d = 2 π/q ranging from 3.14 to 3.65 Å. SAXS pattern of samples and references were recorded in previously filled Ø1.5 mm quartz capillaries. 10 patterns were acquired successively for each sample, thus exposure time was tuned in order to avoid radiation damage. 1D SAXS curves were obtained by circular averaging of the 2D images using Foxtrot software (http://www.synchrotron-soleil.fr/Recherche/LignesLumiere/SWING). As no difference was observed between the 10 images recorded for each samples, the 1D curves were averaged in order to obtain 1 curve for each samples and references with better statistics.

Construction of E. coli strains

Strains were constructed with the λ-red recombination technique[82]. Briefly, the *E. coli* MG1655 strain was transformed and grown with the temperature sensitive pKD46 plasmid containing Red recombinase expression genes, in presence of arabinose. PCR fragments containing the desired mutation plus an antibiotic resistance cassette were electropored to allow recombination. Bacteria were selected for antibiotic resistance and the introduction of the mutation was checked by PCR and sequencing.

First, a reference strain (MG1655 *hfq*:*cat*) was obtained by inserting a chloramphenicol resistance cassette downstream the *hfq* ORF of MG1655, using the pKD3 plasmid as PCR template. Primers sequences were 5'-AATACTTCCGCGCAACAGGACAGCGAAGAAACCGAATAAGGTTCCATGGTCCATATGAATATC-3' (forward) and 5'-AGGATCGCTGGCTCCCCGTGTAAAAAAACAGCCCGAAACCTGTGTAGGCTGGAGCTGCTT-3' (reverse). Then, a second strain deleted for the last 90 *hfq* nucleotides giving a truncated protein with only the first 72 amino-acids (MG1655 *hfq72*:*cat*) was obtained by the same technique but using the previous strain as a template for amplifying the DNA fragment. Recombination was performed in the MG1655 strain. Primers used were 5'-ATTTCTACTGTTGTCCCGTCTCGCCCGGTTTCTCATCACAGTTAAGGTTCCATGGTCCATATGAATATC-3'(forward) and 5'CGCTGGCTCCCCGTGTAAAAAAACAGCCCGAAACC-3' (reverse).

In vitro analysis of plasmid topology

The experiments were performed as described in Tupper *et al*[45]. 5 µg of purified pHSG298 (2675bp, Kan^R, Takara Bio) plasmid was incubated 10 min at room temperature with full

length Hfq in a total volume of 100 µL of buffer containing 10 mM Tris-HCl pH 7.5, 0.1 mM EDTA, 15 mM KCl, 2 mM spermidine, 15% v/v glycerol and 0.1 $g.L^{-1}$ BSA. 18U of calf thymus topoisomerase I (Invitrogen) were then added and the incubation continued for 30 min at 37°C. Proteins were extracted with ultrapure phenol/chloroform/isoamyl alcohol (25:24:1 v/v, Invitrogen) and ethanol precipitation. The DNA was re-suspended in 50 µL Tris-HCl (5 mM, pH8). 2 µL of purified plasmids were electrophoresed in a 1% w/v native agarose gel. Electrophoresis was performed in Tris-Borate-EDTA 1X, 16.5 V/cm during 30min. The gel was then rinsed 15 min with water before staining 30 min with GelRed fluorescent dye (Biotium, Inc.). Images were acquired with a G-box (Syngene) imager.

In vivo analysis of plasmid topology

Strains transformed with pHSG298 were grown overnight in LB with kanamycin at 37°C with good aeration and shaking. The next day, precultures were refreshed by 100X dilution in 50 mL cultures and grown in 250 mL Erlenmeyers flasks at 120 rpm. At 6h and 24h of cultivation time, pHSG298 was purified from 5 mL culture samples with NucleoSpin Plasmid EasyPure extraction kit (Macherey Nagel), and eluted in Tris-HCl (5 mM, pH8) buffer. Each plasmid purification method may have potential biases concerning DNA topology selection, but we preferred to use a kit based on silica column to compare our results with those of Tsui *et al*[2]. In addition, all samples were processed the same way and the only difference was the genetic background. Around 0.3 µg of purified plasmids were electrophoresed in 1% w/v agarose gels in the presence or the absence of 10 $\mu g.mL^{-1}$ chloroquine to separate the plasmids of different linking numbers (Lk). Electrophoresis was performed in Tris-Borate-EDTA 1X, 2.5 V/cm during 16h at 8°C. The gel was then incubated 1h in 10 mM $MgSO_4$ to remove excess of chloroquine (when appropriate), and otherwise stained and observed as above.

Magnetic tweezers experiments

A poly-di-methylsiloxane (PDMS ; Dow-Corning) flow cell with a 2-mm-wide and 100-µm-high channel was mounted on a glass coverslip previously silanized with Sigmacote (Sigma). Anti-digoxigenin (Roche) was flushed into the channel and bound non-specifically for 1h at 37°C, followed by overnight blocking with BSA. The PDMS flow cell was placed beneath two permanent NdFeB magnets. Image acquisition utilized a 60 Hz camera (A10 GigE; JAI) placed on an inverted microscope (Picotwist). Image analysis was performed with Pico software. The 8.8kb DNA molecule held in magnetic tweezers setup was prepared by PCR

from template lambda-DNA, it was then ligated at each end with a 600-bp DNA fragments prepared by PCR from template Litmus28i (NEB; positions 2008 and 2580), one was modified with biotin-tagged dUTP and the other with digoxigenin-tagged dUTP (Roche Diagnostics). This DNA fragment was subsequently attached to a magnetic bead coated with streptavidin prior injection of this construct into the microchannel and attachment to the antidigoxigenin-coated surface. Experiments were carried in buffer A (10mM Tris-HCl pH 7, 1mM EDTA, 0.1 mg/mL BSA). Hfq concentration was chosen as follow: we know that a ratio ~50 monomeric Hfq/DNA molecule allows having a significant coverage of DNA molecule[12,13]. Similarly this assay is done with highly diluted, single DNA molecules to provide information about intramolecular compaction. We previously found that compaction requires an order of magnitude lower of the CTR fragment than the one pertaining to 50% binding saturation[24]. For this reason, we used 60 nM of Hfq for our analysis, corresponding to ~50 Hfq in the volume explored around a 2 µm-long-DNA attached to the glass coverslip.

qPCR plasmid quantification

The different strains transformed with pHSG298 were grown overnight in LB supplemented with kanamycin at 37°C with good aeration and shaking. The next day, precultures were refreshed by 100X dilution in 50 mL cultures and grown in 250 mL Erlenmeyer flaks at 120 RPM. At 6h and 24h of cultivation time, 3 mL samples for each strain were centrifuged at 5000 g. The pellet was rinsed twice with PBS and re-suspended in 0.5 mL PBS 1X. Samples were lysed at 95°C during 5 min, and centrifuged again to remove cellular debris before being subjected to qPCR.

For the quantifications, samples were diluted and mixed with Brilliant III Ultra-Fast SYBR Green qPCR Master Mix (Agilent Technologies) and 300µM primers according to the manufacturer recommendations. Oligonucleotides targeting regions in *mreB* gene (chromosome) and *kan*^{r} (plasmid) DNA sequences, plasmids harbouring target sequences used at known copy numbers for calibration curves and qPCR protocol are described elsewhere[9]. Quantities inferred from Cp were determined and averaged from 8 replicates for the *kan*^{r} gene and were divided by the quantities obtained for *mreB*, giving the number of pHSG298 copies per cell.

Acknowledgements: We are very grateful to N. and L. Bossi for their help with the construction of *hfq* mutant cells, to J.L. Viovy for providing the magnetic tweezers infrastructure and fruitful discussions and to C. Lavelle, V. Militello, C. Sandt, R. Sinden, K. Kubiak and G. Wegrzyn for insightful discussions. SRCD and SAXS measurements on DISCO and SWING beamlines at the SOLEIL Synchrotron were performed under proposals 20171061 # and #20170388.

Funding

This work was supported by CNRS, CEA and synchrotron SOLEIL.

Author Contributions

AM, DP, FW and VA conceived the original idea and designed the experiments. AM and DP purified proteins. AM made topology gels. AM constructed *hfq* mutant strains. AM, DP, OEH, FT, VA and FW performed SRCD experiments. DP, OEH, VA, TB and FW performed SAXS experiments. DP and FG performed FTIR experiments. PR performed magnetic tweezers experiments. AM, FT and FB made qPCR measurements. VA drew figure 9. VA wrote the manuscript. All authors analysed results and commented on the manuscript.

Competing Interests:

The authors declare that they have no competing interests.

References

1 Franze de Fernandez, M. T., Hayward, W. S. & August, J. T. Bacterial proteins required for replication of phage Qb ribonucleic acid. *J. Biol. Chem.* **247**, 824-821 (1972).

2 Tsui, H. C., Leung, H. C. & Winkler, M. E. Characterization of broadly pleiotropic phenotypes caused by an *hfq* insertion mutation in *Escherichia coli* K-12. *Mol. Microbiol.* **13**, 35-49 (1994).

3 Tsui, H. C., Feng, G. & Winkler, M. E. Negative regulation of *mutS* and *mutH* repair gene expression by the Hfq and RpoS global regulators of *Escherichia coli* K-12. *J. Bacteriol.* **179**, 7476-7487 (1997).

4 Gottesman, S. & Storz, G. RNA reflections: converging on Hfq. *RNA* **21**, 511-512 (2015).

5 Gottesman, S. *et al.* Small RNA regulators and the bacterial response to stress. *Cold Spring Harb Symp Quant Biol* **71**, 1-11 (2006).

6 Aiba, H. Mechanism of RNA silencing by Hfq-binding small RNAs. *Curr Opin Microbiol* **10**, 134-139 (2007).

7 Majdalani, N., Cunning, C., Sledjeski, D., Elliott, T. & Gottesman, S. DsrA RNA regulates translation of RpoS message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription. *Proc Natl Acad Sci U S A* **95**, 12462-12467 (1998).

8 Lease, R. A. & Belfort, M. A trans-acting RNA as a control switch in Escherichia coli: DsrA modulates function by forming alternative structures. *Proc Natl Acad Sci U S A* **97**, 9919-9924 (2000).

9 Cayrol, B. *et al.* Riboregulation of the bacterial actin-homolog MreB by DsrA small noncoding RNA. *Integrative biology : quantitative biosciences from nano to macro* **7**, 128-141 (2015).

10 Azam, T. A. & Ishihama, A. Twelve species of the nucleoid-associated protein from Escherichia coli. Sequence recognition specificity and DNA binding affinity. *J Biol Chem* **274**, 33105-33113 (1999).

11 Updegrove, T. B., Correia, J. J., Galletto, R., Bujalowski, W. & Wartell, R. M. E. coli DNA associated with isolated Hfq interacts with Hfq's distal surface and C-terminal domain. *Biochim Biophys Acta* **1799**, 588-596 (2010).

12 Geinguenaud, F. *et al.* Conformational transition of DNA bound to Hfq probed by infrared spectroscopy. *Phys Chem Chem Phys* **13**, 1222-1229 (2011).

13 Cech, G. M. *et al.* The Escherichia Coli Hfq Protein: An Unattended DNA-Transactions Regulator. *Front Mol Biosci* **3**, 36, doi:10.3389/fmolb.2016.00036 (2016).

14 Azam, T. A., Hiraga, S. & Ishihama, A. Two types of localization of the DNA-binding proteins within the Escherichia coli nucleoid. *Genes Cells* **5**, 613-626 (2000).

15 Diestra, E., Cayrol, B., Arluison, V. & Risco, C. Cellular electron microscopy imaging reveals the localization of the Hfq protein close to the bacterial membrane. *PLoS One* **4**, e8301, doi:10.1371/journal.pone.0008301 (2009).

16 Wilusz, C. J. & Wilusz, J. Eukaryotic Lsm proteins: lessons from bacteria. *Nat Struct Mol Biol* **12**, 1031-1036 (2005).

17 Wilusz, C. J. & Wilusz, J. Lsm proteins and Hfq: Life at the 3' end. *RNA Biol* **10**, 592-601 (2013).

18 Brennan, R. G. & Link, T. M. Hfq structure, function and ligand binding. *Curr Opin Microbiol* **10**, 125-133 (2007).

19 Vogel, J. & Luisi, B. F. Hfq and its constellation of RNA. *Nat Rev Microbiol* **9**, 578-589 (2011).

20 Link, T. M., Valentin-Hansen, P. & Brennan, R. G. Structure of Escherichia coli Hfq bound to polyriboadenylate RNA. *Proc Natl Acad Sci U S A* **106**, 19292-19297 (2009).

21 Schumacher, M. A., Pearson, R. F., Moller, T., Valentin-Hansen, P. & Brennan, R. G. Structures of the pleiotropic translational regulator Hfq and an Hfq- RNA complex: a bacterial Sm-like protein. *Embo J* **21**, 3546-3556. (2002).

22 Dimastrogiovanni, D. *et al.* Recognition of the small regulatory RNA RydC by the bacterial Hfq protein. *eLife* **3**, doi:10.7554/eLife.05375 (2014).

23 Jiang, K. *et al.* Effects of Hfq on the conformation and compaction of DNA. *Nucleic Acids Res* **43**, 4332-4341 (2015).

24 Malabirade, A. *et al.* Compaction and condensation of DNA mediated by the C-terminal domain of Hfq. *Nucleic Acids Res* **45**, 7299-7308 (2017).

25 Joyeux, M. Compaction of bacterial genomic DNA: clarifying the concepts. *J Phys Condens Matter* **27**, 383001, doi:10.1088/0953-8984/27/38/383001 (2015).

26 Dai, L., Mu, Y., Nordenskiold, L. & van der Maarel, J. R. Molecular dynamics simulation of multivalent-ion mediated attraction between DNA molecules. *Phys Rev Lett* **100**, 118301, doi:10.1103/PhysRevLett.100.118301 (2008).

27 Sauter, C., Basquin, J. & Suck, D. Sm-like proteins in Eubacteria: the crystal structure of the Hfq protein from Escherichia coli. *Nucleic Acids Res* **31**, 4091-4098 (2003).

28 Nikulin, A. *et al.* Structure of Pseudomonas aeruginosa Hfq protein. *Acta Crystallogr D Biol Crystallogr* **61**, 141-146 (2005).

29 Sauer, E. & Weichenrieder, O. Structural basis for RNA 3'-end recognition by Hfq. *Proc Natl Acad Sci U S A* **108**, 13065-13070 (2011).

30 Stanek, K. A., Patterson-West, J., Randolph, P. S. & Mura, C. Crystal structure and RNA-binding properties of an Hfq homolog from the deep-branching Aquificae: conservation of the lateral RNA-binding mode. *Acta Crystallogr D Struct Biol* **73**, 294-315 (2017).

31 Arluison, V. *et al.* Three-dimensional Structures of Fibrillar Sm Proteins: Hfq and Other Sm-like Proteins. *J Mol Biol* **356**, 86-96 (2006).

32 Beich-Frandsen, M. *et al.* Structural insights into the dynamics and function of the C-terminus of the E. coli RNA chaperone Hfq. *Nucleic Acids Res*, **39**, 4900-4915 (2011).

33 Fortas, E. *et al.* New insight into the structure and function of Hfq C-terminus. *Biosci Rep* **35**, doi:10.1042/BSR20140128 (2015).

34 Partouche, D. *et al.* Techniques to Analyze sRNA Protein Cofactor Self-Assembly In Vitro. *Methods Mol Biol* **1737**, 321-340 (2018).

35 Malabirade, A. *et al.* Membrane association of the bacterial riboregulator Hfq and functional perspectives. *Sci Rep* **7**, 10724, doi:10.1038/s41598-017-11157-5 (2017).

36 Taghbalout, A., Yang, Q. & Arluison, V. The Escherichia coli RNA processing and degradation machinery is compartmentalized within an organized cellular network. *Biochem J* **458**, 11-22 (2014).

37 Cech, G. M. *et al.* Hfq protein deficiency in Escherichia coli affects ColE1-like but not lambda plasmid DNA replication. *Plasmid* **73**, 10-15 (2014).

38 Sukhodolets, M. V. & Garges, S. Interaction of Escherichia coli RNA polymerase with the ribosomal protein S1 and the Sm-like ATPase Hfq. *Biochemistry* **42**, 8022-8034 (2003).

39 Le Derout, J., Boni, I. V., Regnier, P. & Hajnsdorf, E. Hfq affects mRNA levels independently of degradation. *BMC Mol Biol* **11**, 17, doi:1471-2199-11-17 (2010).

40 Rabhi, M. *et al.* The Sm-like RNA chaperone Hfq mediates transcription antitermination at Rho-dependent terminators. *EMBO J* **30**, 2805-2816 (2011).

41 Lindberg, D. J., Wesen, E., Bjorkeroth, J., Rocha, S. & Esbjorner, E. K. Lipid membranes catalyse the fibril formation of the amyloid-beta (1-42) peptide through lipid-

fibril interactions that reinforce secondary pathways. *Biochim Biophys Acta* **1859**, 1921-1929 (2017).

42 Gobeaux, F. & Wien, F. Reversible assembly of a drug peptide into amyloid fibrils: a dynamic circular dichroism study. *Langmuir : the ACS journal of surfaces and colloids*, doi:10.1021/acs.langmuir.8b00094 (2018).

43 Byler, D. M. & Susi, H. Examination of the secondary structure of proteins by deconvolved FTIR spectra. *Biopolymers* **25**, 469-487 (1986).

44 Piccirilli, F. *et al.* Sequential dissociation of insulin amyloids probed by high pressure Fourier transform infrared spectroscopy *Soft mater* **8**, 11863-11870 (2012).

45 Tupper, A. E. *et al.* The chromatin-associated protein H-NS alters DNA topology in vitro. *EMBO J* **13**, 258-268 (1994).

46 Strick, T. R., Allemand, J. F., Bensimon, D., Bensimon, A. & Croquette, V. The elasticity of a single supercoiled DNA molecule. *Science* **271**, 1835-1837 (1996).

47 Marko, J. F. & Siggia, E. D. Stretching DNA. *Macromolecules* **28**, 8759–8770 (1995).

48 Marko, J. F. Torque and dynamics of linking number relaxation in stretched supercoiled DNA. *Phys Rev E Stat Nonlin Soft Matter Phys* **76**, 021926 (2007).

49 Johnson, K. H., Gray, D. M. & Sutherland, J. C. Vacuum UV CD spectra of homopolymer duplexes and triplexes containing A.T or A.U base pairs. *Nucleic Acids Res* **19**, 2275-2280 (1991).

50 Aymami, J., Coll, M., Frederick, C. A., Wang, A. H. & Rich, A. The propeller DNA conformation of poly(dA).poly(dT). *Nucleic Acids Res* **17**, 3229-3245 (1989).

51 Janissen, R. *et al.* Global DNA Compaction in Stationary-Phase Bacteria Does Not Affect Transcription. *Cell* **174**, 1188-1199 e1114, doi:10.1016/j.cell.2018.06.049 (2018).

52 Giraldo, R. Defined DNA sequences promote the assembly of a bacterial protein into distinct amyloid nanostructures. *Proc Natl Acad Sci U S A* **104**, 17388-17393 (2007).

53 Molina-Garcia, L. *et al.* Functional amyloids as inhibitors of plasmid DNA replication. *Sci Rep* **6**, 25425, doi:10.1038/srep25425 (2016).

54 Kajitani, M., Kato, A., Wada, A., Inokuchi, Y. & Ishihama, A. Regulation of the Escherichia coli hfq gene encoding the host factor for phage Q beta. *J Bacteriol* **176**, 531-534 (1994).

55 Hardy, C. D. & Cozzarelli, N. R. A genetic selection for supercoiling mutants of Escherichia coli reveals proteins implicated in chromosome structure. *Mol Microbiol* **57**, 1636-1652 (2005).

56 Lease, R. A., Cusick, M. E. & Belfort, M. Riboregulation in Escherichia coli: DsrA RNA acts by RNA:RNA interactions at multiple loci. *Proc Natl Acad Sci U S A* **95**, 12456-12461 (1998).

57 Lease, R. A., Smith, D., McDonough, K. & Belfort, M. The small noncoding DsrA RNA is an acid resistance regulator in Escherichia coli. *J Bacteriol* **186**, 6179-6185 (2004).

58 Arold, S. T., Leonard, P. G., Parkinson, G. N. & Ladbury, J. E. H-NS forms a superhelical protein scaffold for DNA condensation. *Proc Natl Acad Sci U S A* **107**, 15728-15732 (2010).

59 Dame, R. T., Noom, M. C. & Wuite, G. J. Bacterial chromatin organization by H-NS protein unravelled using dual DNA manipulation. *Nature* **444**, 387-390 (2006).

60 Dame, R. T., Wyman, C. & Goosen, N. H-NS mediated compaction of DNA visualised by atomic force microscopy. *Nucleic Acids Res* **28**, 3504-3510 (2000).

61 Singh, K., Milstein, J. N. & Navarre, W. W. Xenogeneic Silencing and Its Impact on Bacterial Genomes. *Annu Rev Microbiol* **70**, 199-213 (2016).

62 Winardhi, R. S. & Yan, J. Applications of Magnetic Tweezers to Studies of NAPs. *Methods Mol Biol* **1624**, 173-191 (2017).

63 Wiggins, P. A., Dame, R. T., Noom, M. C. & Wuite, G. J. Protein-mediated molecular bridging: a key mechanism in biopolymer organization. *Biophys J* **97**, 1997-2003 (2009).

64 Owen-Hughes, T. A. *et al.* The chromatin-associated protein H-NS interacts with curved DNA to influence DNA topology and gene expression. *Cell* **71**, 255-265 (1992).

65 Japaridze, A. *et al.* Spatial organization of DNA sequences directs the assembly of bacterial chromatin by a nucleoid-associated protein. *J Biol Chem* **292**, 7607-7618 (2017).

66 Gordon, B. R. *et al.* Structural basis for recognition of AT-rich DNA by unrelated xenogeneic silencing proteins. *Proc Natl Acad Sci U S A* **108**, 10690-10695 (2011).

67 Shi, X. & Bennett, G. N. Plasmids bearing hfq and the hns-like gene stpA complement hns mutants in modulating arginine decarboxylase gene expression in Escherichia coli. *J Bacteriol* **176**, 6769-6775 (1994).

68 Scolari, V. F., Bassetti, B., Sclavi, B. & Lagomarsino, M. C. Gene clusters reflecting macrodomain structure respond to nucleoid perturbations. *Mol Biosyst* **7**, 878-888 (2011).

69 Liu, Y., Chen, H., Kenney, L. J. & Yan, J. A divalent switch drives H-NS/DNA-binding conformations between stiffening and bridging modes. *Genes Dev* **24**, 339-344 (2010).

70 del Solar, G., Giraldo, R., Ruiz-Echevarria, M. J., Espinosa, M. & Diaz-Orejas, R. Replication and control of circular bacterial plasmids. *Microbiol Mol Biol Rev* **62**, 434-464 (1998).

71 Brantl, S. Plasmid Replication Control by Antisense RNAs. *Microbiol Spectr* **2**, PLAS-0001-2013, doi:10.1128/microbiolspec.PLAS-0001-2013 (2014).

72 Di Primo, C. Real time analysis of the RNAI-RNAII-Rop complex by surface plasmon resonance: from a decaying surface to a standard kinetic analysis. *J Mol Recognit* **21**, 37-45 (2008).

73 Brantl, S. *Encyclopedia of Microbiology* Vol. 5 (ed M. Schaechter) 409-437 (Academic Press, 2009).

74 Zechiedrich, E. L. & Cozzarelli, N. R. Roles of topoisomerase IV and DNA gyrase in DNA unlinking during replication in Escherichia coli. *Genes Dev* **9**, 2859-2869 (1995).

75 Goers, J. *et al.* Nuclear localization of alpha-synuclein and its interaction with histones. *Biochemistry* **42**, 8465-8471 (2003).

76 Siddiqui, A. *et al.* Selective binding of nuclear alpha-synuclein to the PGC1alpha promoter under conditions of oxidative stress may contribute to losses in mitochondrial function: implications for Parkinson's disease. *Free Radic Biol Med* **53**, 993-1003 (2012).

77 Yu, H., Ren, J. & Qu, X. Time-dependent DNA condensation induced by amyloid beta-peptide. *Biophys J* **92**, 185-191 (2007).

78 Hopkins, J. F., Panja, S., McNeil, S. A. & Woodson, S. A. Effect of salt and RNA structure on annealing and strand displacement by Hfq. *Nucleic Acids Res* **37**, 6205-6213 (2009).

79 Refregiers, M. *et al.* DISCO synchrotron-radiation circular-dichroism endstation at SOLEIL. *J Synchrotron Radiat* **19**, 831-835 (2012).

80 Wien, F. & Wallace, B. A. Calcium fluoride micro cells for synchrotron radiation circular dichroism spectroscopy. *Appl Spectrosc* **59**, 1109-1113 (2005).

81 David, G. & Perez, J. Perez, Combined sampler robot and high-performance liquid chromatography: a fully automated system for biological small-angle X-ray scattering experiments at the Synchrotron SOLEIL SWING beamline. *Journal of Applied Crystallography* **42**, 892-900 (2009).

82 Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**, 6640-6645 (2000).

83 Militello, V. *et al.* Aggregation kinetics of bovine serum albumin studied by FTIR spectroscopy and light scattering. *Biophys Chem* **107**, 175-187 (2004).

84 Steely, H. T., Jr., Gray, D. M. & Ratliff, R. L. CD of homopolymer DNA-RNA hybrid duplexes and triplexes containing A-T or A-U base pairs. *Nucleic Acids Res* **14**, 10071-10090 (1986).

Legends to the figures

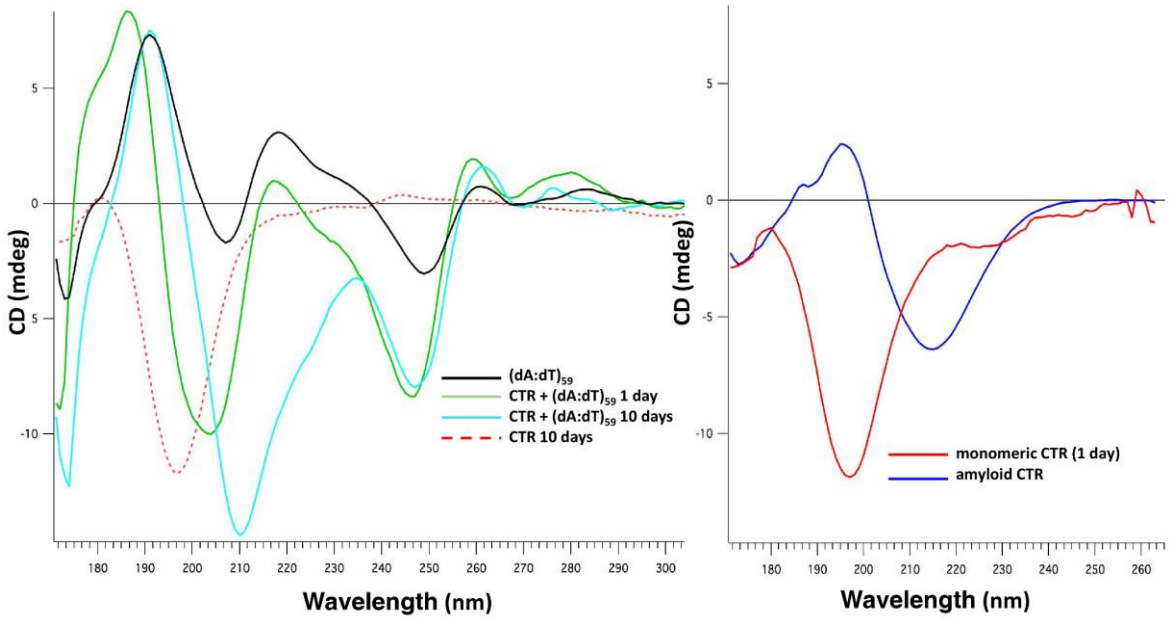


Figure 1. SRCD kinetics of Hfq-CTR self-assembly in the presence of DNA. SRCD experiment was performed as a function of time in order to observe progressive amyloid-like structuring. ***(A)*** $(dA:dT)_{59}$: black line; CTR + $(dA:dT)_{59}$ 1 day: green line; CTR + $(dA:dT)_{59}$ 10 days: cyan line; monomeric CTR 10 days: dashed red line monomeric. ***(B)*** Controls: CTR 1 day: red line; amyloid CTR: blue line. As shown, the typical signal of amyloids at 210-220 nm appears after 7 days, while it takes more than 2 weeks with peptide alone. Note that DNA induces a structural change of the protein and conversely (see figure 6). For this reason we do not substract the spectrum of the peptide alone from the spectrum of DNA:peptide complex.

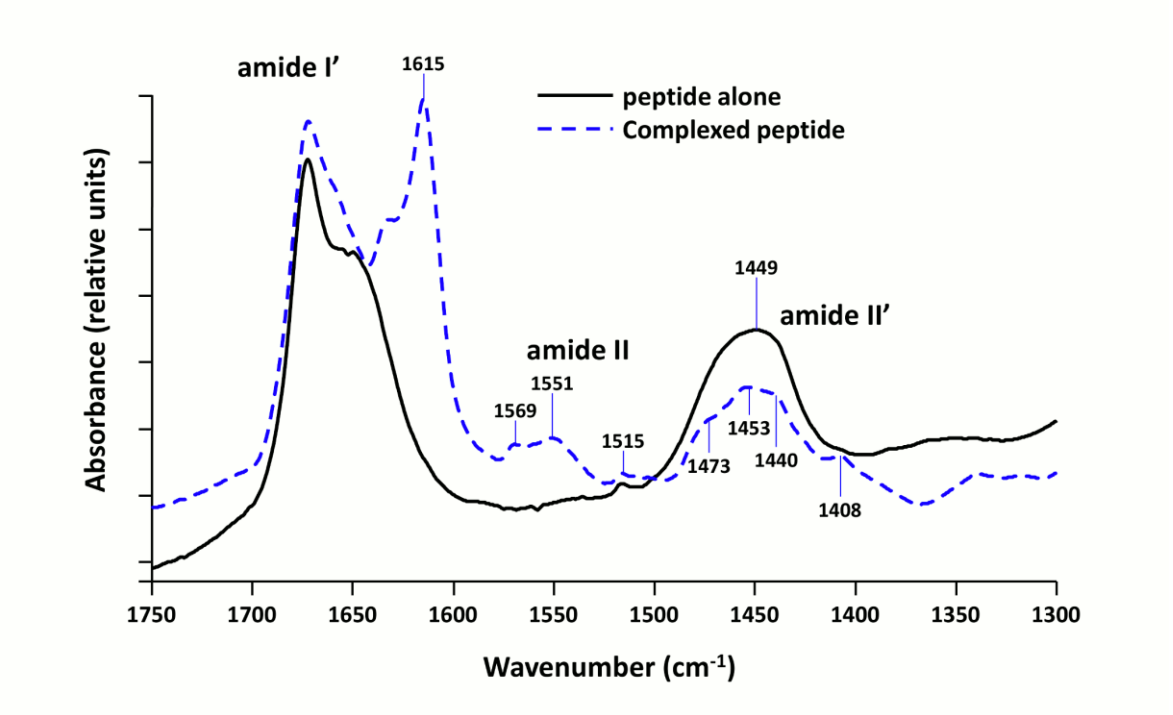


Figure 2. FTIR spectra of Hfq-CTR in the presence or absence of DNA. D_2O buffered samples have been used to avoid the spectral overlaps between Amide I and water bands. Solid black line: peptide alone. Dashed blue line: difference spectrum obtained by subtracting the $(dA:dT)_{59}$ contribution from the complex spectrum. We clearly observe in the amide I' band a contribution at 1615 cm^{-1}, indicative of the formation of the amyloid structure in the presence of DNA. Note in the complex the presence of both Amide II and II′ bands, due to N\H and N\D in-plane bending of the peptidic groups, respectively. Strikingly, the Amide II is present only in the complex, indicating that hydrogens may be entrapped inside the complex and cannot exchange with deuterium[83].

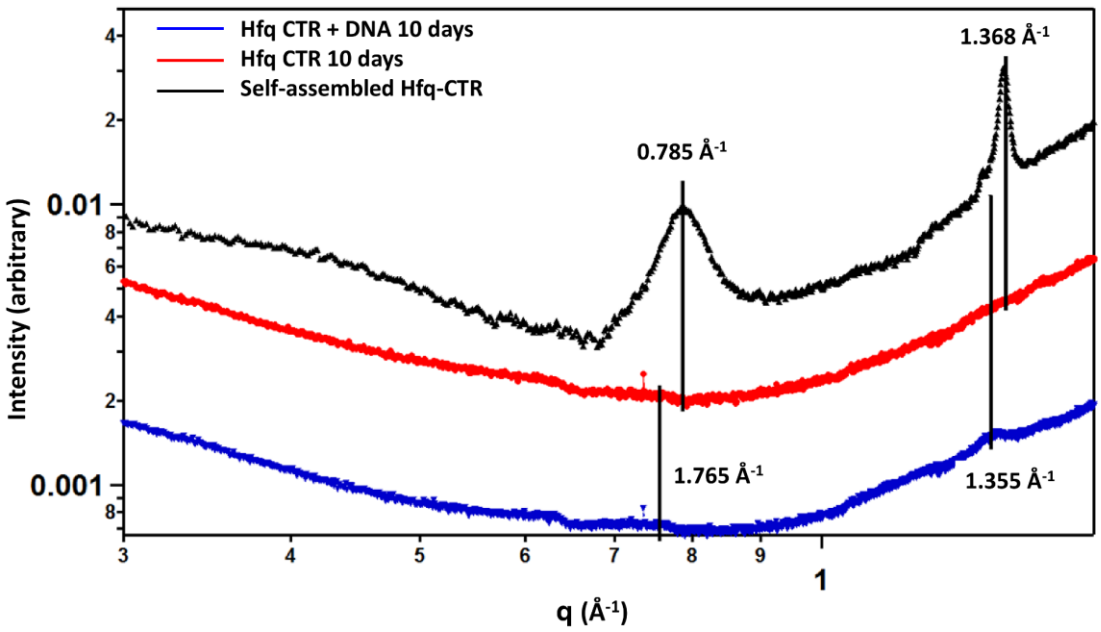


Figure 3: SAXS curves of Hfq-CTR in the presence or absence of DNA. β-Sheet peak corresponding to a cross-β structure are clearly observed for Hfq-CTR in the presence of DNA (incubated 10 days, blue curve) and for Hfq-CTR alone incubated for 6 weeks at 20 mg/mL (higher time and concentration are used to ensure the formation of amyloid fibrils in the absence of cofactor, black curve). But no β-Sheet peak is observed for Hfq-CTR alone after 10 days (red curve). Peaks corresponding to cross-β structure can be seen at 0.765 and 1.355 $Å^{-1}$ for complexed Hfq-CTR, and 0.785 and 1.368 $Å^{-1}$ for non-complexed Hfq-CTR, corresponding to inter-sheet and inter-strand distances (d) of 8.21 and 4.63 and 8.00 Å and 4.59 Å, respectively ($d(Å)=2\pi/q(Å^{-1})$).

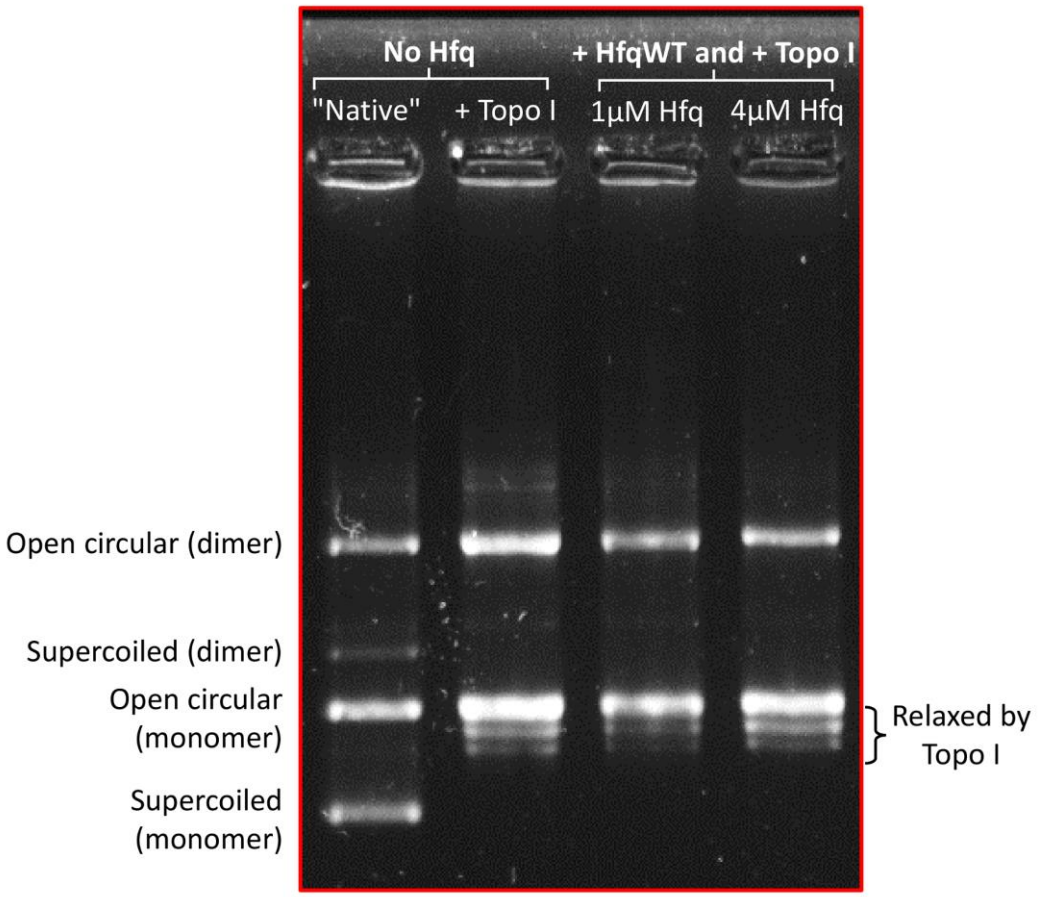


Figure 4: Topology gel of pHSG298 after incubation with Hfq. The presence of purified Hfq does not influence the repartition of topoisomers, confirming that Hfq has no direct effect on overall plasmid topology.

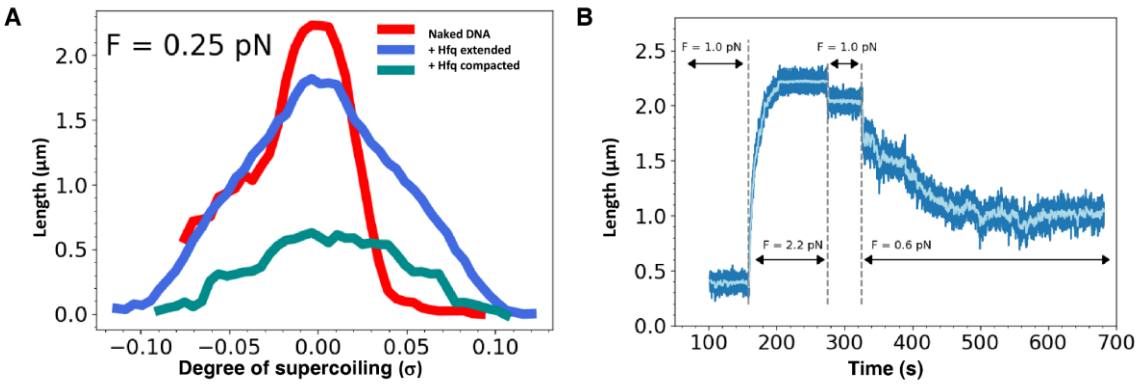


Figure 5. Real-time Hfq-DNA binding followed by magnetic tweezers. The absence of direct effect of Hfq has also been confirmed by tweezers experiments. Note that in this case we used a linear DNA. (A) Extension-versus-rotation curves at 0.25 pN of naked DNA (red), DNA associated with Hfq (present at a concentration of 60 nM) in extended (blue) and compact states (cyan). Compact conformation is accessible through application of supercoiling larger than +/-0.1, or after letting the molecule at low force (below 0.05 pN) for a few minutes. Asymmetry in the response of naked DNA is due to denaturation of the double helix under negative supercoiling. Hfq interactions prevent this denaturation. (B) Change of the length of DNA interacting with 60nM of Hfq submitted to different stretching forces (light blue, running average over a 5s window). At t=100s the molecule is compacted under a force of 1.0 +/-0.1 pN and σ=0. While the force is increased to 2.2 +/-0.1 pN DNA undergoes a continuous lengthening until it reaches a plateau at 2.2 µm. Force was then further decreased, when set to 0.6 +/-0.07 pN molecule relaxed towards a partially compact state with a length of 1.0 µm. As expected, truncated Hfq in the same experiment does not have any effect.

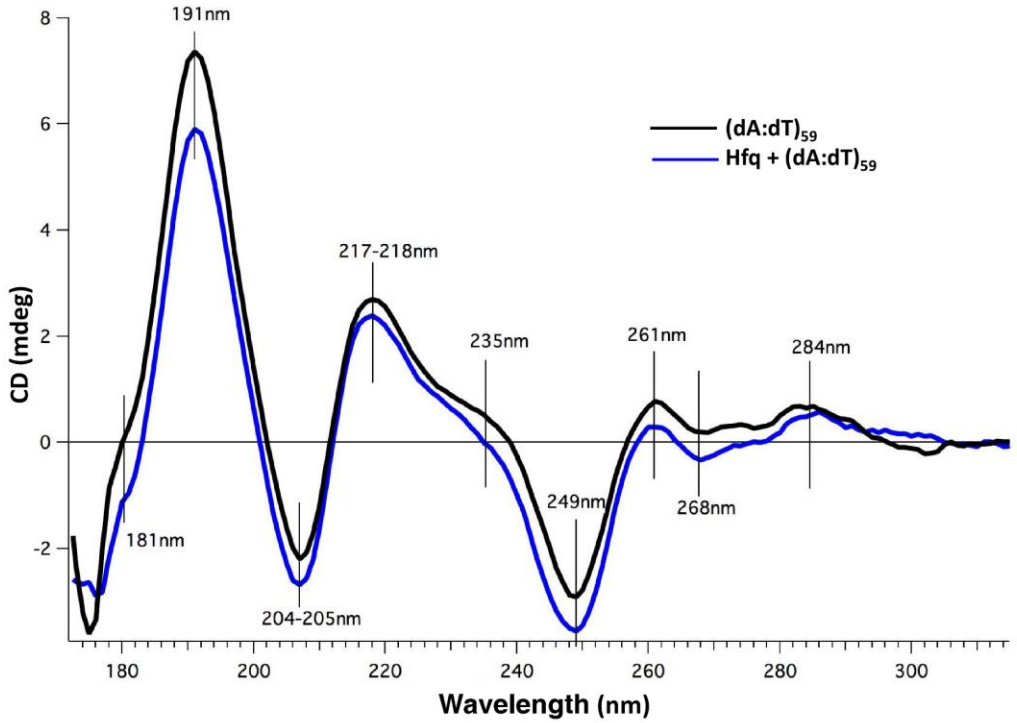


Figure 6: SRCD spectra of DNA interacting with full length Hfq. $(dA:dT)_{59}$ with Hfq (blue line) or without Hfq (black line). The respective contribution of buffer and protein were subtracted. As shown in Figure 1, spectral changes indicate that DNA induces a structural change of the protein and conversely. A low Hfq/bp ratio was chosen in order to minimize the signal of the protein sample and to focus on DNA structural change. As previously described[49], the spectrum of the $(dA:dT)_{59}$ duplex contains positive bands at ~ 217 nm and ~ 191 nm (with a shoulder at ~ 181 nm), and a negative band at ~ 205 nm. Spectral band differences for DNA with and without CTR at 181 and 191 nm (charge transfer and pi-pi*, respectively) correspond to local changes in DNA helix conformation. Further spectral differences include a shoulder at 235 nm and amplitude differences at 249, 261, 268 and 284 nm; the later three positions corresponding to well identified peaks in DNA CD spectroscopy reflecting the base pairing and stacking[49,84]. Smaller magnitudes reflect a weaker degree of base stacking[84]. Note that the amplitude differences are not proportional, which indicates that structural changes are occurring upon interaction between DNA and the CTR of Hfq.

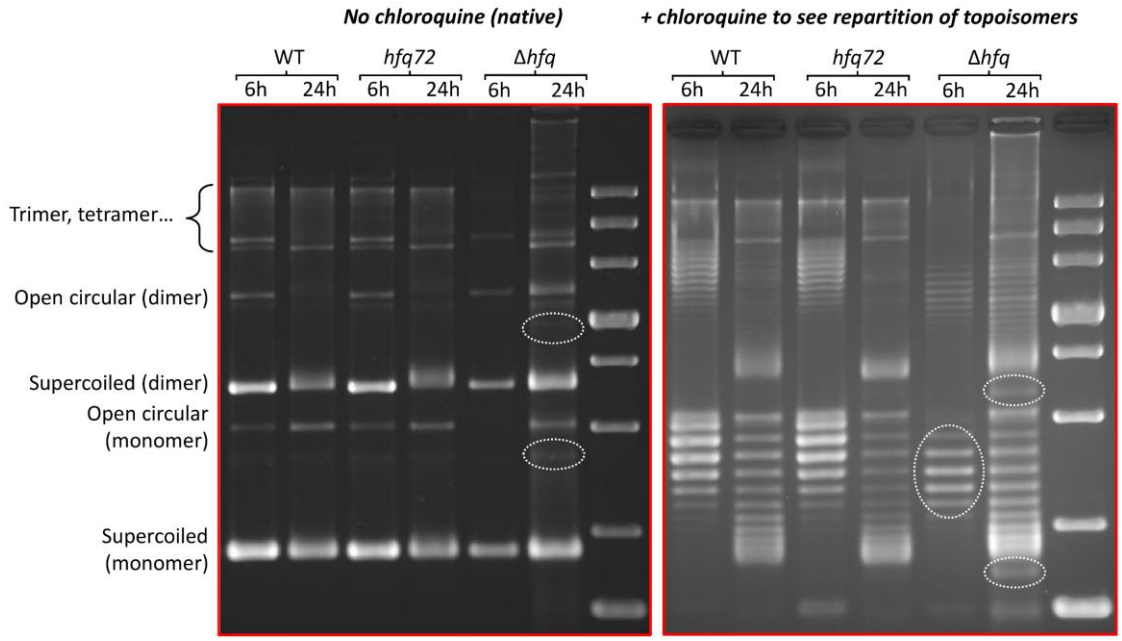


Figure 7: Topology gel of pHSG298 plasmids extracted from different strains. In the absence of chloroquine (native gel) relaxed plasmid goes slower, while in the presence of chloroquine (to see repartition of topoisomers) relaxed plasmid goes faster. Note that left and right parts are different gels stained separately. From this analysis we can conclude that plasmids are usually more relaxed at 24h (stationary phase), that there is no difference between *hfqWT* and *hfq72*, indicating that strong compaction due du the CTR does not affect plasmid topology. Nevertheless, we also confirm that plasmids are more relaxed in Δ*hfq* strain than in others (see dashed ovals). The small bimodal population seen in Δ*hfq* strain at 24h is comparable to the effect observed previously[2]. The presence of multimers is due to the use of MG1655 strain, which is *recA+*.

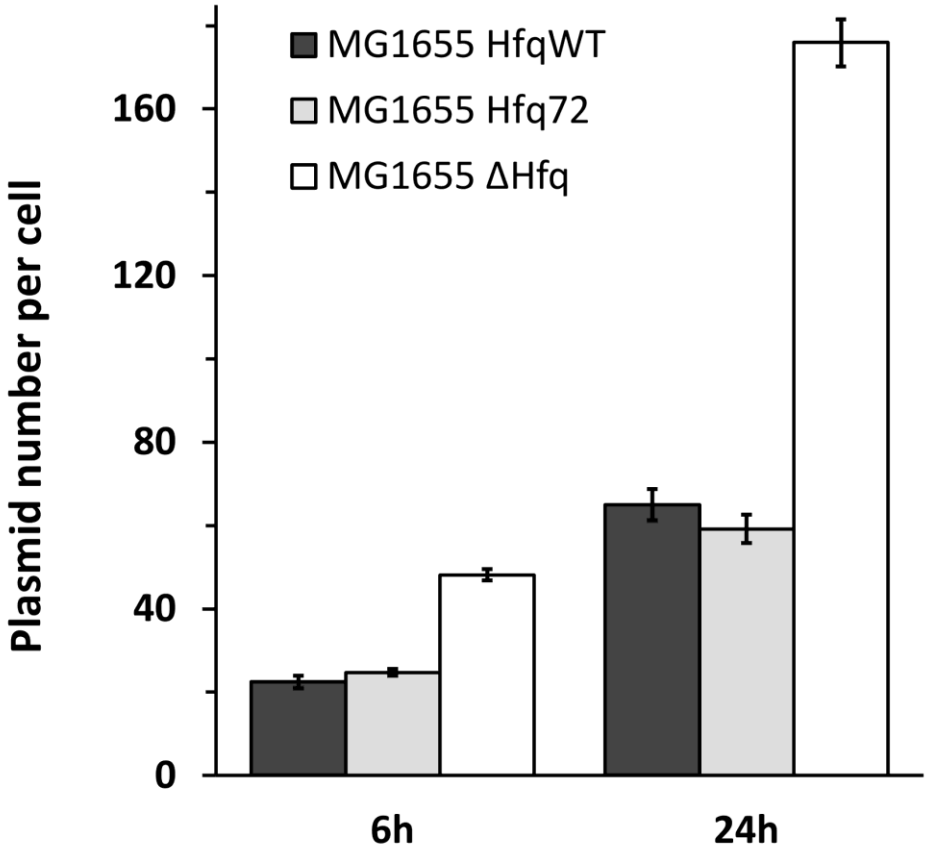


Figure 8. Evaluation of pHSG298 plasmid copy number in *hfq* mutated strains. As expected, a great number of plasmids is observed in late exponential and stationary phases and in the total absence of Hfq. This reconfirms the results of Cech *et al*[37] and shows additionally that deletion of Hfq CTR does not have an impact on plasmid copy number.

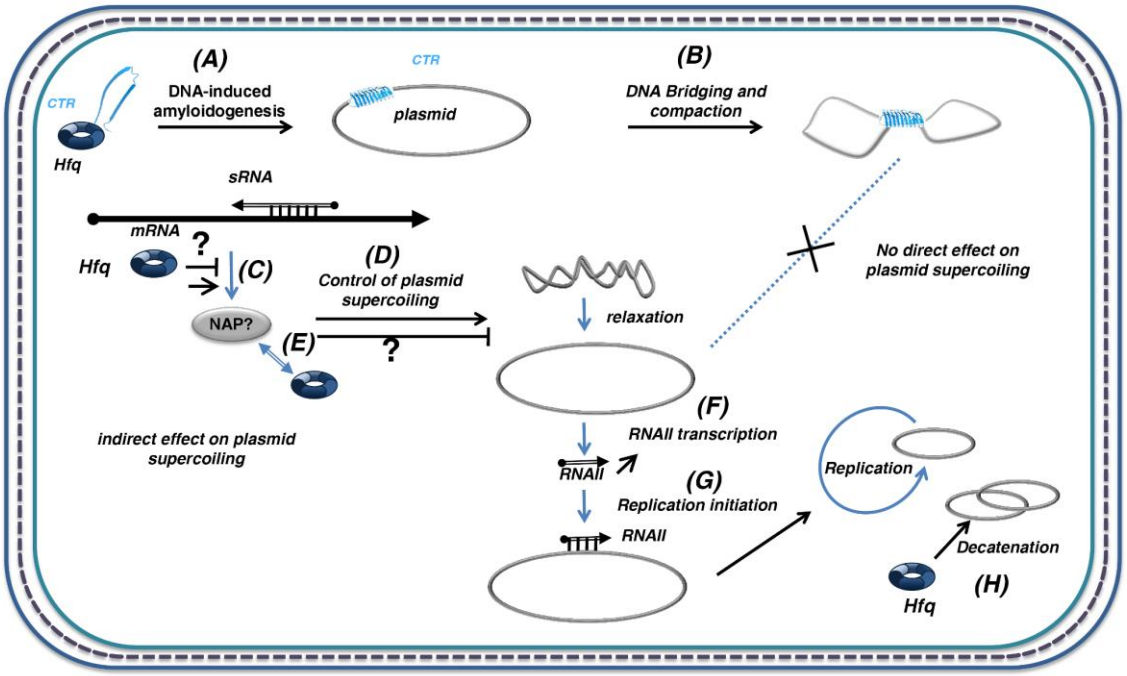


Figure 9. Hfq role in the control of plasmid compaction, supercoiling and replication. Hfq was previously shown to strongly influence DNA compaction (*B)*[23,24]. But Hfq also regulates at the post-transcriptionnal level the expression of various proteins, including NAPs such as H-NS or Dps[56,57] *(C)*. These NAPs can influence DNA topology[55] *(D)*. Here we demonstrate that DNA induces amyloidogenesis of Hfq-CTR (ΣCTR refers to amyloid self-assembly) *(A)*, resulting in DNA bridging and compaction *(B)*. In parallel, sRNA-dependent post-transcriptional regulation may regulate positively or negatively the expression of one NAP, resulting in changes in plasmid supercoiling (positive or negative). Note that Hfq and some NAP may interact physically or functionally *(E)*. The global effect of Hfq deletion on topology results in plasmid relaxation, possibly due to primer (RNAII) transcription (*F)* involved in replication initiation *(G)* or in plasmids decatenation after replication *(H)*. Hence, Hfq deletion influences plasmid replication positively. Plasmids are depicted as grey circles; non coding RNAs as open arrows; mRNAs as thick black lines; 5' and 3' of the mRNA are depicted by a "ball and arrow head", respectively; Hfq as a toroidal hexamer; other proteins (NAPS) as grey ellipses; positive and negative regulation are

indicated by arrows and horizontal bars, respectively; dotted line symbolizes Peptidoglycan (PG) between outer (OM) and inner (IM) membranes.

Annexe D : Article 4

Tailoring the Hydrophobicity of Mesoporous Organosilica for Protein Trapping and Supported Catalysis

Au Laboratoire Léon Brillouin (LLB), ont été étudiés des matériaux contenant des nanocavités ou matrices. Fabriquées et fonctionnalisées au laboratoire, ces matrices à base de silice peuvent être utilisées dans le cas de piégeage ou de catalyse. Leur capacité à adsorber des solutions de protéines, dans cette étude le lysozyme, a été étudiée dans l'objectif de les utiliser comme matériau de confinement, et ce en vue d'analyses ultérieures par diffusion de neutrons. Cette étude, à terme, pourra également être intéressante pour étudier des modèles amyloïdes dont le CTR Hfq, car le confinement pourrait influencer la dynamique de l'assemblage.

L'article présenté dans cette annexe propose une méthode élaborée au LLB pour améliorer le piégeage du lysozyme dans des matériaux organosilicés. Pour ce faire, une fonctionnalisation par différents groupes méthyles ou phényles permet de contrôler l'hydrophobicité de surface de la silice. La présence du groupement phénol en surface permet par exemple de piéger trois fois plus de protéines. La fonctionnalisation de surface favorise l'adsorption des protéines. Cependant, le diamètre des pores de la matrice est aussi un paramètre important. En effet, l'ajustement de la taille des pores conditionne l'entrée des molécules ciblées en leur fournissant un environnement plus ou moins confiné, ce qui pourra être important pour confiner de gros complexes tels que des amyloïdes.

D'un point de vue catalytique, l'immobilisation de cuivre (II) sur les matériaux et la présence de groupement méthylé en surface augmentent le rendement de synthèse de la réaction *Click* de cycloaddition azoture-alcyne (entre le bromobutyne et le benzylazide) dans l'eau. La fonctionnalisation de surface permet d'améliorer l'adsorption des réactifs organiques et l'ajustement de la taille des pores conditionne l'entrée des molécules ciblées en leur fournissant un environnement confiné favorisant l'interaction entre réactifs. L'optimisation de ces conditions améliorerait le rendement des réactions.

Cette méthode ouvre donc une nouvelle voie de synthèse avec ajustement des propriétés du matériau en vue d'applications spécifiques.

L'article joint s'intitule : « Tailoring the Hydrophobicity of Mesoporous Organosilica for Protein Trapping and Supported Catalysis. » par Oriana Osta, Marianne Bombled, **David Partouche**, Florian Gallier, Nadège Lubin-Germain, Nancy Brodie-Linder et Christiane Alba-Simionesco, soumis pour publication.

Tailoring the Hydrophobicity of Mesoporous Organosilica for Protein Trapping and Supported Catalysis.

Oriana Osta*[a], Marianne Bombled[a], David Partouche[a,c], Florian Gallier[b], Nadège Lubin-Germain[b], Nancy Brodie-Linder*[a,b], and Christiane Alba-Simionesco*[a].

[a] Laboratoire Léon Brillouin, CEA Saclay, Bâtiment 563, 91191 Gif-sur-Yvette, France.

[b] Laboratoire de Chimie Biologique, Université de Cergy-Pontoise, 5 mail Gay-Lussac, Neuville sur Oise, 95301 Cergy-Pontoise Cedex, France.

[c] Synchrotron SOLEIL, L'Orme des Merisiers, Saint Aubin, 91192 Gif-sur-Yvette, France.

ABSTRACT

We propose a method to enhance lysozyme trapping and supported-Copper catalysis when confined in organosilica materials. The direct synthesis presented here allows the control of the silica surface hydrophobicity by uniform introduction of methyl or phenyl groups. As a result, the lysozyme trapping is observed to be 3.2 times more efficient with the phenyl-functionalized material than MCM-41. For heterogeneous catalysis, Copper was immobilized on the new organosilica surface. In this case, the presence of methyl groups significantly enhanced the product yield for the catalyzed synthesis of a triazole derivative. This method opens a new route of synthesis where the material properties can be adjusted and dedicated to a specific application.

Molecular trapping and supported catalysis are applications of mesoporous materials with great industrial and scientific impact. A key parameter for their efficiency is the hydrophobicity of the surface, accompanied by a large surface area and the appropriate pore size[1]. Classic examples for applications in molecular trapping are the encapsulation followed by controlled drug delivery in nanomedicine[2], the effective removal of the low concentrations of pollutants[3] and molecular confinement studies[1,4–6] to mention a few. In the area of catalysis, mesoporous silica-supported reactions have been investigated for example, in wastewater treatment[7], in reducing the amount of certain compounds in the mainstream of tobacco smoke[8] and in product yield improvement for organic reactions,[9–11] among other

applications. In a particular case, copper doped nanoporous silica have been found to be a reusable, regioselective catalyst for the cycloaddition reaction between a terminal alkyne and an azide through click chemistry.[10,12] However, in some cases, the yields of reaction are very low and this is attributed to the low interaction between the hydrophobic reactants and the hydrophilic materials.

In this context, MCM-41 materials offer high surface area, ordered structure and thermal stability, but it fails in providing the correct hydrophobic/hydrophilic balance for an efficient interaction with most organic molecules. Frequently, even after post-grafting hydrophobic groups on the surface, this balance is not met due to a total transformation of the surface from hydrophilic to hydrophobic. A partial functionalization with this method has shown to produce inhomogeneous, rough surfaces with the majority of the grafted groups on the external surface or at the extremes or the pores.[9,13,14]

The co-condensation method has been conceived as an alternative to introduce hydrophobic groups to mesoporous silica materials yielding to a more uniform functionalization[9,15]. In this case, a silica precursor is condensed together with an organosilane. This method has been studied to introduce acid, basic and unsaturated groups in the walls of MCM-41-like materials so that they can be directly used for catalytic applications.[14,16–18] Our aim was to control the hydrophobicity of organosilica materials for specific applications by introducing methyl or phenyl groups through co-condensation with methyltriethoxysilane (MTES) or phenyltriethoxysilane (PTES). Such an approach for the synthesis of organosilica molecular sieves[19] has been much less investigated.

Here, lysozyme was chosen as a model to study the effect of the surface hydrophobicity in protein uptake. In addition, the supported copper (click) catalysis of the cycloaddition reaction between two hydrophobic reactants, benzyl azide and 4-bromo-1-butyne, in aqueous media instead of an organic solvent, was tested with the synthesized materials.

For the direct synthesis of mesoporous organosilica, we focused on obtaining a stable material with a highly ordered pore structure. In order to avoid the disruption of the honeycomb-like structure and based on previous research,[20] the percentage of organosilane in the precursor mixture was kept under 20 mol%. The conditions of template washing by ethanol extraction were optimized to obtain a clean material without causing damages to the surface.

Samples were named according to the template used and their fraction of MTES or PTES in the precursor mixture, as follows: C18-M7-Me for 0.93 TEOS/ 0.07 MTES, C18-M15-Me for 0.85 TEOS/ 0.15 MTES, C18-M7-Ph for 0.93 TEOS/ 0.07 PTES, all of them using C_{18}TAB as template, and C12-M7-Me for 0.93 TEOS/ 0.07 MTES using C_{12}TAB as template. Standard MCM-41 C18 and C12 were synthesized and washed under the same conditions for comparison.

Template removal and incorporation of the organic groups in the mesoporous silica was verified by means of FTIR (Figure S1 in ESI). Nitrogen adsorption and desorption isotherms at 77 K for these samples are shown in Figure 1. For standard and methyl modified MCM-41 C18, results show a regular type IV adsorption isotherm. In the case of co-condensation of TEOS with MTES, it was observed an increase of surface area and pore volume with respect to MCM-41 C18 (See Table 1). On the other hand, the co-condensation with PTES produced a less organized material, evidenced by a spread capillary condensation, slightly lower surface area and pore volume than MCM-41 C18.

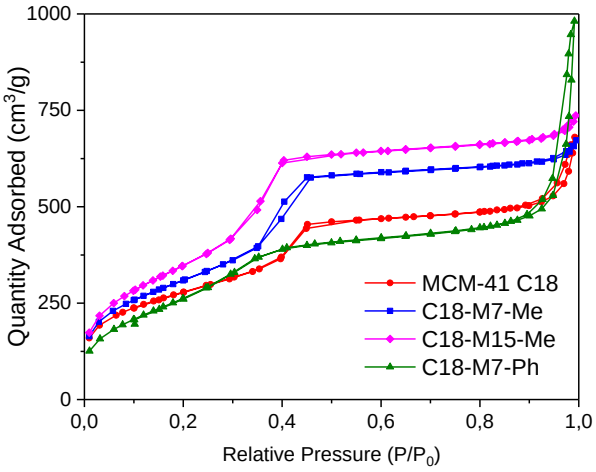


Figure 58. Nitrogen adsorption-desorption isotherms at 77 K of MCM-41 C18 and mesoporous organosilica materials.

The late condensation at P/P_0 higher than 0.9 in samples MCM-41 C18 and C18-M7-Ph was attributed to condensation in the interparticle cavities when the size of the mesoporous silica grains are in the nanometer scale[21], as it was evidenced for C18-M7-Ph by TEM (Figure S1).

These results indicate that upon introduction of MTES at percentages up to 15%, the obtained organosilica maintain an ordered pore structure while increasing their hydrophobic character. To investigate the upper limit of the fraction of MTES in the precursors mixture

that could yield an ordered, highly porous material, an extra sample was synthesized with 20% MTES. Results show that even though a high BET surface area (903 m^2/g) is obtained, there is a significant decrease in the pore volume (0.486 cm^3/g) of the material.

Likewise, sample C12-M7-Me was synthesized in order to explore the hydrophobic modification of small pore mesoporous silica, which is difficult to obtain by post-grafting methods. In this case, a high surface area material, with pore volume slightly lower than the MCM-41 C12 was obtained. The production of such small pore sized material with controlled hydrophobicity, like the one obtained in this work, could be of great scientific interest for molecular confinement studies.

It is worth noting that the affinity constant, C, calculated from the BET equation, decreased 31%, 49% and 61% with respect to the value of MCM-41 C18 (C = 112) for C18-M7-Me, C18-M15-Me and C18-M7-Ph, respectively. This parameter decreases with a decrease on the strength of the adsorbent-adsorbate affinity. Since induced dipole-dipole interactions govern this adsorption, the introduction of organic groups decreases the affinity of nitrogen towards the surface. Then, the observed drop of the C value within the same family of materials is another evidence of their increasing hydrophobicity.

Figure 2 shows the Small-angle X-Ray scattering patterns for the synthesized materials. The observation of three Bragg peaks indexed as (10), (11) and (20) confirm the 2D hexagonal structure p6mm of the standard, and methyl modified materials. The broadening of the (10) peak of the phenyl modified sample could suggest a higher polydispersity in the pore size.

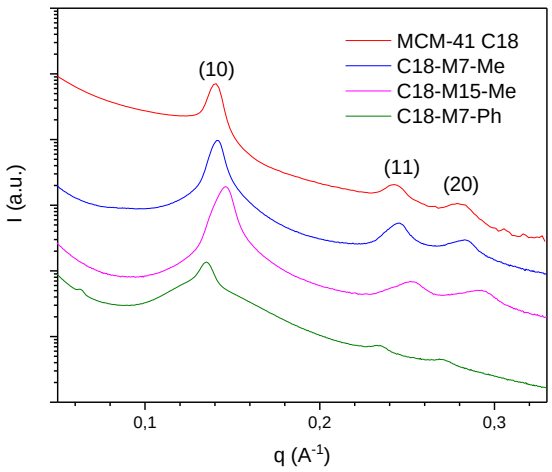


Figure 59. Small angle scattering curves of MCM-41 C18 and mesoporous organosilica materials.

The lattice plane spacing d_{100} was calculated from the SAXS patterns and it is shown in Table 1. This parameter in combination with the pore volume obtained by nitrogen adsorption allowed the calculation of the pore diameter (D_{pore}). This method has been found to be more precise for the determination of the size of small hydrophobic pores than other models that use exclusively nitrogen adsorption data[22]. We found that up to 15% MTES or 7% PTES, the pore diameter does not change significantly with respect to MCM-41, while the post graphing method induces always a decreasing and less predictable size.[22]

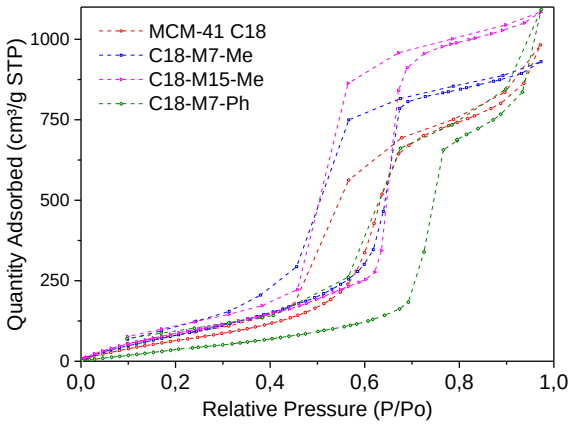


Figure 60. Water adsorption isotherms at 298.15 K of (a) MCM-41 C18, (b) C18-M7-Me, (c) C18-M15-Me and (d) C18-M7-Ph.

Table 1. Mean physical properties of the mesoporous organosilica materials.

Sample	S_{BET} (m^2/g)	V_{pore} (cm^3/g)[a]	d_{10} (nm)	D_{pore} (nm)[d]	Cu(II) (%w/w)
MCM-41 C18	1007	0.712	4.45	4.22	-
C18-M7-Me	1130	0.899	4.48	4.43	-
C18-M15-Me	1284	0.982	4.31	4.32	-
C18-M20-Me	903	0.447	4.05	3.46	-
C18-M7-Ph	989	0.628	4.65	4.30	-
MCM-41 C12	979	0.539	3.55	3.17	-
C12-M7-Me	1134	0.505	3.37	2.97	-
Cu/MCM-41 C18	308	0.286[b]	[c]	-	14.5
Cu/C18-M7-Me	256	0.181	4.46	2.89	9.8
Cu/C18-M15-Me	377	0.227	4.32	3.02	9.9
Cu/C18-M7-Ph	430	0.331	4.55	3.58	9.9

[a] Single point adsorption total pore volume at $P/P_o = 0.5$ [b] Single point adsorption total pore volume at $P/P_o = 0.8$ [c] Not measured [d] Calculated using the KJS model taking into account the lattice plane spacing (d_{10}).[23]

Water adsorption isotherms (Figure 3) show that the total quantity of water increases when the pore volume of the material is higher, regardless of their organic functionalization These results show that the hydrophobicity of the materials is not strong enough to avoid water vapor condensation in the pores.

The stability of the materials towards hydrolysis was studied by performing nitrogen adsorption analysis before and after of water adsorption. Results (Figure S2) clearly show that samples C18-M15-Me and C18-M7-Ph are better stabilized, with a smaller loss of pore volume than C18-M7-Me. The water-resistance found here, could be crucial for better performance of the molecular sieves in aqueous solution.

Transmission electron microscopy images of selected samples are shown in Figure S3. Both C18-M7-Me and C18-M7-Ph show the highly ordered, hexagonal structure, surrounded by mesocellular silica foam (MCF), reported before when 1,3,5-trimethylbenzene was incorporated to the synthesis of the mesoporous silica[24]. A possible explanation is that MCF rises due to the formation of a microemulsion upon addition of the organosilane, which behaves as a swelling agent, due to its hydrophobic character. The abundance of MCF structures in M7-Phenyl explains the increased width of the Bragg's peaks in the SAXS pattern for this sample.

Additionally, in order to verify the presence of methyl groups at the surface of the pore, we have conducted quasielastic neutron scattering experiments in the fixed elastic mode window mode of 2μeV. From the T-dependent of the elastic scans obtained, we could identify a pronounced enhanced mobility of the methyl groups starting at 80K. (Figure S4).

Protein Trapping

The selected organosilica materials were placed in contact with a 0.5 mg/ml solution of lysozyme at pH 7.4 at room temperature during 4h, then they were filtered and the decrease on concentration of the protein in solution was measured by the UV adsorption at 280 nm. This allowed the calculation of the amount of lysozyme adsorbed per milligram of molecular sieve. In the case of methyl-modified materials C18-M7-Me and C18-M15-Me the quantities adsorbed per milligram of organosilica (55 ± 1 μg/mg and 50 ± 5 μg/mg respectively) were lower than for the reference MCM-41 C18 (85 ± 10 μg/mg). Remarkably, in the case of the phenyl-modified sample, C18-M7-Ph, an adsorption 3.2 times higher was observed (270 ± 30 μg/mg).

Lysozyme adsorption is driven by electrostatic interactions between the Si-OH groups and the positively charged protein at pH 7.4. Therefore, the presence of $–CH_3$ instead of –OH groups on the surface could be one of the reasons of the lower quantity adsorbed by the methyl-modified silica. On the other hand, the high amount of lysozyme adsorbed by C18-M7-Ph is attributed to its ability to interact with the protein *via* electrostatic as well as π - π interactions.

FTIR spectra (Figure 4) show peaks centered at 1656 cm^{-1} and 1535 cm^{-1} corresponding to the amide I and amide II bands of lysozyme, respectively. The small peak at 1594 cm^{-1} comes from the aromatic C-C stretch of the phenyl functionalization. The amide I band consists of C=O stretching coupled with N-H bending and C-H stretching components. As these stretching modes are affected by the hydrogen bonds present in the different conformations (α-helix, parallel and antiparallel β-sheets, etc.), the shape of the amide I band gives information about the secondary structure of the protein[25]. It is then possible to perform studies of protein structure under different confined conditions using these new molecular sieves.

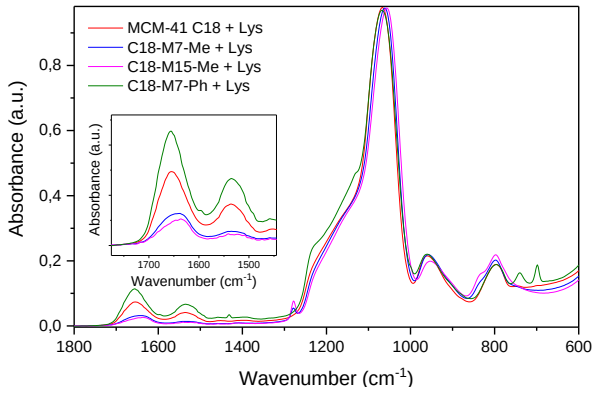


Figure 4. FTIR (ATR) Spectra of the lysozyme adsorbed on the materials. Insert : Zoom of the amide I and II bands of lysozyme.

Catalysis of the alkyne-azide cycloaddition reaction

The azide-alkyne cycloaddition reaction at ambient temperature, as shown in figure 5, typically requires a copper (I) species as catalyst. This species is prompt to disproportionate, making it necessary to add an auxiliary reducing species, as ascorbate, to have an efficient catalysis. Additionally, an extraction or filtration step has to be performed to purify the

product[26]. Another alternative is to use copper loaded on nanoporous silica for heterogeneous catalysis.[12]

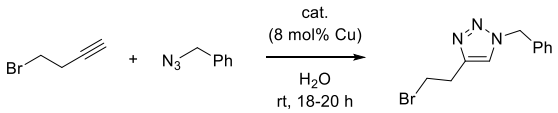


Figure 5. Cycloaddition reaction between bromobutyne and benzylazide in the presence of a copper catalyst in water.

Copper (II) was immobilized on the materials using the method developed by Brodie-Linder *et al*[27]. N_2 adsorption isotherms for the copper post-grafted materials (Figure 6) show a lower total amount adsorbed than the pristine samples.

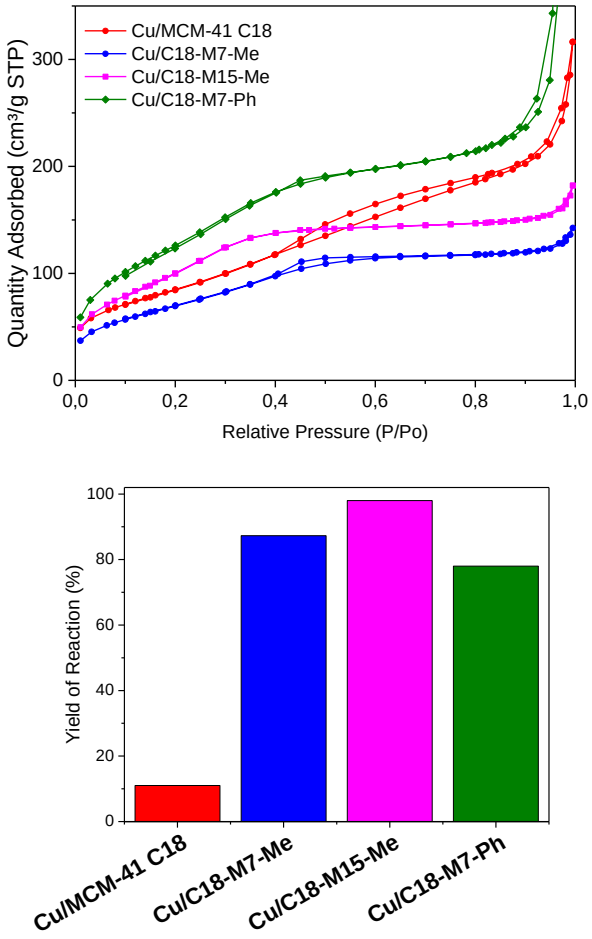


Figure 6. Nitrogen adsorption isotherms at 77 K (up) and yields of the cycloaddition reaction between bromobutyne and benzylazide in water (down) for the different silica-supported copper catalysts.

The irregular shape of the isotherm of Cu/MCM-41 C18 indicates more important wall degradation due to the contact with the basic copper solution. In this sense, the organic groups seem to protect the surface from basic hydrolysis, same as observed after water adsorption, Cu/C18-M7-Ph being the least affected in terms of surface area and pore volume. There is also an expected decrease on the pore size in all samples due to the incorporation of copper on the surface.

The materials were tested as catalysts for this reaction, with no addition of reducing agents or further ligands. Figure 6 shows the yields of reaction obtained by means of NMR spectroscopy. A vast increase in the yield of reaction is observed for the methyl- and phenyl-modified samples Cu/C18-M7-Me, Cu/C18-M15-Me and Cu/C18-M7-Ph (87%, >99% and 78% respectively) with respect to the yield of 11% obtained when Cu/MCM-41 C18 is used as catalyst.

These results confirm that the surface area and the organic functionalization of the materials play an important role in the efficiency of the catalyst. In this case, the presence of methyl or phenyl groups enhances the adsorption of the organic reactants. Another important factor is the pore diameter of the catalyst. It has been observed before that a tight fit that allows the entrance of the targeted molecules while also providing a confined environment for the reactants to interact, is a key factor in obtaining high reaction yields. Total transformation of the reactants on Cu/C18-M15-Me (pore size of 3.02 nm and a surface area of 377 m^2/g) indicates that this material provides an optimum hydrophobic nanoenvironment for this particular reaction in water. A control experiment using copper-doped materials with pores blocked by surfactant (Figure S6) allowed confirmation that most of the conversion takes place inside the pores of the hydrophobically modified materials.

Conclusion

By means of a direct synthesis of TEOS with MTES or PTES, mesoporous organosilica materials were synthesized with different degrees of hydrophobicity for testing their applications in lysozyme trapping and supported copper catalysis. A complete characterization confirmed the large surface area, ordered pore structure and organic functionalization of the materials, providing a partial hydrophobic environment that depends on the percentage of organosilane present in the precursor mixture. It was observed that such

a difference in the nanoenvironments had a significant impact on the performance of the materials in the selected applications.

The quantity of lysozyme adsorbed decreased for all the methyl-modified materials, with respect to the original MCM-41 C18. In contrast, the lysozyme uptake by the phenyl-modified sample, C18-M7-Ph, was more than three times higher than the reference MCM-41. This enhancement is attributed to π-π interactions between the organosilica and the protein.

The copper doped materials were proven to be active for regiospecific catalysis of the cycloaddition reaction between bromobutyne and benzyl azide, using water as solvent. In particular, C18-M15-Me has provided an ideal surface for this specific reaction, with an obtained yield of reaction of >99%. These unprecedented results for a Cu (II) supported catalyst, using such a facile synthesis encourages the testing of the materials for the synthesis of further triazole derivatives in water.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

Detailed experimental procedure, FTIR spectra, N2 adsorption isotherms before and after water adsorption, TEM images, Neutron Backscattering Experiments, yields of catalysis reaction with blocked pores and deconvolution of the FTIR amide I band for lysozyme in solution and on C18-M7-Ph. (PDF)

AUTHOR INFORMATION

Corresponding Authors

*E-mail: christiane.alba-simionesco@cea.fr

*E-mail: nancy.brodie-linder@u-cergy.fr, nancy.linder@cea.fr

*E-mail: oriana.osta@cea.fr

ACKNOWLEDGMENT

The authors thank Veronique Arluison for helpful discussion on the confined lysozyme, Annie Brulet for the SAXS experiments and Henriette Wase Hansen and Bernhard Frick for inelastic neutron scattering on IN16B at the ILL (Grenoble, France)

REFERENCES

(1) Lei, C.; Shin, Y.; Magnuson, J. K.; Fryxell, G.; Lasure, L. L.; Elliott, D. C.; Liu, J.; Ackerman, E. J. Characterization of Functionalized Nanoporous Supports for Protein Confinement. *Nanotechnology* **2006**, *17*, 5531–5538.

(2) Charnay, C.; Bégu, S.; Tourné-Péteilh, C.; Nicole, L.; Lerner, D. A.; Devoisselle, J. M. Inclusion of Ibuprofen in Mesoporous Templated Silica: Drug Loading and Release Property. *Eur. J. Pharm. Biopharm.* **2004**, *57* (3), 533–540.

(3) Cashin, V. B.; Eldridge, D. S.; Yu, A.; Zhao, D. Surface Functionalization and Manipulation of Mesoporous Silica Adsorbents for Improved Removal of Pollutants: A Review. *Environ. Sci. Water Res. Technol.* **2018**, *4* (2), 110–128.

(4) Dosseh, G.; Xia, Y.; Alba-Simionesco, C. Cyclohexane and Benzene Confined in MCM-41 and SBA-15: Confinement Effects on Freezing and Melting. *J. Phys. Chem. B* **2003**, *107* (26), 6445–6453.

(5) Azaïs, T.; Laurent, G.; Panesar, K.; Nossov, A.; Guenneau, F.; Sanfeliu Cano, C.; Tourné-Péteilh, C.; Devoisselle, J.-M.; Babonneau, F. Implication of Water Molecules at the Silica–Ibuprofen Interface in Silica-Based Drug Delivery Systems Obtained through Incipient Wetness Impregnation. *J. Phys. Chem. C* **2017**, *121* (48), 26833–26839.

(6) Brodrecht, M.; Kumari, B.; Breitzke, H.; Gutmann, T.; Buntkowsky, G. Chemically Modified Silica Materials as Model Systems for the Characterization of Water-Surface Interactions. *Z. Für Phys. Chem.* **2018**, *232* (7–8), 1127–1146.

(7) Wu, Q.; Hu, X.; Yue, P. L.; Zhao, X. S.; Lu, G. Q. Copper/MCM-41 as Catalyst for the Wet Oxidation of Phenol. *Appl. Catal. B Environ.* **2001**, *32*, 151–156.

(8) Marcilla Gomis, A. Effect of Mesoporous Catalysts on the Mainstream Tobacco Smoke of 3R4F and 1R5F Reference Cigarettes. *Am. J. Chem. Eng.* **2015**, *3* (1), 1–18.

(9) Mbaraka, I.; Shanks, B. Design of Multifunctionalized Mesoporous Silicas for Esterification of Fatty Acid. *J. Catal.* **2005**, *229*, 365–373.

(10) Roy, S.; Chatterjee, T.; Pramanik, M.; Roy, A. S.; Bhaumik, A.; Islam, S. M. Cu(II)-Anchored Functionalized Mesoporous SBA-15: An Efficient and Recyclable Catalyst for the One-Pot Click Reaction in Water. *J. Mol. Catal. Chem.* **2014**, *386*, 78–85.

(11) Rostamnia, S.; Doustkhah, E. Nanoporous Silica-Supported Organocatalyst: A Heterogeneous and Green Hybrid Catalyst for Organic Transformations. *RSC Adv.* **2014**, *4*, 28238–28248.

(12) Jlalia, I.; Gallier, F.; Brodie-Linder, N.; Uziel, J.; Augé, J.; Lubin-Germain, N. Copper(II) SBA-15: A Reusable Catalyst for Azide-Alkyne Cycloaddition. *J. Mol. Catal. Chem.* **2014**, *393*, 56–61.

(13) Wang, X.; Lin, K. S. K.; Chan, J. C. C.; Cheng, S. Direct Synthesis and Catalytic Applications of Ordered Large Pore Aminopropyl-Functionalized SBA-15 Mesoporous Materials. *J. Phys. Chem. B* **2005**, *109*, 1763–1769.

(14) Yokoi, T.; Yoshitake, H.; Tatsumi, T. Synthesis of Amino-Functionalized MCM-41 via Direct Co-Condensation and Post-Synthesis Grafting Methods Using Mono-, Di- and Tri-Amino-Organoalkoxysilanes. *J. Mater. Chem.* **2004**, *14* (6), 951.

(15) Sayari, A.; Hamoudi, S. Periodic Mesoporous Silica-Based Organic-Inorganic Nanocomposite Materials. *Chem. Mater.* **2001**, *13*, 3151–3168.

(16) Shang, F.; Sun, J.; Wu, S.; Yang, Y.; Kan, Q.; Guan, J. Direct Synthesis of Acid–Base Bifunctional Mesoporous MCM-41 Silica and Its Catalytic Reactivity in Deacetalization–Knoevenagel Reactions. *Microporous Mesoporous Mater.* **2010**, *134*, 44–50.

(17) Macquarrie, D. J. Direct Preparation of Organically Modified MCM-Type Materials. Preparation and Characterisation of Aminopropyl–MCM and 2-Cyanoethyl–MCM. *Chem. Commun.* **1996**, No. 16, 1961–1962.

(18) Fowler, C. E.; Burkett, S. L.; Mann, S. Synthesis and Characterization of Ordered Organo–Silica–Surfactant Mesophases with Functionalized MCM-41-Type Architecture. *Chem. Commun.* **1997**, No. 18, 1769–1770.

(19) Putz, A.-M.; Wang, K.; Len, A.; Plocek, J.; Bezdicka, P.; Kopitsa, G. P.; Khamova, T. V.; Ianăşi, C.; Săcărescu, L.; Mitróová, Z. Mesoporous Silica Obtained with Methyltriethoxysilane as Co-Precursor in Alkaline Medium. *Appl. Surf. Sci.* **2017**, *424*, 275–281.

(20) Burkett, S. L.; Sims, S. D.; Mann, S. Synthesis of Hybrid Inorganic–Organic Mesoporous Silica by Co-Condensation of Siloxane and Organosiloxane Precursors. *Chem. Commun.* **1996**, *11*, 1367–1368.

(21) Marei, N. N.; Nassar, N. N.; Vitale, G. The Effect of the Nanosize on Surface Properties of NiO Nanoparticles for the Adsorption of Quinolin-65. *Phys. Chem. Chem. Phys.* **2016**, *18*, 6839–6849.

(22) Schoeffel, M.; Brodie–Linder, N.; Audonnet, F.; Alba–Simionesco, C. Wall Thickness Determination of Hydrophobically Functionalized MCM-41 Materials. *J Mater Chem* **2012**, *22*, 557–567.

(23) Kruk, M.; Jaroniec, M.; Sayari, A. Adsorption Study of Surface and Structural Properties of MCM-41 Materials of Different Pore Sizes. *J. Phys. Chem. B* **1997**, *101*, 583–589.

(24) Xin, C.; Zhao, N.; Zhan, H.; Xiao, F.; Wei, W.; Sun, Y. Phase Transition of Silica in the TMB–P123–H2O–TEOS Quadru-Component System: A Feasible Route to Different Mesostructured Materials. *J. Colloid Interface Sci.* **2014**, *433*, 176–182.

(25) Adams, S.; Higgins, A. M.; Jones, R. A. L. Surface-Mediated Folding and Misfolding of Proteins at Lipid/Water Interfaces. *Langmuir* **2002**, *18*, 4854–4861.

(26) Himo, F.; Lovell, T.; Hilgraf, R.; Rostovtsev, V. V.; Noodleman, L.; Sharpless, K. B.; Fokin, V. V. Copper(I)-Catalyzed Synthesis of Azoles. DFT Study Predicts Unprecedented Reactivity and Intermediates. *J. Am. Chem. Soc.* **2005**, *127*, 210–216.

(27) Brodie-Linder, N.; Besse, R.; Audonnet, F.; LeCaer, S.; Deschamps, J.; Impéror-Clerc, M.; Alba-Simionesco, C. The Key to Control Cu II Loading in Silica Based Mesoporous Materials. *Microporous Mesoporous Mater.* **2010**, *132*, 518–525.

université PARIS-SACLAY
ÉCOLE DOCTORALE
INTERFACES
Approches Interdisciplinaires : Fondements, Applications et Innovation

Titre : Analyse de l'assemblage de peptides amyloïdes bactériens.

Mots clés : amyloïde fonctionnel, nano-spectro-imagerie AFM-IR, tecnique corrélative.

Résumé : Hfq est une protéine bactérienne qui a un rôle pléïotropique. La principale fonction de la protéine Hfq bactérienne consiste à répondre aux stress que peut rencontrer la bactérie lors d'un changement environnemental, en utilisant essentiellement un contrôle post-transcriptionnel. La protéine, par sa capacité à interagir avec les ARN et notamment les petits ARN non codant, permet ainsi une régulation rapide de l'expression génétique. En outre la protéine interagit aussi avec l'ADN qu'elle aide à se structurer. Les mutations dans le gène qui code pour Hfq ont des effets pléïotropes c'est-à-dire déterminant plusieurs caractères phénotypiques.

D'un point de vue structural, la protéine adopte un repliement de type Sm, caractérisé par un oligomère toroïdal reposant sur la formation d'un feuillet β continu à 30 brins. Cependant, outre cette région Sm N-terminale, Hfq possède également une région C-terminale (CTR) de taille et de séquence variables selon les bactéries.

Mon travail de thèse a porté sur l'analyse de cette région CTR chez la bactérie *Escherichia coli*. Cette région a en effet la capacité de former une structure de type amyloïde : structures auto-assemblées in vivo, à proximité de la membrane interne et dans le nucléoïde. Par l'utilisation de diverses techniques physico-chimiques (microscopie moléculaire, spectroscopie et microscopie infrarouge, dichroïsme circulaire et diffusion aux petits angles), mon travail a consisté à caractériser l'assemblage de cette région de Hfq ainsi que les facteurs l'influençant en particulier la présence d'acide nucléique. Une partie de mon travail de thèse a aussi consisté à mettre en place une méthode d'imagerie corrélative innovante permettant d'analyser la signature chimique et morphologique d'une fibre amyloïde unique. Mon travail a enfin porté sur l'analyse de l'effet de composés inhibant l'agrégation de la structure amyloïde, ce qui pourrait constituer une piste pour développer une nouvelle classe d'antibiotiques.

Title : Analysis of the assembly of bacterial amyloid peptides

Keywords : Functional amyloid, nanospectroscopy AFM-IR, correlative technique.

Abstract : Hfq is a pleiotropic bacterial protein that determines several phenotypic characteristics. Its main function is to facilitate responses to stresses that bacteria may encounter during environmental changes, mainly by using post-transcriptional genetic control. The protein, by its capacity to interact with RNA, in particular small non-coding RNA, enables a rapid regulation of gene expression. In addition, the protein also interacts with DNA and compacts it.

From a structural point of view, the protein adopts an Sm-like fold, characterized by a toroidal oligomer formed by a continuous 30-stranded β sheet. Besides its conserved N-terminal Sm domain, Hfq also possesses a C-terminal region (CTR) that can vary in size and sequence between bacteria.

My PhD work focused on the analysis of this CTR region in *Escherichia coli* bacteria.

Indeed, this region has the capacity to form an amyloid structure. This structural dynamic is related to the formation of self-assembled structures in vivo, in the proximity of the inner membrane and in the nucleoid. Using various physicochemical techniques (molecular microscopy, spectroscopy and infrared microscopy, circular dichroism and small angle X-ray scattering), my work consisted in characterizing the assembly of this region of Hfq, as well as the factors influencing its assembly (in particular, the presence of nucleic acids). A part of my work consisted in setting up an innovative correlative–imaging method to analyze the chemical and morphological signature of a single amyloid fibre. Finally, my work focused on the analysis of the effect of compounds that inhibit the aggregation of the amyloid structure, which could constitute a new way to develop a novel class of antibiotics.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France