



HAL
open science

Dynamique de l'agrégation protéique chez la bactérie **Escherichia coli**

Anne-Sophie Coquel

► **To cite this version:**

Anne-Sophie Coquel. Dynamique de l'agrégation protéique chez la bactérie *Escherichia coli*. Sciences agricoles. INSA de Lyon, 2012. Français. NNT : 2012ISAL0108 . tel-00778887v2

HAL Id: tel-00778887

<https://theses.hal.science/tel-00778887v2>

Submitted on 24 May 2013

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.

ANNE-SOPHIE COQUEL

**DYNAMIQUE DE L'AGREGATION PROTEIQUE
CHEZ LA BACTERIE *ESCHERICHIA COLI***

THESE

PRESENTEE DEVANT L'INSTITUT NATIONAL DES SCIENCES APPLIQUEES DE LYON (INSA)

POUR OBTENIR

LE GRADE DE DOCTEUR

DANS LA SPECIALITE INFORMATIQUE
ECOLE DOCTORALE INFO-MATHS (EDA 512)

SOUTENUE LE 16 NOVEMBRE 2012 DEVANT LE JURY COMPOSE DE

DR OLIVIER GANDRILLON	RAPPORTEUR	UMR5534, LYON
DR CHARLES KERVANN	RAPPORTEUR	INRIA, RENNES
PR GUILLAUME BESLON	EXAMINATEUR	INSA, INRIA, LYON
DR LIONEL MOISAN	EXAMINATEUR	CNRS UMR8145, PARIS
DR ABRAM AERTSEN	EXAMINATEUR	UNIVERSITE CATHOLIQUE, LEUVEN
DR HUGUES BERRY	DIRECTEUR DE THESE	INRIA, LYON
DR ARIEL LINDNER	DIRECTEUR DE THESE	INSERM U1001, PARIS



Remerciements

Je tiens tout d'abord à remercier mes deux directeurs de thèse Hugues Berry et Ariel Lindner, pour m'avoir fait confiance dans ce projet et m'avoir donné la possibilité de réaliser cette thèse.

J'exprime également mes remerciements à Olivier Gandrillon et Charles Kervrann pour avoir accepté d'être rapporteurs de cette thèse, ainsi que Guillaume Beslon, Lionel Moisan et Abram Aertsen pour en être les examinateurs.

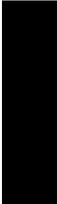
Je remercie tous mes collègues de l'INSERM U1001 et de l'équipe Beagle-INRIA pour avoir été là dans les moments de doute, pour le support technique, leurs conseils et leur expérience et pour tous les bons moments passés ensemble à déguster des pâtisseries ou organiser une raclette party.

Je remercie également toutes les personnes avec lesquelles j'ai eu l'occasion de collaborer. Merci à Fatima Gasset-Rosa et Rafael Giraldo, Thomas Julou, Mariella Dimiccoli, Maël Primet, Jean-pascal Jacob, Sander Govers, pour nos discussions, leur patience et leur gentillesse.

Pour leurs encouragements et leur soutien, je remercie, Olivier Détour, ma famille et mes ami(e)s qui ont toujours cru en moi et m'ont poussée à surfer sur la vague.

Pour les relectures et son avis de non-scientifique, je remercie un scénariste qui saura se reconnaître, et qui a trouvé une certaine poésie dans ce manuscrit scientifique.

Ces remerciements ne seraient pas complets si je ne remerciais pas les WOD et l'équipe de la box Reebok CrossFit Louvre ainsi que les cours et coaches du CMG Denfert qui m'ont permis de souffler au cours de mes journées de rédaction.



Résumé

L'agrégation protéique joue un rôle clé dans la dégénérescence cellulaire et est notamment reliée à de nombreuses maladies humaines en lien avec le vieillissement telles que les maladies d'Alzheimer et Parkinson ou encore la maladie du prion. Chez la bactérie *Escherichia coli*, l'accumulation de dommages sous forme d'agrégats protéiques et leur ségrégation asymétrique au pôle ont permis de démontrer des signes de vieillissement.

Cette thèse s'est concentrée sur l'étude de la dynamique spatiale des agrégats protéiques *in vivo* chez la bactérie *E. coli*. Les agrégats protéiques peuvent être classifiés comme corps d'inclusion dont on dit souvent qu'ils sont amorphes ou comme amyloïdes dont le niveau de structuration est très élevée par la présence de nombreux feuillet β . Combinant une double approche théorique et expérimentale, basée sur la modélisation et la microscopie time-lapse et microfluidique, nous avons étudié le mécanisme gouvernant le mouvement des agrégats protéiques et la transmission verticale d'agrégats de type prionoïde sur plusieurs dizaines de générations.

Nos résultats indiquent clairement que les agrégats protéiques sont régis par un mouvement Brownien de diffusion avec un coefficient de diffusion dépendant de la taille de la molécule. L'étude de protéinopathie amyloïde a démontré l'existence de lignages propageant deux types d'agrégats : globulaire ou en forme de "comet-like". Les lignées présentant les agrégats sous forme globulaire indiquent une augmentation de la taille des agrégats jusqu'à inhibition de la division cellulaire tandis que la forme "comet-like" est moins préjudiciable à la croissance. Nous avons également observé à faible fréquence des lignées contenant les deux types d'agrégats avec un changement de type d'agrégat. A partir d'un agrégat globulaire, des agrégats "comet-like" peuvent naître.

Mots clés : agrégation protéique, *Escherichia coli*, diffusion, prionoïde, croissance.



Abstract

Protein aggregation plays a key role in cell decline and leads to several human disease linked to ageing like Alzheimer or Parkinson disease and prion disease. In *Escherichia coli* bacteria, accumulation of damaged proteins and their asymmetric segregation allowed to show ageing signs. This thesis is focused on the *in vivo* spatial dynamics of protein aggregates in *E. coli*. Protein aggregates can be classified as inclusion bodies and they are amorphous or amyloid with a high order level due to β sheets. Combining a double theoretical and experimental approach, based on modeling and time-lapse and microfluidic microscopy, we studied the mechanism governing the motion of protein aggregates and the long-term vertical transmission of prionoid aggregates for about 10 generations.

Our results show clearly that Brownian diffusion governs the motion of protein aggregates and the diffusion coefficient depends on the molecule size. The amyloid proteinopathy study shows the existence of lineages propagating two kind of aggregates : globular or comet-like. Lineages maintaining globular aggregates present an increase of the aggregate size until inhibition of the growth rate while comet-like aggregates are mildly detrimental to growth. We observed also at low frequency in some lineages the presence of both aggregates and a switch between them. Globular foci give born to comet-like aggregates.

Keywords : protein aggregation, *Escherichia coli*, diffusion, prionoid, growth.



Avant-Propos

L'intérêt pour l'agrégation protéique a largement augmenté depuis la découverte de leur lien avec de nombreuses maladies neurodégénératives (Alzheimer, Parkinson, Huntington, prion). Ces maladies ont un impact économique et social de plus en plus élevé. Le coût des dépenses liées aux démences, dont Alzheimer est la plus importante, dépasse les 600 milliards de dollars, soit 1% du PIB mondial. En 2011, on dénombre 35 millions de personnes atteintes de la maladie d'Alzheimer et ce chiffre pourrait doubler d'ici 2030 (Rapport publié par Alzheimer's Disease International, ADI).

Les agrégats protéiques adoptent une conformation anormalement repliée, permettant ainsi la formation de structures complexes et organisées telles que les fibres amyloïdes. Ainsi, l'agrégation protéique, ou accumulation de dommages protéiques dans une cellule, pourrait être responsable en partie de la dégénérescence cellulaire et du vieillissement.

L'étude l'agrégation protéique est un problème interdisciplinaire relevant des domaines de la biologie expérimentale, biologie moléculaire, microscopie, physique. Il existe différents niveaux d'observation depuis la composition des agrégats protéiques et leur structure ou encore leur mouvement et localisation dans la cellule. Comprendre les mécanismes régissant l'agrégation protéique implique d'étudier la distribution spatiale d'agrégats protéiques mais aussi leur déplacement et leur corrélation avec le vieillissement. Est ce que la localisation polaire des agrégats protéiques observés chez la bactérie *Escherichia coli* est stochastique ? Est ce que le mouvement des agrégats protéiques est régi par un mécanisme passif de diffusion ou nécessite-t-il de l'ATP ? Dans la littérature, il existe très peu d'études quantitatives *in vivo* du mouvement des agrégats protéiques.

Avant d'évoquer le projet de thèse à proprement parler, j'introduirai le repliement des protéines et les mécanismes entraînant l'agrégation protéique ainsi que son lien avec le vieillissement. Enfin, une troisième partie se focalisera sur la diffusion des molécules dans une cellule, mécanisme passif qui gouverne la localisation de la majorité des molécules. Dans un deuxième chapitre, j'évoquerai les différents outils méthodologiques utilisés. Le troisième chapitre présentera les articles soumis. Je terminerai enfin par une discussion et les perspectives de ce projet.

Table des matières

1	Introduction	1
I	L'agrégation protéique	2
I.1	De la conformation native au repliement incorrect	2
I.1.1	Le repliement des protéines	2
I.1.2	Paysage énergétique et conformation native	4
I.1.3	Effet de l'encombrement moléculaire sur le repliement des protéines	5
I.1.4	Pourquoi les protéines s'agrègent ?	7
I.1.5	Le système contrôle qualité des protéines chez la bactérie	9
	I.1.5.1 Les chaperonnes	9
	I.1.5.2 Les protéases	12
I.2	Les différents types d'agrégats	12
I.2.1	Les maladies associées	13
I.2.2	Les corps d'inclusion	13
I.2.3	Les amyloïdes	14
I.2.4	Les prions et les prionoides	14
	I.2.4.1 Les prions chez la levure <i>Saccharomyces cerevisiae</i>	15
	I.2.4.2 Les prions chez la bactérie <i>Escherichia coli</i>	16
I.3	Vieillissement et agrégation protéique chez les organismes unicellulaires	16
I.3.1	Le vieillissement chez les organismes unicellulaires	17
	I.3.1.1 Le vieillissement chez la levure <i>Saccharomyces cerevisiae</i>	17
	I.3.1.2 Le vieillissement chez la bactérie <i>Caulobacter crescentus</i>	18
	I.3.1.3 Le vieillissement chez la levure <i>Schizosaccharomyces pombe</i>	19
	I.3.1.4 Le vieillissement chez la bactérie <i>Escherichia coli</i>	20
I.3.2	Ségrégation asymétrique des agrégats protéiques chez les organismes unicellulaires et vieillissement	23
	I.3.2.1 Chez la levure, <i>S. cerevisiae</i>	24
	I.3.2.2 Chez la bactérie <i>E. coli</i>	25
I.3.3	Asymétrie versus symétrie	28
I.3.4	Mobilité des agrégats protéiques	28

II	La diffusion	29
II.1	Histoire de la diffusion	29
II.1.1	Le mouvement Brownien	29
II.1.2	Equation de la diffusion : Stokes-Einstein	32
II.2	Caractérisation de la diffusion	33
II.2.1	Marche aléatoire	33
II.2.2	Encombrement moléculaire et diffusion	33
II.2.3	Différents types de diffusion	34
II.3	Mesures de la diffusion	35
II.3.1	Ordres de grandeur	35
II.3.2	Chez la bactérie <i>E. coli</i>	35
II.3.3	Différences entre <i>in vivo</i> et <i>in vitro</i>	37
III	Objectifs	39
2	Méthodologie	41
I	Microscopie	42
I.1	Microscopie optique	42
I.1.1	Les bases	42
I.1.2	Microscope à contraste de phase	43
I.1.3	Microscope à fluorescence	43
I.1.4	Résolution optique	44
I.2	Technologies	45
I.2.1	Microscopie time-lapse	45
I.2.2	Système microfluidique	45
I.3	Suivi de molécules <i>in vivo</i>	47
II	Analyses d'images	49
II.1	Cellst (Cell Segmentation and Tracking)	50
II.1.1	La segmentation	51
II.1.2	Le suivi des cellules	51
II.1.3	Caractérisation et suivi des agrégats	52
II.2	Bacteria Analyzer	54
II.2.1	Traitement d'images	54
II.2.2	La segmentation et le suivi des cellules	54
II.2.3	Caractérisation et suivi des agrégats	56
III	La modélisation	58
III.1	Le modèle multi-agents (SMA)	58
III.2	Le modèle basé individus (IBM)	59
III.3	Notre modèle	59
3	Résultats	63
I	Article : Dynamique de l'agrégation protéique chez <i>E. coli</i>	64
I.1	Résumé	64
I.2	Article	64
II	Article : Effet de la protéine chaperonne DnaK sur la transmission du prionoïde RepA-WH1	104
II.1	Résumé	104

II.2	Article	104
4	Discussion	143
I	Dynamique des agrégats protéiques	144
I.1	Résumé des résultats obtenus	145
I.2	Localisation et mécanisme passif	145
I.3	Ségrégation des agrégats protéiques	149
I.3.1	Ségrégation asymétrique	149
I.3.2	Ségrégation stochastique	151
II	Toxicité des agrégats protéiques	152
II.1	Résumé des résultats obtenus	152
II.2	Influence de la forme des agrégats protéiques sur la toxicité	153
II.3	Influence de la taille des agrégats protéiques sur la toxicité	154
II.4	<i>Escherichia coli</i> , modèle d'étude des prions	154
III	Conclusion	155

Table des figures

1.1	Les différentes hypothèse de chemins de repliement d'une protéine.	3
1.2	Paysage énergétique en regard du repliement et de l'agrégation protéique.	4
1.3	Notion de volume exclu.	6
1.4	Vue d'ensemble du repliement et de l'agrégation protéique	8
1.5	Les différents classes de protéines chaperonnes	11
1.6	La levure <i>S. cerevisiae</i>	17
1.7	Durée de vie chronologique et réplivative chez la levure <i>S. cerevisiae</i>	17
1.8	La bactérie <i>C. crescentus</i>	18
1.9	Cycle de vie chez la bactérie <i>C. crescentus</i> et vieillissement	19
1.10	La levure <i>S. pombe</i>	19
1.11	La bactérie <i>E. coli</i>	20
1.12	Cycle de vie <i>Escherichia coli</i>	21
1.13	Effets de divisions consécutives à partir du vieux ou nouveau pôle sur la croissance	21
1.14	Dispositif microfluidique	22
1.15	Stabilité du taux de croissance à long terme chez <i>E. coli</i>	22
1.16	Taux de mortalité en fonction de l'âge chez <i>E. coli</i>	23
1.17	Accumulation de dommages chez la levure chez <i>S. cerevisiae</i>	24
1.18	Localisation <i>in vivo</i> de la protéine chaperonne ibpA fusionnée à la protéine fluo- rescente YFP.	25
1.19	Distribution des agrégats protéiques à leur première apparition.	25
1.20	Localisation au pôle de l'agrégat protéique.	26
1.21	Corrélation négative entre le taux de croissance cellulaire et la fluorescence des agrégats protéiques.	26
1.22	Corrélation entre le vieillissement et la présence d'agrégats protéiques.	27
1.23	Marches aléatoires d'une particule dans un espace à une dimension	31
1.24	Marche aléatoire d'une particule sur un réseau en deux dimensions.	31
1.25	Trajectoire d'une particule suivant une marche aléatoire dans un espace à deux dimensions.	32

1.26	Cycle cellulaire de la bactérie <i>Caulobacter crescentus</i> et localisation de la protéine PleC	36
1.27	Courbe du MSD en fonction du temps pour les données brutes et les données corrigées par rapport à la projection 2d	36
1.28	Tableau des coefficients de diffusion en 2d de l'histidine kinase PleC en fonction du type de cellule et de la position selon l'axe de la cellule	37
1.29	Coefficients de diffusion de multimètres (1 à 5) de GFP dans le cytoplasme de la bactérie <i>E. coli</i>	38
2.1	La lentille convergente.	42
2.2	Le microscope à contraste de phase.	43
2.3	Le microscope à fluorescence.	44
2.4	Disque d'Airy	44
2.5	Résolution latérale et ouverture numérique	45
2.6	Système de microscopie microfluidique	46
2.7	Cellule mère et évacuation des cellules	47
2.8	Images de fluorescence de la bactérie exprimant la protéine ClpX fusionnées à différents marqueurs fluorescents.	48
2.9	Vue d'ensemble de l'algorithme implémenté dans Cellst	50
2.10	Vue d'ensemble du processus de dilatation	51
2.11	Notion de risque	52
2.12	Squelettes actifs	53
2.13	Squelettes actifs	53
2.14	Séquence d'images créée pour un canal en fonction du temps.	54
2.15	Exemple de segmentation pour un canal en fonction du temps.	55
2.16	Exemple de lignage pour un canal en fonction du temps	56
2.17	Types d'agrégats protéiques	56
2.18	Représentation géométrique de notre modèle	60
2.19	Coordonnées sphériques d'un point M	61
4.1	Différentes étapes de formation des agrégats protéiques et leur localisation.	146
4.2	Mouvement diffusif des agrégats protéiques chez la levure <i>S. cerevisiae</i>	147
4.3	Diffusion des agrégats dans le cytoplasme de <i>S. cerevisiae</i>	148
4.4	Agrégation protéique à long terme chez la bactérie <i>E. coli</i>	150
4.5	Formes des agrégats protéiques	153



Abréviations

2d : 2 dimensions

3d : 3 dimensions

ADN : Acide DésoxyriboNucléique

ATP : Adénosine TriPhosphate

GFP : Green Fluorescent Protein (protéine fluorescente verte)

Hsp : Heat Shock Protein (protéine de choc thermique)

IBM : Individual Based Model (Modèle Basé Individu)

IbpA : Inclusion Body Protein A

IbpB : Inclusion Body Protein B

LPA : Large Polar Aggregate

MSD : Mean Square Displacement (carré moyen des déplacements)

NA : Numerical aperture (ouverture numérique)

PrP^c : Protein Prion Control

PrP^{sc} : Protein Prion Scrapy

PDMS : Polydimethylsiloxane

PSF : Point Spread Function (fonction d'étalement du point)

RFP : Red Fluorescent Protein (protéine fluorescente rouge)

SMA : Système Multi-Agents

SPT : Single Particle Tracking

YFP : Yellow Fluorescent Protein (protéine fluorescente jaune)

Introduction : agrégation protéique et dégénérescence cellulaire

Cette introduction portera sur deux mécanismes fondamentaux pour les protéines dans la cellule qui sont le repliement des protéines et l'agrégation. L'agrégation protéique est un mécanisme impliqué dans de nombreuses maladies liées au vieillissement chez l'Homme, mais qu'en est-il chez les organismes unicellulaires, organismes très utilisés en recherche médicale car on dispose de nombreuses connaissances et outils moléculaires et génétiques. Nous nous intéresserons également au mécanisme principal régissant la localisation des molécules : la diffusion.

Sommaire

I	L'agrégation protéique	2
I.1	De la conformation native au repliement incorrect	2
I.2	Les différents types d'agrégats	12
I.3	Vieillessement et agrégation protéique chez les organismes unicellulaires	16
II	La diffusion	29
II.1	Histoire de la diffusion	29
II.2	Caractérisation de la diffusion	33
II.3	Mesures de la diffusion	35
III	Objectifs	39

I L'agrégation protéique

Les protéines sont impliquées dans tous les processus biologiques d'un organisme et jouent donc un rôle clé. Chacune des protéines doit se replier dans une structure tridimensionnelle unique et spécifique qui détermine sa fonction dans l'organisme, la conformation native. Le repliement des protéines est un problème central en biologie. Les protéines nouvellement synthétisées émergent du ribosome dans un environnement encombré composé de polypeptides et molécules très diverses. Dans ces conditions d'encombrement et de par la vitesse de traduction relativement lente (15 à 75 secondes pour une protéine de 300 acides aminés [Hartl and Hayer-Hartl, 2009]), les protéines naissantes sont très vulnérables à un repliement anormal (misfolding) et à l'agrégation qui sont deux événements empêchant l'acquisition d'une conformation native correcte et concernent environ 30% des protéines [Schubert et al., 2000]. Chez la bactérie *Escherichia coli*, la vitesse de synthèse peut atteindre 20 acides aminés par seconde [Uemura et al., 2008] et une chaîne polypeptidique émerge du ribosome toutes les 35 secondes [Baneyx and Mujacic, 2004].

I.1 De la conformation native au repliement incorrect

La structure d'une protéine fait intervenir quatre niveaux hiérarchiques. La structure primaire est le niveau d'organisation le plus élémentaire et correspond à la séquence d'acides aminés. Le second niveau est la structure secondaire qui correspond à l'enchaînement de liaisons peptiques possédant des angles particuliers ainsi que des liaisons hydrogènes particulières [Ramachandran and Sasisekharan, 1968]. La structure secondaire implique la présence d'hélices α et de feuillets β . La structure tertiaire correspond à l'organisation tridimensionnelle ou conformation native de la protéine et est biologiquement active. La structure de la protéine est organisée de façon à exposer les résidus polaires et enfouir les parties hydrophobes. Le dernier niveau est la structure quaternaire qui résulte de l'assemblage de plusieurs chaînes polypeptidiques. La conformation d'une protéine repose sur l'équilibre entre différentes forces d'interaction qui favorisent la forme repliée (interactions hydrophobes, liaisons hydrogènes, interactions ioniques, interactions de van der Waals) et l'entropie conformationnelle de la chaîne polypeptidique qui s'oppose au repliement [Chan and Dill, 1990].

I.1.1 Le repliement des protéines

En 1973, Anfinsen [Anfinsen, 1973] montre *in vitro* que beaucoup de protéines dénaturées se replient spontanément dans leur forme native. Ils montrent que la structure tertiaire des protéines est déterminée par la séquence primaire et donc l'enchaînement des acides aminés. De plus, selon Anfinsen, *l'état natif correspond à la structure caractérisée par l'énergie libre la plus basse dans des conditions données*".

Comment une protéine recherche-t-elle sa conformation native parmi toutes les conformations possibles? Considérons une protéine de 100 acides aminés et émettons l'hypothèse que chaque résidu peut prendre deux conformations possibles. Il y a ainsi 2^{100} conformations possibles. En admettant que le temps nécessaire pour passer d'une conformation à l'autre est de 10^{-11} secondes (temps qui correspond au mouvement de vibration des liaisons atomiques, en accord avec les lois de la physique), il faudrait entre 10^{19} et 10^{11} années afin de passer toutes les conformations

possibles [Dobson and Karplus, 1999], ce qui est incompatible avec l'échelle du temps de vie. Le temps de doublement de la bactérie *E.coli* étant de 20 à 30 min, il est évident que d'autres mécanismes interviennent afin de faciliter le repliement des protéines. En effet, le repliement des protéines se déroule sur une échelle de temps variant de quelques microsecondes à quelques minutes [Kubelka et al., 2004].

Ce paradoxe a été évoqué pour la première fois par Levinthal (Figure 1.1 A, page 3) en 1968 [Levinthal, 1968] qui supposait qu'une protéine ne cherche pas sa structure de manière aléatoire mais qu'elle suit des chemins de repliement. A chaque étape, elle adopte une structure qui mène vers un minimum local d'énergie.

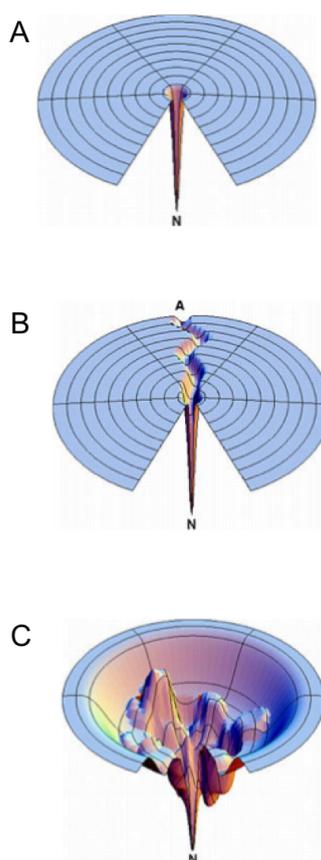


FIGURE 1.1 – Les différentes hypothèses de chemins de repliement d'une protéine. A) Le paradoxe de Levinthal. N est la conformation native. La chaîne polypeptidique cherche au hasard sa conformation native ce qui peut-être plus long que l'âge présumé de l'univers. **B)** Existence d'un chemin préférentiel permettant de guider la recherche de la conformation native. **C)** Existence de barrières énergétiques permettant de trouver le minimum énergétique et donc la conformation native en un minimum de temps.

Face au paradoxe de Levinthal, deux solutions possibles :

- l'existence d'intermédiaires de repliement indiquant un chemin préférentiel pour l'acquisition de la conformation native [Baldwin, 1997] (Figure 1.1 B, page 3).

- l’existence d’un paysage énergétique en entonnoir dans lequel la protéine recherche son état natif [Dill and Chan, 1997] (Figure 1.1 C, page 3).

I.1.2 Paysage énergétique et conformation native

Initialement, le repliement des protéines était décrit comme une succession d’étapes permettant d’acquies la conformation native. Les formes non repliées ou partiellement repliées sont hétérogènes d’un point de vue de leur conformation et il est donc très probable qu’il n’existe pas qu’un seul chemin pour atteindre la conformation native. Les modèles concernant le repliement des protéines [Wolynes et al., 1995] évoluent désormais vers l’existence d’un paysage énergétique dans lequel la chaîne polypeptidique rechercherait sa conformation native. Le repliement peut ainsi être décrit comme une série de transitions entre des états d’énergie ayant pour effet de guider la protéine vers l’état conformationnel de plus basse énergie à savoir la conformation native. Cependant, d’un état à l’autre les différences d’énergie sont très faibles et il existe donc un risque permanent de repliement dans une forme incorrecte même pour les protéines ayant acquis leur conformation native. Le risque est d’autant plus élevé lors de l’exposition à un stress (thermique, osmotique, pH, etc.). La chaîne polypeptidique cherche ainsi au hasard sa conformation native en ayant pour priorité une minimalisation de l’entropie [Chiti and Dobson, 2009] [Jahn and Radford, 2005] (Figure 1.2, page 4). Du fait que les conformations natives sont plus stables que les conformations non-natives, cela réduit le nombre de conformations possibles.

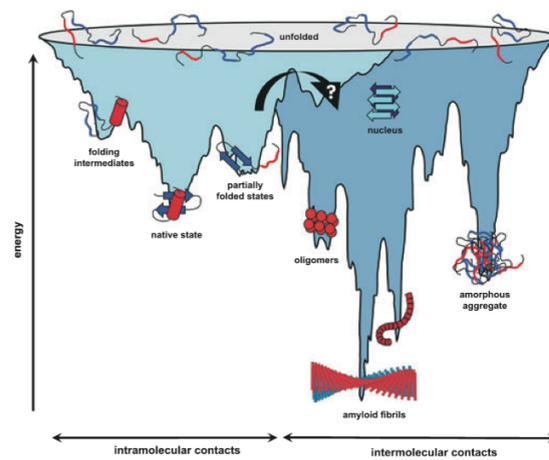


FIGURE 1.2 – **Paysage énergétique en regard du repliement et de l’agrégation protéique.** La surface indique une multitude de conformations possibles qui convergent par un effet d’entonnoir vers la conformation native ou la formation d’agrégats. (Figure extraite de [Jahn and Radford, 2005])

L'existence d'intermédiaires conformationnels semble dépendante de la longueur de la protéine. Pour des protéines de grandes tailles (supérieure à 120 acides aminés, environ 13 kDa), de nombreuses études ont décrits l'existence d'intermédiaires [Kim and Baldwin, 1990] [Jamin and Baldwin, 1998]. En revanche, pour des protéines de taille inférieure à 100 acides aminés soit environ 12 kDa, aucun intermédiaire n'est détectable par les expériences de cinétique de repliement [Krantz et al., 2002] et cela peut s'expliquer par leur cinétique de repliement très rapide de l'ordre de la microseconde.

La stabilité d'une protéine est déterminée par la différence d'énergie libre entre la conformation native et un état non replié. Les états de transition et les barrières énergétiques déterminent ainsi la durée de vie et la stabilité cinétique des intermédiaires. L'entropie a un rôle fondamental en tant que force principale dirigeant le processus de repliement. Le délicat équilibre entre l'enthalpie et l'entropie fait que l'état replié d'une protéine, stable dans certaines conditions, est instable dans d'autres. L'état de transition correspond à la barrière d'énergie qui sépare l'état non replié de l'état replié. Dans l'état de transition, les interactions qui favorisent le repliement, atteignent un seuil critique qui permet de compenser la perte d'entropie associée à la formation des structures natives. La différence entre configuration native et non native est la présence d'interactions intra-moléculaires qui sont principalement de nature hydrophobe. Certains intermédiaires peuvent être piégés dans des états de basse énergie qui sont plus stables. Il en résulte une conformation mal repliée et une protéine non fonctionnelle ou encore la formation d'agrégats favorisée par des interactions intramoléculaires [Chiti and Dobson, 2009] [Jahn and Radford, 2005] (Figure 1.2, page 4).

I.1.3 Effet de l'encombrement moléculaire sur le repliement des protéines

Les protéines émergent du ribosome dans un environnement encombré contenant de nombreuses macromolécules, protéines et polypeptides diverses. Cette caractéristique est ubiquitaire de l'intérieur des cellules. On parle d'environnement encombré et non d'environnement concentré [Minton, 1992] [Zimmerman and Trach, 1991] car il n'y a pas qu'une seule espèce de molécules présente à concentration élevée mais un ensemble de molécules dont chacune est présente à faible concentration. Toutes ces macromolécules occupent un volume significatif dans la cellule qui varie entre 7 et 40% selon le milieu et représente une concentration entre 200 et 300 g/l [Ellis, 2001a] [Zimmerman and Trach, 1991].

L'encombrement moléculaire décrit le fait que la concentration totale des macromolécules dans la cellule est si élevée qu'une portion significative du volume est occupé physiquement et est indisponible pour d'autres molécules.

Dans une cellule, on distingue un continuum d'interactions entre les macromolécules comprenant des interactions spécifiques qui sont propres à chaque protéine et dépendent de la structure des protéines et des interactions non spécifiques qui dépendent des propriétés globales (charge nette ou forme globale) des molécules et des acides aminés exposés.

Les molécules n'étant pas interpénétrables, l'encombrement moléculaire induit des interactions stériques non-spécifiques et répulsives entre les macromolécules. L'encombrement va ainsi mo-

difier les équilibres des réactions telles que le repliement des protéines, la dimérisation ou encore la formation d'amyloïdes. On observe une augmentation des associations entre protéines [Rivas et al., 1999] et des monomères entre eux [Rivas et al., 2001].

Les premières études théoriques semi-quantitatives afin de quantifier les effets de l'encombrement cytoplasmique ont été réalisées dans les années 1980 par Minton [Minton, 1983]. Il a été suggéré que l'encombrement pourrait avoir un impact important sur la stabilité des protéines, les interactions ou encore le repliement des protéines principalement à cause de l'effet de volume exclu.

Le volume exclu est une interaction non spécifique et représente le volume inaccessible à une molécule du à la présence d'autres macromolécules. La figure 1.3 (page 6) représente la notion de volume exclu. Les macromolécules d'encombrement sont représentées par des sphères de rayon r_s et colorées en noir. Le volume total occupé par ces sphères est de 30%. Le volume libre, non occupé par les macromolécules, dans ce cas est donc de 70%. Si l'on cherche maintenant à intégrer une molécule de rayon r_t , le volume accessible correspond à la région de l'espace dans lequel on peut mettre le centre de la molécule. Si $r_t \ll r_s$ (Figure 1.3 A, page 6), le volume accessible correspond quasiment au volume libre en bleu. Dans le cas où $r_t \sim r_s$, le volume accessible à la particule est très fortement réduit par rapport au volume libre (Figure 1.3 B, page 6).

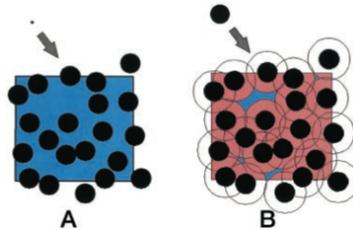


FIGURE 1.3 – **Notion de volume exclu.** En rose et noir se trouvent les volumes exclus et en bleu le volume accessible à des macromolécules sphériques. **A)** Si $r_t \ll r_s$, le volume accessible correspond quasiment au volume libre en bleu. **B)** Dans le cas où $r_t \sim r_s$, le volume accessible à la particule est très fortement réduit par rapport au volume libre (Figure extraite de [Minton, 2001]).

L'effet de l'encombrement sur la vitesse de repliement va dépendre du degré d'encombrement. Les interactions non spécifiques entre les macromolécules contribuent de façon significative à l'énergie libre totale. Il a été montré que de hautes concentrations de macromolécules qui ne participent pas directement aux réactions induisent des changements d'un ordre de grandeur ou plus sur les vitesses et les équilibres des réactions et notamment sur le repliement des protéines [Hatters et al., 2002] [Silow and Oliveberg, 2003]. De plus, l'encombrement moléculaire contribue à augmenter localement la concentration protéique et donc à favoriser l'agrégation par une augmentation de la probabilité de rencontre.

En 1999, van den Berg [van den Berg et al., 1999] montre que l'encombrement moléculaire favorise l'agrégation du lysozyme sous sa forme réduite (le repliement est plus lent que pour la forme oxydée) et empêche tout repliement correct tandis que la forme oxydée du lysozyme est grandement affectée par l'encombrement moléculaire mais le repliement correct reste possible.

Dans ces conditions d'encombrement moléculaire, les protéines nouvellement synthétisées sont spécialement vulnérable à un mauvais repliement et à l'agrégation qui sont deux réactions entravant l'acquisition de la conformation native.

I.1.4 Pourquoi les protéines s'agrègent ?

Le repliement des protéines est une réaction du premier ordre tandis que l'agrégation protéique est une réaction du second ordre dépendante de la concentration protéique. Afin d'être fonctionnelle, une protéine doit acquérir sa conformation native. Pour cela elle passe au travers de plusieurs intermédiaires structuraux (Figure 1.4, page 8). Ces intermédiaires ont pour caractéristique commune d'exposer les régions hydrophobes qui sont normalement enfouies dans la conformation native. Il convient de distinguer mauvais repliement et agrégation qui sont deux notions différentes [Ellis and Minton, 2006]. Le mauvais repliement concerne des chaînes polypeptidiques qui ont acquis une conformation partiellement repliée suffisamment stable mais ne sont pas capables d'acquérir leur conformation native. Ces protéines mal repliées ont une susceptibilité à l'agrégation très variable [Dill and Chan, 1997] (Figure 1.4, page 8). L'exposition des parties hydrophobes favorise l'agrégation. Le piégeage d'autres protéines au sein d'agrégats protéiques a été suggéré comme une des raisons possibles de toxicité associée aux maladies neurodégénératives [Nucifora et al., 2001] [Steffan et al., 2000]. Les agrégats protéiques présentent différents niveaux d'organisation structurale. Les fibres amyloïdes sont très structurées par la présence de feuilletts β et ne peuvent pas être dégradées. En revanche, les agrégats moins ordonnés maintiennent la possibilité d'être replié correctement ou encore dégradé.

Plusieurs facteurs intrinsèques et extrinsèques causent l'agrégation protéique [Lindner and Demarez, 2009].

Quels sont les facteurs intrinsèques ?

- Des mutations dans la séquence primaire de la protéine entraînent une tendance de la protéine affectée à mal se replier et s'agréger [Mogk et al., 2003] et éventuellement à provoquer des maladies. Ces taux d'erreurs transcriptionnelles sont de l'ordre de 10^{-5} à 10^{-6} erreurs par nucléotide [Blank et al., 1986]. De plus, des mutations au sein du système contrôle qualité des protéines peuvent aussi entraîner l'agrégation. Par exemple, une mutation de la chaperonne Hsp α cristalline provoque la cataracte [Andley, 2006].
- Des erreurs de traductions [Mogk et al., 2003] [Drummond and Wilke, 2008]. Le taux d'erreur lors de la traduction est d'environ 10^{-3} à 10^{-4} erreurs par nucléotide [Loftfield and Vanderjagt, 1972] ce qui implique qu'environ 20% des protéines ne sont pas traduites correctement et vont augmenter leur propension à l'agrégation.
- Les stress environnementaux tels que le stress oxydatif [Stadtman and Levine, 2000] ou le choc thermique [Parsell et al., 1994].
- Une saturation du système contrôle qualité des protéines favorise également l'accumulation d'agrégats protéiques.

Quels sont les facteurs extrinsèques ?

Les facteurs environnementaux peuvent promouvoir l'agrégation et des maladies neurodégénératives [Coppede et al., 2006]. Plusieurs études démontrent que les métaux lourds tel que le fer ou le zinc favorisent l'agrégation [Uversky et al., 2002]. Les pesticides pourraient également être

impliqués dans la maladie de Parkinson [Langston et al., 1983].

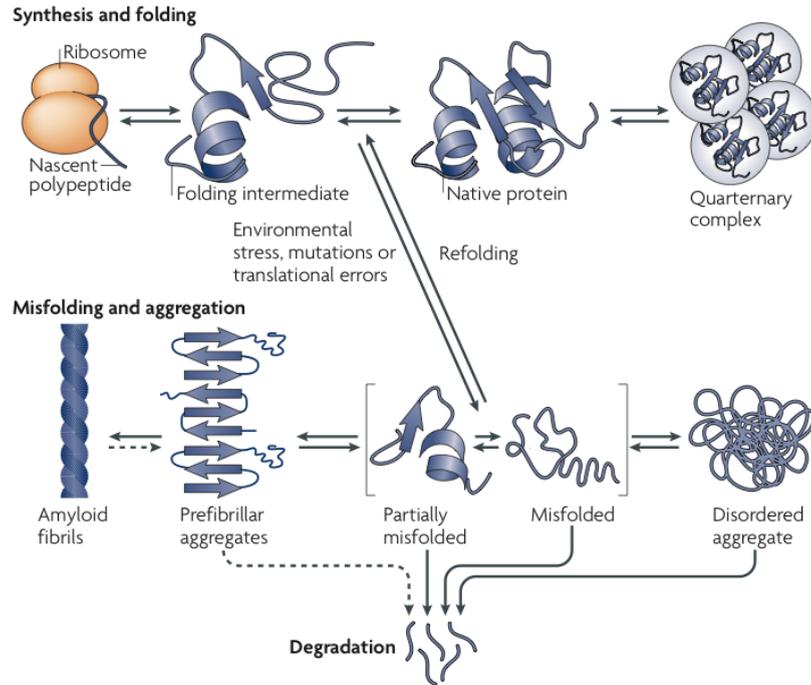


FIGURE 1.4 – **Vue d'ensemble du repliement et de l'agrégation protéique.** Avant d'obtenir sa conformation tridimensionnelle définitive, une protéine passe par différents intermédiaires. Des mutations dans la séquence du nouveau polypeptide synthétisé ou des erreurs de traductions peuvent entraîner un mauvais repliement (misfolding) de la protéine. Il existe différents degrés de repliement incorrect. Une protéine mal repliée peut-être repliée dans sa conformation native ou dégradée ou encore peut s'agréger. (Figure extraite de [Tyedmers et al., 2010])

L'agrégation protéique a des conséquences importantes pour la cellule en termes de :

- réduction du recyclage des acides aminés
- recrutement et immobilisation des protéines chaperonnes et protéases
- "kidnapping" de protéines fonctionnelles
- formation de polypeptides toxiques
- perte de fonction de la molécule

Afin de prévenir le mauvais repliement et l'agrégation, la cellule déploie différentes stratégies. Une première stratégie protectrice semble être la présence de résidus glycine et proline dans la séquence primaire [Monsellier and Chiti, 2007]. Certains acides aminés favoriseraient plus l'agrégation que d'autres. Une autre perspective suggère que les mécanismes d'agrégation et de repliement sont en compétition [Tartaglia and Vendruscolo, 2010]. Des stratégies de régulation sont également possibles comme la régulation de la la concentration protéique ou de la traduction [Wickner et al., 1999]. Différentes études ont démontré une corrélation entre concentration protéique, solubilité et propension à l'agrégation [Tartaglia et al., 2007]. En fin, la cellule dispose

également d'un puissant système contrôle qualité des protéines qui sera détaillé dans la prochaine section.

I.1.5 Le système contrôle qualité des protéines chez la bactérie

Afin de faire face aux conditions d'encombrement moléculaire et éviter un mauvais repliement protéique ou encore l'agrégation, les organismes depuis les bactéries jusqu'à l'Homme, disposent d'un système contrôle qualité des protéines. Ce système est également très utile face aux conditions hostiles que peuvent rencontrer les organismes. Les bactéries supportent ainsi des écarts de température, des variations de pH et d'osmolarité ou des stress chimiques.

Ce système se compose d'un ensemble de protéines chaperonnes et de protéases qui est nécessaire au repliement *de novo* ainsi qu'au contrôle qualité des protéines [Hartl, 1996] [Bukau and Horwich, 1998]. La fonction des différents composants dans le contrôle qualité dépend de leur capacité à replier correctement ou à dégrader les protéines anormalement repliées. Le contrôle qualité se déroule en conditions de croissance normale pour la cellule mais devient particulièrement important en conditions de stress comme un choc thermique.

La concentration d'une protéine native résulte d'un équilibre entre le taux de synthèse protéique, la dilution et le taux de dégradation. Les situations de stress telles que le choc thermique qui est un des stress les plus étudiés, affectent cet équilibre en diminuant la stabilité du repliement natif ce qui augmente le nombre de molécules repliées anormalement et entraîne ainsi l'agrégation.

I.1.5.1 Les chaperonnes

Les protéines chaperonnes sont des protéines dont la fonction est d'assister d'autres protéines dans leur maturation, en leur assurant un repliement tridimensionnel adéquat et limiter le processus d'agrégation qui peut se produire lors de la recherche de la conformation native, la chaîne polypeptidique étant piégée dans un état d'énergie stable mais n'a pas acquis sa conformation native ou après exposition à un stress environnemental (élévation de la température, stress oxydatif, anoxie, éthanol, dénaturants chimiques ...), ou enfin d'assister le processus de dégradation protéique. Une surexpression de ces protéines chaperonnes augmente la tolérance au stress chez *E. coli* [Thomas and Baneyx, 1998].

Certaines chaperonnes sont exprimées constitutivement tandis que d'autres sont induites à la suite d'un stress. Les chaperonnes induites à la suite d'un choc thermique furent les premières à être découvertes par Ritossa en 1962 et furent nommées protéines de choc thermique : Heat shock protein, Hsp [Ritossa, 1962]. La réponse au choc thermique est une réponse universelle de tous les organismes face à un stress thermique, chimique ou oxydatif. La structure des protéines est en effet sensible à la chaleur, elles se dénaturent et perdent leur action biologique. Lors d'un stress thermique, les protéines se déplient à cause d'une augmentation de l'entropie, qui représente le degré de désordre, bien supérieure à l'enthalpie, qui se définit comme la quantité de chaleur échangée par le système au cours d'une réaction réalisée à pression constante.

Le rôle des protéines chaperonnes est de prévenir les dommages potentiellement causés par une perte de fonction protéique causée par un mauvais repliement tridimensionnel. D'autres chaperonnes sont impliquées dans le repliement de protéines nouvellement synthétisées alors qu'elles sont extraites du ribosome. Bien que la plupart des protéines aient la capacité de se replier correctement sans intervention de chaperonnes, certaines en sont incapables. Lorsqu'un polypeptide

adopte sa conformation native, les acides aminés portant une chaîne latérale hydrophobe (non polaire) se rassemblent au coeur de la protéine afin de s'éloigner de l'eau. Des températures élevées ou divers traitements chimiques dénaturent une protéine entraînant un changement de conformation ou un dépliement de la protéine et lui font perdre sa conformation ainsi que sa fonction.

Les protéines chaperonnes ont une affinité pour les zones hydrophobes accessibles du polypeptide. Cette interaction empêche des associations indésirables avec d'autres protéines et permet la poursuite du processus de repliement.

Chez les bactéries, on distingue trois classes de protéines chaperonnes (Figure 1.5, page 11) :

- Les **Folding chaperonnes** sont constituées par le système DnaK qui inclue les co-chaperonnes DnaJ et GrpE ainsi que le système GroEL/ES qui collabore étroitement avec le système DnaK [Buchberger et al., 1996] [Nishihara et al., 1998]. Elles sont ATP-dépendantes et leur rôle est de limiter et inverser le repliement incorrect ainsi que l'agrégation [Bukau and Horwich, 1998] [Slepenkov and Witt, 2002] [Arsene et al., 2000]. DnaK avec l'aide de sa co-chaperonne DnaJ se lie à la protéine afin d'exposer les régions hydrophobes de la protéine et ainsi par cycles répétés de "binding and release", la protéine est soit repliée correctement soit désagrégée. Le système GroEL/ES [Ellis and van der Vies, 1991] [Gething and Sambrook, 1992], (respectivement chaperonine 60 et 10 chez les eucaryotes) [Hemmingsen et al., 1988], permet de capturer et séquestrer les protéines repliées partiellement dans une cavité centrale de GroEL où le repliement aura lieu [Hartl, 1996] [Hartl and Hayer-Hartl, 2002].

En outre, l'encombrement moléculaire peut expliquer pourquoi il existe des chaperonnes malgré la capacité de beaucoup de protéines à adopter leur conformation native par elle-même [Anfinsen, 1973]. L'encombrement moléculaire faciliterait l'association entre les protéines chaperonnes et leurs cibles [Martin and Hartl, 1997]. L'hypothèse de la cage d'Anfinsen suggère que la chaperonne GroEL augmente son efficacité de repliement des protéines dénaturées en attachant la chaîne polypeptidique partiellement foldée à l'intérieur de sa structure oligomérique afin de la protéger de l'agrégation avec une autre chaîne polypeptidique. Ainsi, les chaperonnes peuvent apparaître comme des "tubes expérimentaux" dans lesquels chaque polypeptidique peut se replier [Ellis, 2001b]. Elles permettent d'isoler les protéines les unes des autres et de combattre l'encombrement moléculaire.

- Les **Holding chaperonnes** sont représentées principalement par les chaperonnes IbpA et IbpB. Elles permettent de maintenir les protéines dans un état intermédiaire, partiellement foldé, dans l'attente de l'intervention des folding chaperonnes. En d'autres termes, IbpA and IbpB protègent les protéines dénaturées contre une agrégation irréversible [Kuczynska-Wisnik et al., 2002]. Ces chaperonnes sont associées à des corps d'inclusion [Allen et al., 1992] [Carrio et al., 2005] [Lindner et al., 2008] ou des protéines endogènes [Laskowska et al., 1996]. Malgré l'homologie de séquence entre ces 2 protéines chaperonnes (50%), elles se comportent très différemment. IbpA est très hydrophobe et est retrouvé dans la fraction insoluble. Elle a également tendance à former des multimères *in vitro* et des fibrilles en l'absence de sa co-chaperonne IbpB [Ratajczak et al., 2010]. Elle interagit avec les agrégats protéiques afin de faciliter le repliement. En revanche, IbpB est principale-

I. L'AGRÉGATION PROTÉIQUE

ment soluble quand elle est produite en absence d'IbpA [Kuczynska-Wisnik et al., 2002]. Actuellement, aucune données ne montre une interaction directe d'IbpB avec les agrégats protéiques. Par contre, elle semble impliquée dans les mécanismes de désagrégation [Ratajczak et al., 2009].

- Les **Disaggregating chaperonnes** dont ClpB est la plus caractérisée, travaillent en collaboration avec le système DnaK et les chaperonnes IbpA/B [Mogk et al., 2003], ces chaperonnes étant nécessaires à l'acquisition de la thermotolérance [Tomoyasu et al., 2001]. ClpB permet de dégrader les protéines repliées incorrectement et donc non fonctionnelles lorsque les tentatives de repliement via les holding et holding chaperonnes ont échoué. Ces chaperonnes sont d'autant plus importantes lors de l'exposition à un stress. Augmenter l'expression du système DnaK et de la chaperonne ClpB est suffisant pour restaurer la thermotolérance dans des cellules dont le gène *rpoH*, codant pour le facteur sigma32 qui est impliquée dans la croissance en condition de choc thermique [Tomoyasu et al., 2001], a été délété.

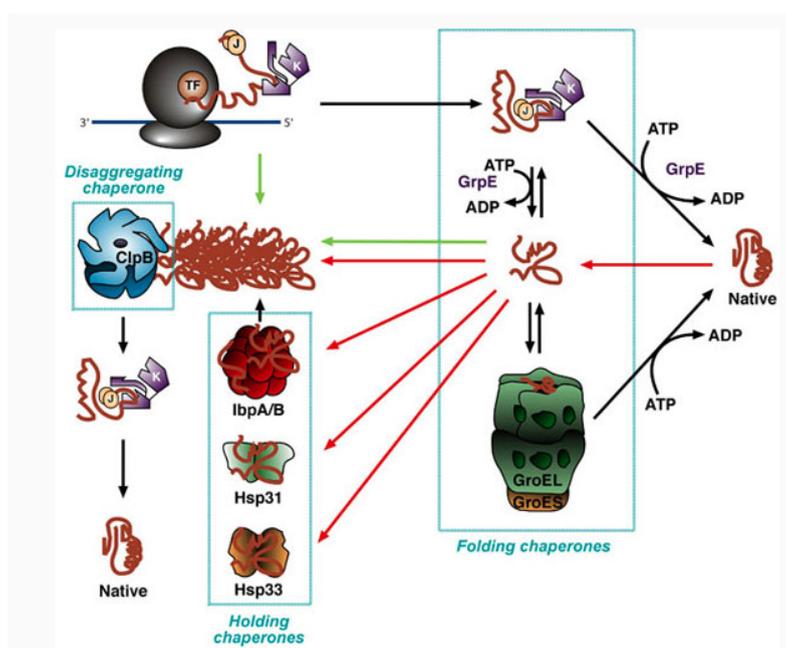


FIGURE 1.5 – Les différents classes de protéines chaperonnes. Le système DnaK inclut les chaperonnes DnaK, DnaJ et GrpE ainsi que les chaperonnes GroEL/S qui assistent les protéines dans l'acquisition de leur conformation native. Les petites protéines de choc thermique IbpA/B se lient aux protéines mal repliées durant la phase d'agrégation et la protéine ClpB déclenche le processus de désagrégation en collaboration avec le système DnaK. (Figure extraite de [Baneyx and Mujacic, 2004])

De part leur thermotolérance accrue par rapport aux autres protéines, les protéines chaperonnes peuvent jouer leur rôle lors d'un stress sans être dénaturée ou s'agréger même si certaines peuvent s'agréger entre elles telle que IbpA. Elles ont ainsi pour rôle d'interagir avec la conformation non-native des autres protéines. Certaines chaperonnes permettent d'éviter l'agrégation et d'autres

ont un rôle d'assistant dans le repliement ou encore la solubilisation des agrégats. L'agrégation est également minimisée par l'action des protéases dont le rôle est de dégrader les protéines mal repliées et donc non fonctionnelles [Huang et al., 2001].

I.1.5.2 Les protéases

L'hydrolyse des protéines qui n'ont pas réussi à atteindre leur conformation native joue un rôle crucial dans le système contrôle qualité, en empêchant l'agrégation de protéines mais aussi en permettant le recyclage des acides aminés. La protéolyse étant irréversible, la dégradation protéique doit être régulée avec attention.

Les principales protéases chez *E. coli* sont ClpP et Lon [Maurizi, 1992] [Rosen et al., 2002]. Ces protéases, ATP dépendantes, reconnaissent les parties hydrophobes exposées et dégradent les polypeptides mal repliés localisés dans la fraction cellulaire soluble mais aussi ceux appartenant à un agrégat protéique.

La dégradation des protéines s'effectue en trois étapes chez la bactérie [Dougan et al., 2002] :

- La reconnaissance du substrat.
- La fixation du substrat à la protéine endommagée.
- Le transfert du substrat dans la "chambre de dégradation" afin d'effectuer la protéolyse.

Lon fut la première à être découverte [Swamy and Goldberg, 1981]. C'est une sérine protéase tétramérique de 87 kDa contenant trois domaines fonctionnels : le domaine N-terminal est impliqué dans la reconnaissance du substrat et la liaison, le domaine central est responsable de l'activité ATPase et le domaine C-terminal a une activité protéolytique.

La protéase CLpP est une protéine organisée en deux heptamères de 23 kDa chacune. Cette protéase forme un complexe avec les protéines ClpA et ClpX qui permettent de transférer les substrats dans le canal central de ClpP.

I.2 Les différents types d'agrégats

Afin d'être fonctionnelle, une protéine nécessite d'avoir une structure tridimensionnelle correcte. Cependant, comme nous l'avons vu précédemment, la recherche de la conformation native dans le paysage énergétique ou l'exposition à un stress tel que le choc thermique ou oxydatif, peut entraîner un repliement incorrect de la protéine avec exposition des parties hydrophobes de la molécule. La protéine peut alors s'agréger avec d'autres molécules mal repliées et perdre ainsi son activité biologique.

Les agrégats protéiques ont été initialement classés grâce à la microscopie électronique. On distingue ceux qui présentent une structure qui s'apparente soit à des corps d'inclusion soit à des amyloïdes. Ces derniers sont impliqués dans de nombreuses maladies neurodégénératives humaines.

I.2.1 Les maladies associées

La conformation native n'est pas seulement une caractéristique essentielle à la conversion de l'information génétique en activité biologique mais c'est aussi une caractéristique clé dans le contrôle et la localisation de cette activité [Dobson and Ellis, 1998]. Ceci implique l'idée que l'échec du repliement des protéines dans la forme attendue crée une protéine avec des propriétés différentes ou inactive et une propension à l'agrégation plus importante. Ces deux facteurs sont impliqués dans de nombreuses maladies humaines [Dobson, 2001] [Azriel and Gazit, 2001] [Chiti and Dobson, 2006], comme la maladie de Creutzfeldt-Jakob ou encore les maladies de type amylose comme la maladie d'Alzheimer [Harper and Lansbury, 1997], mais aussi le diabète de type II [Kahn et al., 1999].

Les fibrilles amyloïdes formées au cours de ces différentes pathologies concernent des protéines différentes et présentent des séquences d'acides aminés différentes. Cependant, elles ont toutes une structure composée de feuillets β croisés dans laquelle le brin β individuel est orienté perpendiculairement à l'axe long de la fibre. Des études ont montré que la formation d'intermédiaires de conformation serait une étape importante dans la formation des fibres amyloïdes [Hornemann and Glockshuber, 1998].

Pendant longtemps, l'agrégation protéique chez les procaryotes a été vue comme un mécanisme non spécifique non relié à la formation d'amyloïdes chez les eucaryotes. Peu d'intérêt leur était porté à l'exception de leur rôle dans la production de protéines recombinantes. Cependant leur formation et structure partagent des caractéristiques communes avec les fibrilles amyloïdes liées aux maladies humaines telles qu'Alzheimer [Wang et al., 2008] [Morell et al., 2008] et en font des modèles très intéressants pour l'étude de l'agrégation protéique.

I.2.2 Les corps d'inclusion

Les corps d'inclusion sont des agrégats protéiques insolubles qui ont été observés pour la première fois chez la bactérie *E. coli* via la surexpression de protéines recombinantes [Carrio et al., 1998]. Un des problèmes associés à la surexpression de protéines dans *E. coli* est leur séquestration dans des agrégats insolubles appelés corps d'inclusion. Les premiers cas reportés, concernaient les chaînes A et B de l'insuline [Williams et al., 1982]. On a d'abord pensé que les corps d'inclusion ne concernaient que les protéines hétérologues exprimées dans *E. coli* mais on observe aussi leur formation lors de la surexpression de protéines endogènes comme par exemple la surexpression de sous-unités de l'ARN polymérase [Gribskov and Burgess, 1983].

On pensait ces corps d'inclusion amorphes, sans ordre, mais il semblerait qu'ils contiennent également des feuillets β comme les amyloïdes [de Groot et al., 2009] que l'on peut retrouver dans certaines maladies humaines [Ventura, 2005].

La protéine chaperonne DnaK possède un rôle clé dans l'extraction et la dégradation des corps d'inclusion [Martinez-Alonso et al., 2009]. Une précédente étude a montré qu'en absence des protéines chaperonnes DnaK ou GroEL, on observe une diminution importante de la croissance de la cellule [Gonzalez-Montalban et al., 2005] et qu'en absence de DnaK, les corps d'inclusion observés sont de taille plus importante dans la cellule mère [Carrio and Villaverde, 2003].

I.2.3 Les amyloïdes

Les amyloïdes sont formés par l'agrégation de protéines solubles avec une structure très ordonnée constituée de feuillets β croisés c'est-à-dire que les feuillets β intermoléculaires sont perpendiculaires à l'axe de la fibre [Chiti and Dobson, 2006] [Sawaya et al., 2007]. La formation des fibres amyloïdes fait principalement intervenir un phénomène de nucléation-polymérisation [Lomakin et al., 1997] [Wetzel, 2006], qui permet dans une première étape la formation d'un noyau stable à l'aide des feuillets β . Une seconde étape permet la croissance du noyau par fixation de nouvelles protéines.

Les fibres amyloïdes sont impliquées dans de nombreuses maladies humaines neurodégénératives telles que les maladies d'Alzheimer [Hartl et al., 2011], Parkinson, et diabète de type II. Chaque maladie est associée à des protéines particulières et des agrégats de ces protéines semblent directement ou indirectement être à l'origine des manifestations pathologiques associées à ces maladies.

Une récente étude [Ma et al., 2012] sur l'effet de l'encombrement moléculaire en regard de la formation de fibres amyloïdes démontre qu'il n'est pas possible d'établir de règles universelles. En effet, ils observent des résultats opposés en fonction des molécules étudiées. L'encombrement moléculaire promeut la formation de fibres amyloïdes de protéines amyloïdogènes telles que la protéine Tau humaine impliquée dans la maladie d'Alzheimer ou la protéine prion PrP humaine. A l'opposé pour des protéines non-amyloïdogènes comme la protéine prion du lapin ou encore le lysozyme du blanc d'oeuf, l'agrégation est inhibée.

Il est à noter que la formation d'amyloïdes ne représente pas toujours un danger pour la cellule. Il a été suggéré que les fibres amyloïdes pouvaient avoir une fonction bénéfique dans le cadre de maladies et représentaient une réponse de défense cellulaire par la séquestration des espèces toxiques dans des fibres amyloïdes non toxiques [Hardy and Selkoe, 2002]. Les fibres amyloïdes jouent un rôle important un rôle de stockage d'hormones protéiques et peptiques dans des granules sécrétoires du système endocrinien [Maji et al., 2009].

I.2.4 Les prions et les prionoides

Le terme prion a été employé pour la première fois par Prusiner [Prusiner, 1982] pour désigner les agents infectieux responsables de certaines maladies neurodégénératives chez l'Homme et les animaux. L'agent infectieux serait la protéine PrP. La forme cellulaire normale, PrP^c (protein prion control) est une protéine membranaire dont la demi-vie est de 2 à 3 heures *in vivo* [Baldwin et al., 1995]. La forme infectieuse, PrP^{sc} (protein prion scrappy), provoque l'apparition chez son équivalent cellulaire normal de changements de conformation importants qui conduisent à la formation d'agrégats protéiques cytotoxiques constitués presque uniquement de la forme infectieuse. Les deux isoformes qui ont des structures primaires identiques se différencient par des propriétés physico-chimiques distinctes. La forme prion de la protéine PrP^{sc} est plus stable et résistante aux protéases. Ces variations s'expliquent par leur structure tridimensionnelle. La protéine PrP^c est composée à 42% d'hélices α et de très peu de feuillets β tandis que la protéine PrP^{sc} présente 43% de feuillets β et 30% d'hélices α [Pan et al., 1993].

L'expression du gène PRNP (codant pour la protéine PrP) n'est pas modifiée et continue d'exprimer une protéine dont la structure primaire est normale. Ce sont les conformations secondaires,

tertiaires et quaternaires qui sont altérées afin d'avoir une protéine prion.

Les prions sont une sous-classe des amyloïdes et jouent un rôle important car ils génèrent de nouveaux phénotypes en modifiant la fonction normale des protéines et la réponse cellulaire. En plus d'avoir une structure stable et ordonnée grâce aux feuilletts β , ils ont la capacité de s'auto-assembler, sont hérités verticalement lors de la division cellulaire et peuvent être également transférés horizontalement à un autre organisme et agissent alors comme agent infectieux [Krishnan and Lindquist, 2005] [Cobb and Surewicz, 2009].

La chaperonne GroEL pourrait faciliter le recrutement de PrP^c au sein d'agrégats de PrP^{sc} et promeut également l'agrégation PrP^c en absence de template [DeBurman et al., 1997] [Stockel and Hartl, 2001].

I.2.4.1 Les prions chez la levure *Saccharomyces cerevisiae*

Chez la levure *Saccharomyces cerevisiae*, les deux premières protéines découvertes dont le comportement était proche des prions de mammifères sont les protéines Sup35p [Cox, 1965] et Ure2p [Lacroute, 1971]. Cependant, ce n'est qu'en 1994 que Wickner montra leur comportement de type prion [Wickner, 1994]. Depuis de nombreux autres prions ont été décrits : [PIN⁺] [Derkatch et al., 2001], [MOT3] [Alberti et al., 2009], etc.

[PSI⁺] est la forme prion de la protéine Sup35, un composant du complexe de terminaison de la traduction chez la levure. C'est le prion le plus étudié. Les agrégats [PSI⁺] séquestrent la protéine soluble Sup35 ce qui a pour conséquence une défaillance de la terminaison de la traduction. Ce prion peut être dénaturé au moyen de guanine hydrochloride. Une superexpression de la protéine chaperonne Hsp104 permet également de traiter ce prion. Ceci suggère que la chaperonne Hsp104 doit être exprimée à un niveau intermédiaire afin de permettre la propagation du prion [Crow and Li, 2011].

[URE3] est la forme prion de la protéine Ure2, une protéine impliquée dans la régulation du métabolisme de l'azote. Ce prion peut également être dénaturé par de la guanine hydrochloride.

Ces deux prions possèdent des caractéristiques communes comme des régions riches en Q/N (glutamine/asparagine) qui sont indispensables pour la maintenance et la propagation du prion [Crow and Li, 2011] ou encore une curabilité réversible. En effet, le taux d'agrégation dépend du contenu en arginine et glutamine [Fuentelba et al., 2010]. En outre, leur désagrégation est facilitée lorsque les résidus asparagine sont mutés en résidus glutamine car cette mutation entraîne une diminution de structuration de l'agrégat [Halfmann et al., 2011]. Chez la levure, les prions sont dépendants des chaperonnes et notamment de la Hsp104 qui est responsable de la propagation du prion en le découpant en petits fragments et permet ainsi la formation de nouveaux agrégats à partir de ces fragments. En absence de la chaperonne Hsp104, les agrégats "prion" atteignent une taille trop importante pour être transmis à la cellule fille lors de la division [Kryndushkin et al., 2003].

Sous sa forme anormale, le prion de levure s'assemble en fibres et est capable de contaminer une levure saine. Ce prion est inoffensif pour l'Homme et en reste d'autant plus intéressant comme modèle d'étude des prions. Les prions sont souvent associés à des maladies mais il a été montré

récemment que chez la levure, ils peuvent être bénéfiques. Ils agissent alors comme des éléments génétiques [Uptain and Lindquist, 2002] et interrupteurs moléculaires pour contrôler certaines voies cellulaires et peuvent influencer l'adaptation lors de l'exposition à un stress qui implique la survie de l'organisme [Halfmann et al., 2012].

I.2.4.2 Les prions chez la bactérie *Escherichia coli*

A l'heure d'aujourd'hui, aucune protéine prion n'a été observée chez la bactérie *E. coli*. Cependant, elle pourrait constituer un modèle de choix pour l'étude de ceux-ci.

Quand le caractère infectieux est absent on parle de prionoïde. Il y a transmission verticale mais absence d'infection [Aguzzi and Rajendran, 2009].

Récemment, l'équipe de Giraldo [Fernandez-Tresguerres et al., 2010] a développé un modèle synthétique permettant d'étudier la transmissibilité des protéinopathies amyloïdes chez la bactérie *E. coli*. En utilisant la protéine initiant la réplication RepA du plasmide pPS10 de *Pseudomonas savastanoi* fusionnée à la protéine fluorescente rouge RFP et en modifiant le domaine terminal WH1, ils observent la formation *in vitro* d'agrégats de type amyloïdes [Giraldo, 2007]. La protéine de réplication RepA est un répresseur de sa transcription ou un inhibiteur de sa réplication selon qu'elle se trouve sous la forme de dimères ou de monomères respectivement. La conversion du complexe RepA-WH1 de dimère à monomère entraîne un repliement partiel de certaines sous-unités et favorise l'agrégation. Ce complexe se comporte comme certains prions chez les mammifères mais n'est pas infectieux et est ainsi caractérisé comme prionoïde. L'introduction d'une mutation ponctuelle dans le domaine WH1 a permis la création d'un mutant (A31V) dont le niveau d'agrégation était augmenté [Giraldo, 2007] et ainsi d'étudier et caractériser l'impact de ce type d'agrégats (cette étude fait partie de ce projet de thèse).

Par ailleurs, en utilisant le prion [*PSI*⁺], prion de la protéine Sup35, Garrity et ses collaborateurs [Garrity et al., 2010] démontrent que le cytoplasme de la bactérie *E. coli* contient tous les éléments moléculaires nécessaires qui permettent la formation d'agrégats de protéines sous forme prion.

Ces résultats permettent d'envisager l'utilisation de la bactérie *E. coli* comme modèle d'étude de la formation des prions.

I.3 Vieillessement et agrégation protéique chez les organismes unicellulaires

Comme nous l'avons vu précédemment, l'agrégation protéique est impliquée dans de nombreuses maladies humaines liées au vieillissement. Mais qu'en est-il chez les organismes unicellulaires? Les organismes unicellulaires sont également sujets au vieillissement même si pendant longtemps, on pensait les bactéries immortelles. La levure *S. cerevisiae* fut le premier organisme unicellulaire étudié à montrer des signes de vieillissement. Jusqu'à récemment, on suggérait que la présence d'un stade juvénile était une condition nécessaire au vieillissement et que les levures telles que *S. pombe* ou les bactéries à division symétrique comme *E. coli* ne vieillissaient probablement pas car il n'était pas possible de différencier les parents des descendants. Cependant, d'autres facteurs tels que l'agrégation protéique permettent de différencier la cellule mère de la

cellule fille et sont impliqués ainsi dans le vieillissement des organismes.

I.3.1 Le vieillissement chez les organismes unicellulaires

I.3.1.1 Le vieillissement chez la levure *Saccharomyces cerevisiae*

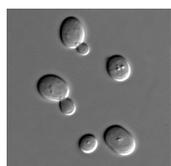


FIGURE 1.6 – La levure *S. cerevisiae*

La levure *Saccharomyces cerevisiae* est l'organisme eucaryote le plus simple appartenant au règne des champignons et à la famille des Saccharomycètes. De forme ronde, elle mesure de 5 à 10 μm de diamètre. Une cellule mère donne naissance à une cellule fille de taille plus petite qui donnera naissance à son tour après avoir atteint la taille adulte (Figure 1.6, page 17). On notera qu'avec l'âge il devient difficile de distinguer la cellule mère de la cellule fille car la taille de cette dernière augmente avec l'âge [Kennedy et al., 1994].

La toute première étude sur le vieillissement de *S. cerevisiae* a été publiée il y a plus de 50 ans et montrait qu'une cellule avait un nombre de divisions fini [Mortimer and Johnston, 1959]. Ceci fût la preuve qu'une levure est mortelle. Dans cette étude, le vieillissement a été analysé selon la durée de vie réplivative.

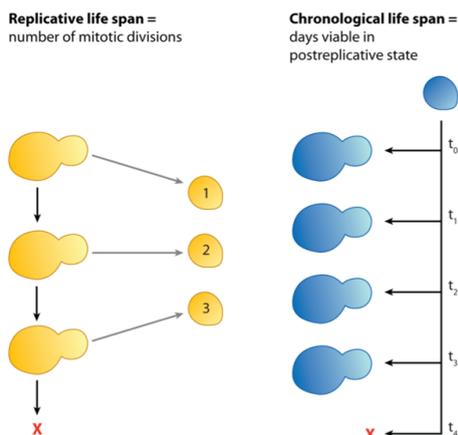


FIGURE 1.7 – Durée de vie chronologique (chronological lifespan) et réplivative (replicative lifespan) chez la levure *S. cerevisiae*. La durée de vie chronologique est le temps pendant lequel une cellule peut survivre sans entrer en division et la durée de vie réplivative est le nombre de fois où une cellule peut entrer en division mitotique. (Figure extraite de [Steinkraus et al., 2008])

Chez les organismes unicellulaires, il existe deux méthodes de mesure de la longévité (Figure 1.7,

page 17) :

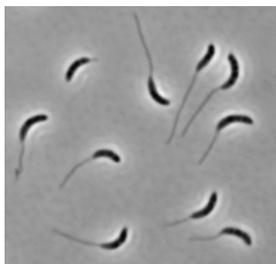
- **La longévité réplivative** : la durée de vie réplivative se mesure en nombre de divisions accomplies par une cellule avant sa mort [Jazwinski, 1990].
- **La longévité chronologique** : la durée de vie chronologique se mesure en unité de temps lorsque les cellules se trouvent dans un état post-mitotique et qu'elles ne se divisent plus [Fabrizio and Longo, 2003].

Les travaux de Qin [Qin and Lu, 2006] démontrent la présence de variations génotypiques dans le vieillissement réplivative et chronologique entre des isolats naturels et des souches de laboratoire. Ils observent un nombre de divisions fini variant de 19 à 36 divisions et une longévité variant de 3 à 17 jours. La longévité réplivative est plus longue de 30% chez les isolats naturels que les souches de laboratoire. En revanche, la longévité chronologique est significativement plus courte chez les isolats naturels. Aucune corrélation n'a pu être établie entre longévité réplivative et chronologique.

La particularité de la levure *S. cerevisiae* de se diviser asymétriquement en a fait un modèle de choix pour l'étude de la durée de vie réplivative grâce à sa micromanipulation. En effet, le vieillissement réplivative se mesure en supprimant par microdissection à l'aide d'un microscope optique les cellules filles qui se distinguent aisément de la cellule mère par leur petite taille [Steinkraus et al., 2008].

Lors de la division, la cellule fille gagne en capacité réplivative par rapport à la cellule mère dont elle est issue ; la cellule fille est plus jeune que la cellule mère [Kennedy et al., 1994]. Toutefois, elle ne récupère pas la totalité du potentiel réplivative maximal. Une fille issue d'une mère jeune vit plus longtemps qu'une fille issue d'une cellule mère vieille [Kennedy et al., 1994].

I.3.1.2 Le vieillissement chez la bactérie *Caulobacter crescentus*



La bactérie aquatique *C. crescentus* mesure 2 à 3 μm de longueur pour 0.7 μm de section. En 2003, Ackermann [Ackermann et al., 2003] montre pour la première fois que les bactéries sont aussi sujettes au vieillissement réplivative.

FIGURE 1.8 – La bactérie *C. crescentus*

La bactérie *C. crescentus* est une bactérie à division asymétrique avec une phase juvénile. Les cellules sont successivement mobiles durant 30 à 45 minutes avant de perdre leur flagelle, puis fixées par un pédoncule à un substrat solide (Figure 1.9 A, page 19). Les cellules propagatrices, propulsées par un flagelle, ne peuvent pas se diviser. Seules les cellules pédonculées peuvent le faire. Leurs descendants sont mobiles, mais se fixent au bout d'un certain temps. Au début, les cellules mères se divisent régulièrement et rapidement mais avec le temps, les divisions s'espacent et certaines cellules finissent par arrêter de se diviser (Figure 1.9 B, page 19).

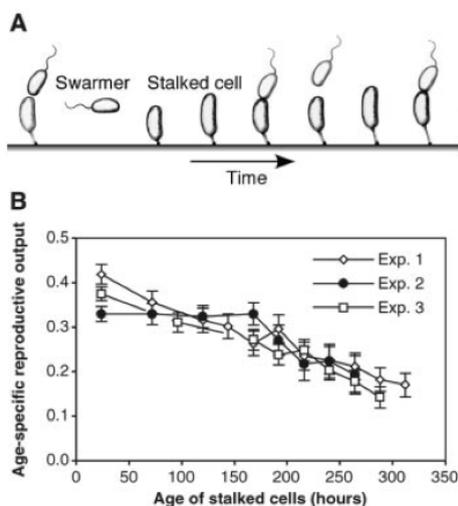


FIGURE 1.9 – Cycle de vie chez la bactérie *C. crescentus* et vieillissement. **A)** Swarmer cell est une cellule mobile qui perd son flagelle après 30-45 minutes pour se différencier en stalked cell ou cellule fixée qui peut se diviser. **B)** Nombre de descendants produits par les cellules fixées (stalked cell) par heure et en fonction de l'âge. Les cellules produisent de moins en moins de descendants avec le temps et sont sujettes au vieillissement. (Figure extraite de [Ackermann et al., 2003]).

I.3.1.3 Le vieillissement chez la levure *Schizosaccharomyces pombe*

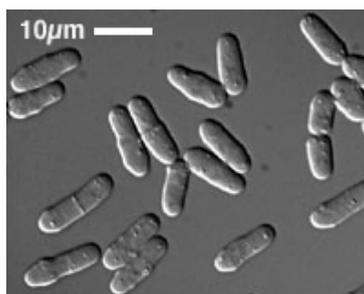


FIGURE 1.10 – La levure *S. pombe*

La levure *S. pombe*, aussi appelée levure à fission est un organisme unicellulaire eucaryote qui se présente sous forme cylindrique de longueur 6 à 10 µm. En 1999, Barker *et al.* [Barker and Walmsley, 1999] démontrèrent des signes de vieillissement réplcatif. Cependant, leurs résultats sont plutôt controversés car ils sont partis du postulat d'une asymétrie morphologique alors que cette bactérie se divise par fission binaire, donnant naissance à deux cellules filles identiques d'un point de vue morphologique. De plus, leurs données ne permettaient pas de définir si le phénotype d'asymétrie observé était un phénotype liée au vieillissement. En revanche, ils ont démontré que *S. pombe* avait un nombre fini de divisions. D'autres expériences ont été réalisées par M. Coelho et I. Tolic-Norrelykke [Coelho, 2012] et il ne leur a pas été possible

de démontrer des signes de vieillissement réplcatif dans des cellules uniques. La question du vieillissement chez des organismes à division symétrique reste ainsi d'actualité.

I.3.1.4 Le vieillissement chez la bactérie *Escherichia coli*

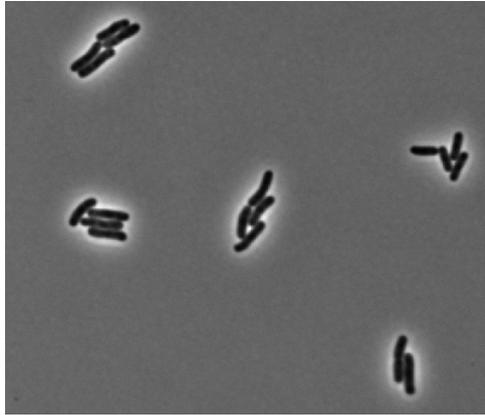


FIGURE 1.11 – La bactérie *Escherichia coli* est une bactérie en forme de bâtonnet qui mesure entre 3 et 6 μm de long et 1 μm de section.

Jusqu'à il y a peu de temps, on pensait que la bactérie *E. coli* était immortelle car celle-ci se divisait de façon symétrique et donnaient ainsi naissance à deux cellules filles identiques d'un point de vue morphologique. Il n'était donc pas possible de distinguer la cellule mère de la cellule fille d'un simple coup d'oeil contrairement à la levure *S. cerevisiae* ou la bactérie *C. crescentus* qui se divise de manière asymétrique et présente une phase juvénile. Un organisme vieillit à partir du moment où il est possible de distinguer les parents des descendants et donc d'observer une différenciation morphologique et/ou physiologique.

En 2005, Stewart [Stewart et al., 2005] montre que malgré une division symétrique morphologiquement, la bactérie *E. coli* présente des signes de vieillissement et une division asymétrique d'un point de vue physiologique.

Lorsqu'elle se divise, il y a formation d'un septum au centre ce qui permet la création de deux nouveaux pôles, les pôles déjà présents seront nommés les vieux pôles (Figure 1.12, page 21). Chaque cellule fille hérite ainsi d'un vieux pôle (déjà présent chez la cellule mère) et d'un nouveau pôle (créé lors de la division de la cellule mère). Il est ainsi possible de définir l'âge de chaque cellule comme étant celui du vieux pôle.

De plus, Stewart [Stewart et al., 2005] met en évidence une diminution du taux de croissance d'environ 1% par génération au cours des 7 premières générations. En étudiant de plus près les cellules filles qui héritaient consécutivement du vieux pôle et celles du nouveau pôle, il a montré que les cellules héritant du vieux pôle voient leur croissance diminuer au fil des générations par rapport à celles héritant du nouveau pôle et inversement les cellules héritant du nouveau pôle voient une légère augmentation de leur croissance avec l'âge par rapport aux cellules héritant du vieux pôle (Figure 1.13, page 21).

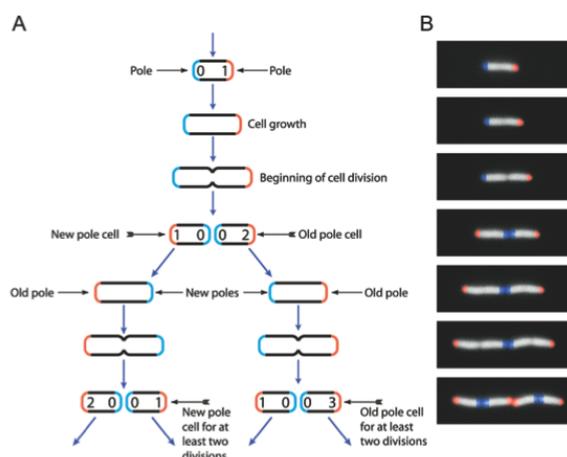


FIGURE 1.12 – **Cycle de vie de la bactérie *E. coli***. Lors de la division, deux nouveaux pôles (en bleu) sont formés, un pour chaque cellule fille et deux autres pôles sont pré-existants, les vieux pôles (en rouge). Les couleurs bleu et rouge ont été ajoutées afin de faciliter la distinction. **A)** Le nombre de divisions depuis que le pôle a été créé, est indiqué à l'intérieur de chaque pôle. Il est ainsi possible de déterminer un âge en nombre de divisions à partir du vieux pôle. **B)** Images time-lapse correspondant aux différentes étapes de la figure A. (Figure extraite de [Stewart et al., 2005])

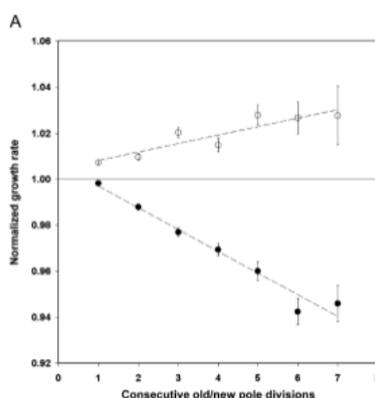


FIGURE 1.13 – **Effets de divisions consécutives à partir du vieux ou nouveau pôle sur la croissance**. Le taux de croissance représenté sur l'axe y est normalisé par le taux de croissance de toutes les cellules appartenant à une même génération dans le film. Sur l'axe des x, les divisions consécutives sont observées soit à partir du nouveau pôle (cercles vides) ou à partir du vieux pôle (cercles pleins). Pour chaque cellule héritant du vieux pôle, il a été défini un âge en nombre de divisions consécutives et de même pour les cellules héritant du nouveau pôle. En comparant ces valeurs avec les taux de croissance respectifs des cellules, ils ont montré que plus le vieux pôle est vieux et plus lent est le taux de croissance et que des cellules héritant consécutivement du nouveau pôle ont un taux de croissance plus élevé. (Figure extraite de [Stewart et al., 2005])

Le système expérimental de Stewart *et al.* ne permettait d'observer qu'un maximum de 9 générations sans superposition des cellules. Afin de regarder les effets du vieillissement à plus long terme, Wang *et al.* [Wang et al., 2010] ont développé un dispositif microfluidique (Figure 1.14, page 22) permettant de conserver et d'observer la cellule mère et donc le vieux pôle pendant environ 200 de générations.

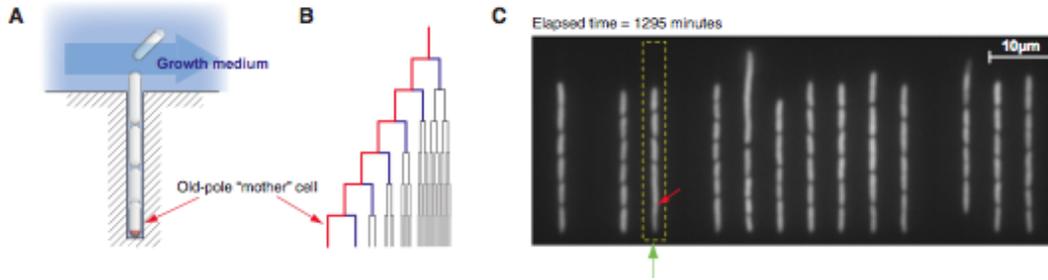


FIGURE 1.14 – **Dispositif microfluidique.** **A)** Illustration schématique du dispositif microfluidique "mother machine". La cellule mère avec le vieux pôle est piégée au fond d'un canal. Un canal principal perpendiculaire aux canaux où se trouvent les bactéries permet la circulation du milieu permettant de nourrir les bactéries et également d'évacuer les cellules sortant des canaux **B)** En rouge est représentée la branche de l'arbre correspondant à la cellule mère avec le vieux pôle d'origine. **C)** Image en sortie du microscope du dispositif microfluidique. (Figure extraite de [Wang et al., 2010])

Les résultats ne montrent aucune diminution significative du taux de croissance à long terme (Figure 1.15, page 22) contrairement à ce que l'on aurait pu attendre des résultats de Stewart. En revanche, les résultats sur les premières générations sont en accord avec ceux obtenus par Stewart et indique une baisse du taux de croissance.

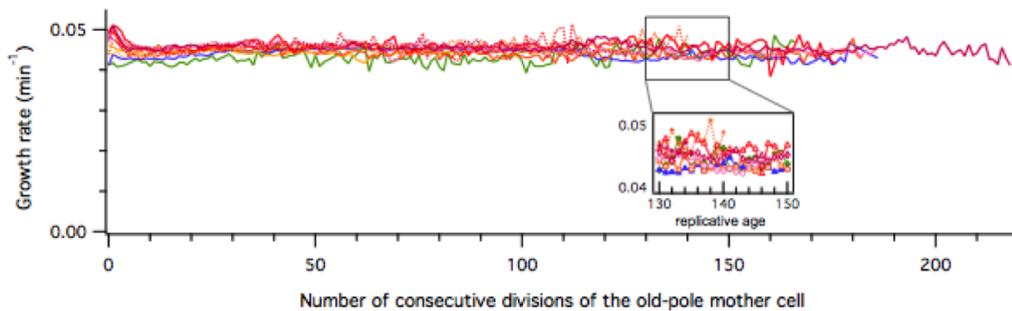


FIGURE 1.15 – **Stabilité du taux de croissance à long terme chez *E. coli*.** Le taux d'élongation montre une baisse comme dans les résultats de Stewart sur les premières générations et se stabilise ensuite sur plus de 200 générations pour les trois souches utilisées. Le taux de croissance montre des fluctuations rapides à court terme et suit une distribution gaussienne. Ces onze courbes représentent plus de 10^5 cellules mères individuelles. (Figure extraite et modifiée à partir de [Wang et al., 2010])

Cette stabilité du taux de croissance semble incompatible avec un phénomène de vieillissement car ce dernier implique une détérioration progressive de l'état physiologique de la cellule. Chez *E. coli*, la croissance cellulaire serait ainsi découplée de la mortalité.

Une autre définition du vieillissement est l'augmentation du taux de mortalité avec l'âge. Wang *et al.* ont ainsi montré qu'*E. coli* était bien sujette au vieillissement en démontrant une augmentation importante du taux de mortalité avec l'âge (Figure 1.16, page 23) .

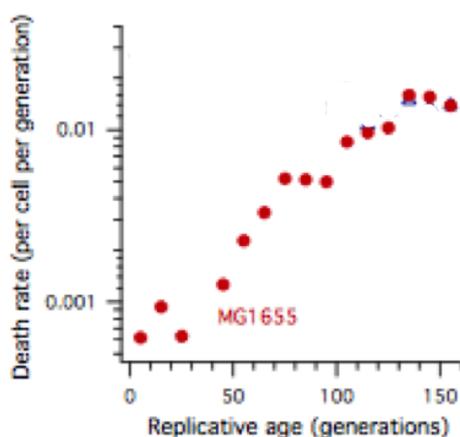


FIGURE 1.16 – Taux de mortalité en fonction de l'âge chez *E. coli*. La souche MG1655 montre une augmentation du taux de mortalité avec le temps (Figure extraite et modifiée à partir de [Wang et al., 2010])

I.3.2 Ségrégation asymétrique des agrégats protéiques chez les organismes unicellulaires et vieillissement

Le vieillissement se caractérise par une différenciation entre parents et descendants d'un point de vue morphologique ou physiologique. Chez de nombreux organismes unicellulaires, on observe une distribution asymétrique des dommages protéiques (protéines carbonylées, oxydées, mal repliées, etc.) lors de la division. Cette différenciation physiologique a ainsi permis de montrer que des organismes à division symétrique vieillissaient également. La rétention des dommages dans la cellule mère permet de maintenir des niveaux faibles de dommages au sein de la population du fait de la génération de nouvelles cellules indemnes de dommages. La ségrégation des agrégats entre la cellule mère et la cellule fille fait l'objet de nombreuses études chez la levure ou encore chez la bactérie et pourrait être impliquée dans le vieillissement, notamment chez les organismes à division symétrique qui ne permettent pas de différenciation morphologique entre les parents et les descendants.

I.3.2.1 Chez la levure, *S. cerevisiae*

La levure *S. cerevisiae* se divise de manière asymétrique afin de générer une nouvelle cellule avec un pouvoir prolifératif non entamé. Il est à noter que la durée de vie répliquative de la cellule mère est d'environ 20 à 30 divisions en conditions de laboratoire [Mortimer and Johnston, 1959]. De plus, une asymétrie physiologique est observée par une retenue préférentielle des protéines carbonylées, des ADN_r circulaires extrachromosomiques ou ERC (fragments d'ADN ribosomal circulaires. L'ADN ribosomal est constitué par environ 200 répétitions en tandem des séquences codant pour les ARN ribosomaux) ou encore des agrégats protéiques dans la cellule mère. Le gène Sir2 semble impliqué dans cette ségrégation asymétrique des agrégats protéiques car sa mutation suffit à renverser le mécanisme et entraîner une distribution symétrique des dommages entre la cellule fille et la cellule mère [Aguilaniu et al., 2003]. D'un point de vue du vieillissement répliquatif (voire la définition page 17), les dommages sont ainsi retenus de façon asymétrique dans la cellule mère afin que la cellule fille dispose d'un pouvoir prolifératif total, alors que d'un point de vue chronologique, les dommages s'accumuleraient dans les cellules n'étant plus en capacité de se diviser, jusqu'à un certain seuil qui entraînerait la mort de la cellule (Figure 1.17, page 24).

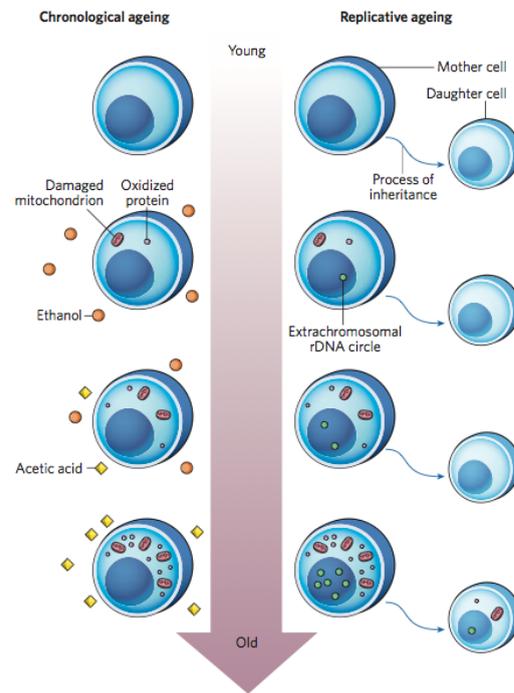


FIGURE 1.17 – Accumulation de dommages chez la levure chez *S. cerevisiae*. d'un point de vue du vieillissement répliquatif et chronologique (Figure extraite de [Kaeberlein, 2010]).

I.3.2.2 Chez la bactérie *E. coli*

La bactérie *E. coli* est une bactérie qui se divise symétriquement en deux cellules filles. Les deux cellules filles ne montrent aucune différence morphologique. Cependant, il existe une asymétrie physiologique et il est possible d'attribuer un âge à la cellule en fonction du vieux pôle (Figure 1.12, page 21). La première étude de suivi d'agrégats protéiques fut réalisée en 2008 par l'équipe de Lindner *et al.* [Lindner et al., 2008] qui a utilisé la protéine chaperonne IbpA ("holding chaperonne" se fixant aux agrégats) fusionnée à la protéine fluorescente YFP (Yellow Fluorescent Protein) afin de suivre les agrégats protéiques en conditions naturelles dans la bactérie *E. coli*. Ils ont démontré que cette fusion colocalisait avec le corps d'inclusion, les corps d'inclusion étaient visibles en contraste de phase et se superposaient aux foci observés en fluorescence (Figure 1.18, page 25).

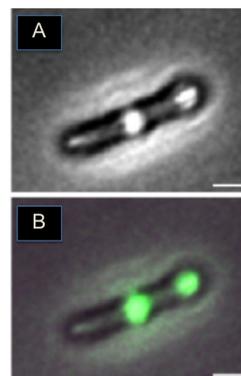


FIGURE 1.18 – Localisation *in vivo* de la protéine chaperonne *ibpA* fusionnée à la protéine fluorescente YFP par surexpression d'une protéine insoluble. **A)** Image de contraste de phase d'une cellule d'*E. coli* présentant des foci. **B)** Superposition de l'image de contraste de phase (C) avec l'image de fluorescence (*ibpA*-YFP est en vert). (Figure extraite de [Lindner et al., 2008]).

Ces auteurs ont montré que les agrégats protéiques avaient une probabilité égale d'apparaître à l'un des pôle (31% pour le vieux pôle et 30% pour le nouveau pôle) ou au centre de la cellule (28%). (Figure 1.19, page 25).

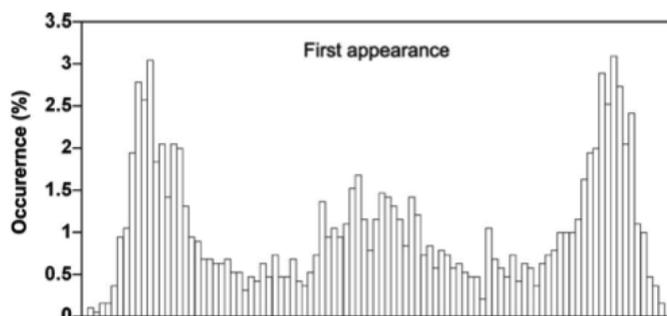


FIGURE 1.19 – Distribution des agrégats protéiques à leur première apparition. A leur première apparition, les agrégats sont au pôle ou au centre de la cellule. (Figure extraite de [Lindner et al., 2008]).

Avec le temps, la population d'agrégats localisée au milieu de la cellule décroît pour se localiser préférentiellement au vieux pôle. En une génération un nouveau pôle devient un vieux pôle et ainsi en deux générations, un agrégat apparaissant au centre de la cellule, est localisé au vieux pôle (Figure 1.20, page 26). Les foci observés bougeant très peu de leur localisation d'origine, leur localisation dans la cellule mère détermine laquelle des deux cellules filles héritera de l'agrégat protéique.

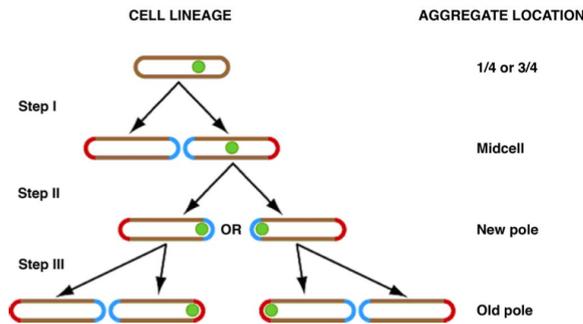


FIGURE 1.20 – **Localisation au pôle de l'agrégat protéique.** Les pôles bleus représentent le nouveau pôle formé lors de la division et les pôles rouges, les vieux pôles. Les agrégats protéiques sont représentés par les cercles verts. Les agrégats se situant en position centrale avant la division sont retrouvés au nouveau pôle à la division suivante et enfin au vieux pôle car le nouveau pôle devient un vieux pôle. (Figure extraite de [Lindner et al., 2008]).

D'autre part, les cellules contenant un vieux pôle montre une fluorescence élevée due à la présence d'un agrégat protéique mais aussi un taux de croissance diminué. Il existe ainsi une corrélation négative entre le taux de croissance et l'intensité de fluorescence (Figure 1.21, page 26).

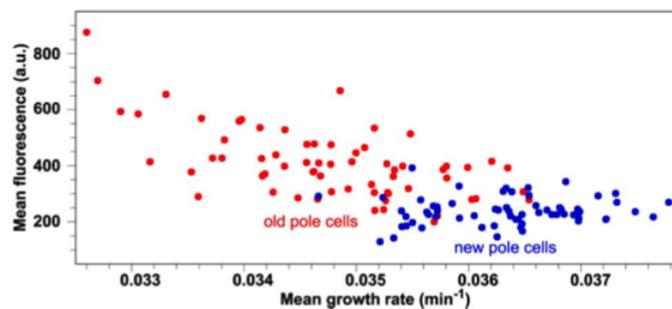


FIGURE 1.21 – **Corrélation négative entre le taux de croissance cellulaire et la fluorescence des agrégats protéiques.** Le vieux pôle présente des agrégats avec une intensité plus élevée et un taux de croissance plus faible alors que le nouveau pôle ne présente pas d'agrégats protéiques ou de très faible intensité et un taux de croissance plus rapide. (Figure extraite de [Lindner et al., 2008]).

Afin de déterminer si la présence d'un corps d'inclusion était associée à la perte de fitness, Lindner

et ses collaborateurs ont suivi dès la première division, des cellules naissantes héritant du nouveau pôle. Les cellules produites ont été divisées en deux sous-groupes : l'agrégat présent dans la cellule mère est localisé au vieux pôle (population 1) ou au nouveau pôle (population 2) (Figure 1.22, page 27). Il en résulte que la présence des corps d'inclusion est corrélée négativement avec le taux de croissance. Dans la population 1 où l'agrégat est hérité par le vieux pôle, les descendants héritant du nouveau pôle croissent plus rapidement et ceux héritant du vieux pôle croissent plus lentement qu'avant la division. En revanche, pour la population 2 (agrégat hérité par le nouveau pôle), les descendants héritant du vieux pôle croissent aussi plus lentement qu'avant la division même si la différence est moins importante que pour la population 1 mais ceux héritant du nouveau pôle ne voient pas d'augmentation de leur taux de croissance. La rétention de l'agrégat protéique dans la cellule héritant du nouveau pôle (population 2) ne permet pas de renverser totalement la différence de taux de croissance observée entre les cellules filles. Si l'agrégation avait été le seul facteur responsable du vieillissement, on aurait observé un effet du vieillissement de +4% pour la population 2. Ceci indique que l'agrégation protéique n'est pas le seul facteur en cause dans le vieillissement chez la bactérie *E. coli*.

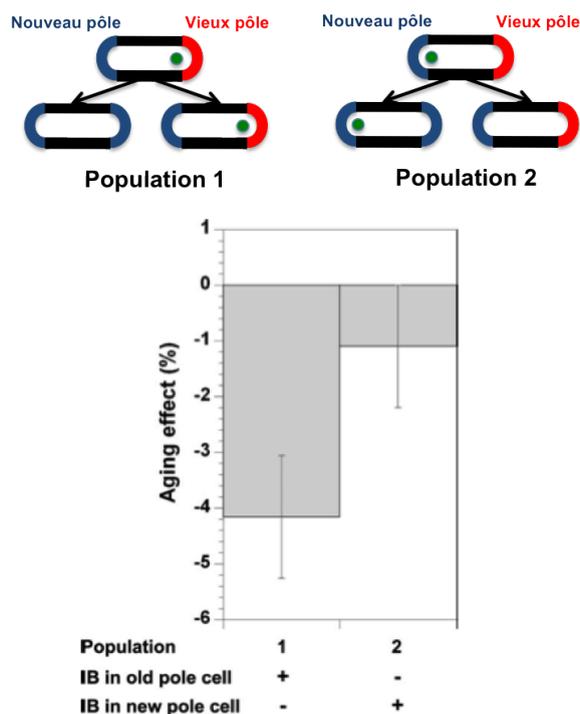


FIGURE 1.22 – **Corrélation entre le vieillissement et la présence d'agrégats protéiques.** L'effet du vieillissement a été calculé à partir de la différence du taux de croissance relatif (GR) entre les cellule fille héritant de l'agrégat protéique situé au nouveau pôle et de celle héritant de l'agrégat protéique situé au vieux pôle, de cellules mères naissantes qui héritent de l'agrégat protéique au vieux pôle (population 1) ou au nouveau pôle (population 2). (Figure extraite de [Lindner et al., 2008]).

I.3.3 Asymétrie versus symétrie

L'évolution aurait sélectionné une ségrégation asymétrique des dommages dans le but de perpétuer la population [Lynch, 2005] [Lindner et al., 2008]. Ségréger les dommages dans la cellule fille la plus âgée (cellule mère) et laisser l'autre (cellule fille) indemne de dommages permettrait la réjuvenation.

Une distribution asymétrique des dommages permettrait à la cellule indemne de ces dommages d'avoir une capacité supérieure pour faire face rapidement à l'accumulation de dommages qui mettent en péril les fonctions vitales et ainsi d'avoir une croissance plus rapide. L'allocation de la majeure partie des dommages à une cellule, la cellule mère en laissant indemne de dommages l'autre cellule serait le prix à payer pour survivre à des taux de dommages importants rencontrés dans la nature lors de stress environnementaux.

Afin d'élucider les avantages et désavantages de la division symétrique ou asymétrique, plusieurs modèles [Watve et al., 2006] [Ackermann et al., 2007] [Erjavec et al., 2008] ont été réalisés et suggèrent qu'une répartition symétrique aurait pour conséquence une croissance plus lente à cause de l'accumulation de vieux composants cellulaires.

La ségrégation des dommages dépend des conditions environnementales qui vont influencer sur le fonctionnement de la cellule. Si la production des dommages est plus rapide que la croissance de la cellule, on observe une ségrégation des dommages car le système contrôle qualité des protéines est saturé et ne peut réparer toutes les protéines endommagées [Paul, 2007]. A l'inverse, lorsque le taux de dommages n'excède pas un certain seuil, les dommages sont dilués entre les deux cellules filles, et ne nécessite pas de distribution asymétrique [Chao, 2010].

I.3.4 Mobilité des agrégats protéiques

Chez la bactérie *E. coli*, nous venons de voir que les agrégats protéiques apparaissent à l'un des pôles ou au centre de la cellule pour ensuite être localisés préférentiellement au vieux pôle en 2 générations. Ce mécanisme permet ainsi à la cellule mère de donner naissance à une cellule fille indemne de dommages, phénomène de réjuvenation, et à une seconde cellule fille qui héritera de l'agrégat protéique. Quels sont les mécanismes qui gouvernent cette localisation et la rétention des agrégats protéiques dans la cellule mère ? Est ce que les agrégats sont simplement retenus dans la cellule mère par un mécanisme passif de diffusion et la création de volumes confinés ou est ce qu'un mécanisme actif est nécessaire ? Chez la bactérie, la diffusion est le mécanisme le plus utilisé afin d'adresser les molécules à leur destination finale.

II La diffusion

Dans la cellule, la diffusion est un mécanisme important car elle concerne non seulement les mécanismes de transport mais aussi la cinétique de nombreuses réactions biochimiques. La diffusion résulte du mouvement aléatoire des molécules provoqué par le transfert d'énergie associé aux collisions. Les collisions entre molécules ont lieu à très grande fréquence : des milliards de fois par seconde. Chaque collision et le transfert d'énergie qui en résulte modifient de manière aléatoire les mouvements relatifs des molécules les unes par rapport aux autres. L'addition de tous ces mouvements relatifs se traduit macroscopiquement par le phénomène de diffusion. En dépit de ce caractère aléatoire, il en résulte un mouvement d'ensemble des molécules qui se traduit par la dispersion ou la migration de ces molécules sans aucune autre action que l'énergie thermique. Cette diffusion est le résultat de fluctuations thermiques dans la suspension, souvent désignées sous le nom de mouvement Brownien. Une conséquence macroscopique de ces mouvements aléatoires est le mouvement d'ensemble des molécules des zones les plus concentrées vers les moins concentrées.

II.1 Histoire de la diffusion

II.1.1 Le mouvement Brownien

Dans un article publié en 1828 [Brown, 1828], le botaniste écossais Robert Brown décrit le mouvement aléatoire de différentes particules suffisamment fines pour être en suspension dans l'eau. Il constate une constante et irrégulière agitation. Après avoir réitéré ses expériences avec des grains de pollen provenant de différentes plantes vivantes et mortes ainsi que sur des particules inorganiques comme du verre broyé très finement, Brown en est arrivé à la conclusion que les mouvements aléatoires n'étaient pas une propriété biologique des particules mais relevaient de la physique.

Le physicien Joseph Delsaux donne une première explication hypothétique du mouvement Brownien dans des notes qu'il publie en 1877. Selon lui, le mouvement Brownien serait le résultat des mouvements moléculaires calorifiques du liquide ambiant. Ainsi un grain de pollen suspendu dans un fluide est soumis à de multiples collisions par les molécules qui constituent le liquide. Si le grain est suffisamment petit, ces chocs aléatoires finissent par faire bouger le grain bien que leur effet moyen soit nul. Ce fut le physicien Louis Gouy [Gouy, 1888] qui fut le premier en 1888 à montrer que le mouvement Brownien est inversement proportionnel à la viscosité du milieu dans lequel évoluent les particules. Cependant, Albert Einstein sera celui qui formulera une théorie complète en 1905.

Au cours sa thèse [Bachelier, 1900], Louis Bachelier met en évidence le caractère Markovien du mouvement Brownien à savoir que le déplacement d'une particule Brownienne après un instant t dépend uniquement de l'endroit où elle était à l'instant t et non de comment elle est arrivée à cet endroit.

En 1905, Henri Poincaré écrit dans son ouvrage "Les valeurs de la science" (<http://www.ac-nancy-metz.fr/enseign/phil/textesph/Valeurdelascience.pdf>) sans avoir connaissance des travaux d'Albert Einstein :

Le biologiste, armé de son microscope, a remarqué il y a longtemps dans ses préparations des mouvements désordonnés des petites particules en suspension ; c'est le mouvement Brownien. Le biologiste a cru d'abord que c'était un phénomène vital,

mais il a vu bientôt que les corps inanimés ne dansaient pas avec moins d'ardeur que les autres ; il a alors passé la main aux physiciens ...

Cette même année 1905, Albert Einstein [Einstein, 1905] fut le premier à proposer une théorie quantitative du mouvement Brownien en démontrant que les incréments $X(t + \tau) - X(t)$ du mouvement Brownien sont des variables aléatoires qui suivent une loi de Gauss. Il montre que les mouvements des grains de pollen de Brown résultent des impulsions transmises à la particule par les chocs avec les molécules du milieu soumise à l'agitation thermique. Il montre ainsi que chaque grain de pollen se déplace autour d'une valeur constante et que son déplacement moyen est nul : $\langle x \rangle = 0$. Il démontre également que le déplacement quadratique moyen $\langle x^2 \rangle$ d'une particule sphérique de rayon r est donné par l'équation :

$$\langle x^2 \rangle = \frac{RT\tau}{3N\pi\eta r}d \quad (1.1)$$

avec R la constante de gaz parfait, T la température absolue, N le nombre d'Avogadro, η la viscosité du milieu, τ le temps nécessaire à la particule se déplacer et d le nombre de dimensions.

A partir de cette relation, Jean-Baptiste Perrin [Perrin, 1913] a apporté la confirmation expérimentale de la théorie du mouvement Brownien d'Einstein en validant son équation quelles que soient les conditions. Il a ainsi pu déterminer la valeur de la constante d'Avogadro N à partir du diamètre des particules et de leur déplacement.

En 1906, en parallèle des expériences d'Einstein, Smoluchowski menait des études très similaires sur le mouvement Brownien et interprète le mouvement Brownien comme une marche aléatoire sur un réseau. Perrin émettra également la même hypothèse que Smoluchowski à savoir que les grains de pollen de Brown suivent une marche aléatoire à cause de leur mouvement erratique [Perrin, 1913].

Une marche aléatoire est un modèle mathématique permettant de rendre compte de certains phénomènes naturels tels que le mouvement Brownien. L'idée a été évoquée pour la première fois en 1905 par le biostatisticien Karl Pearson [Pearson, 1905] qui étudiait la migration d'une population de moustiques dans une forêt.

Le modèle le plus simple de marche aléatoire et celui d'une marche aléatoire à une dimension (Figure 1.21, page 26). Une molécule n'a que deux choix possibles : aller à gauche ou à droite si l'on se place sur un réseau par exemple, avec p la probabilité d'aller à gauche à chaque instant et $q = p - 1$ la probabilité que la molécule se déplace à droite à chaque instant. Dans le cas d'un mouvement Brownien, la marche aléatoire est isotrope, et la probabilité de se déplacer à gauche ou à droite est la même. On a ainsi une équiprobabilité :

$$p = q = \frac{1}{2} \quad (1.2)$$

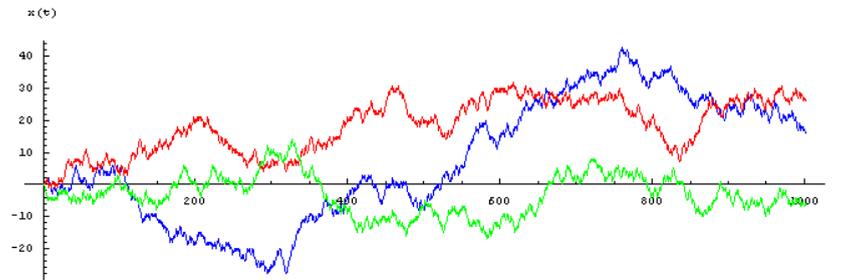


FIGURE 1.23 – **Marches aléatoires d’une particule dans un espace à une dimension.** La figure représente 3 marches aléatoires isotropes indépendantes pour une particule avec pour condition initiale $x(0)=0$. (Figure extraite de http://fr.wikipedia.org/wiki/Fichier:Walk1d_0.png). Chaque particule se déplace autour d’une valeur constante et leur déplacement moyen est nul.

Dans le cas d’une marche aléatoire à deux dimensions sur lattice carré, une particule dispose de 4 mouvements possibles : gauche, droite, en haut ou en bas (Figure 1.24, page 31 ; Figure 1.25, page 32).

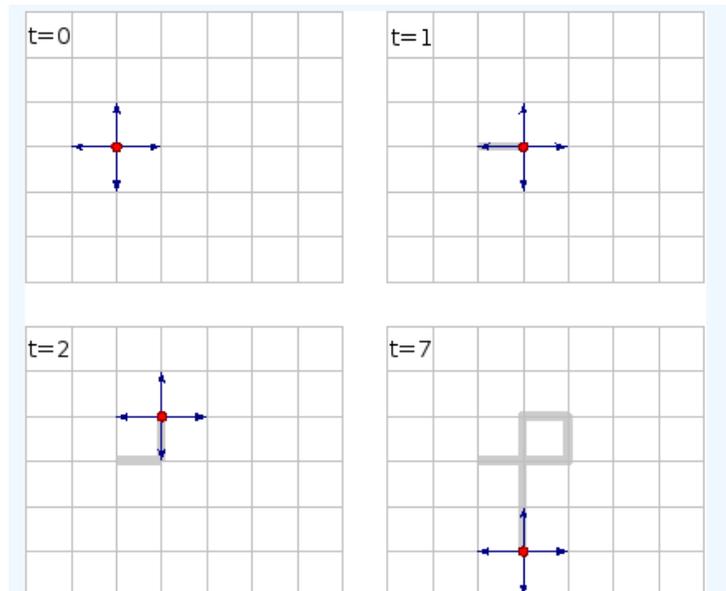


FIGURE 1.24 – **Marche aléatoire d’une particule sur un réseau en deux dimensions.** (Figure extraite de <http://planetesciences.fr/gd/Wendelin-Werner,-le-seigneur-des-zigzags.htm>)

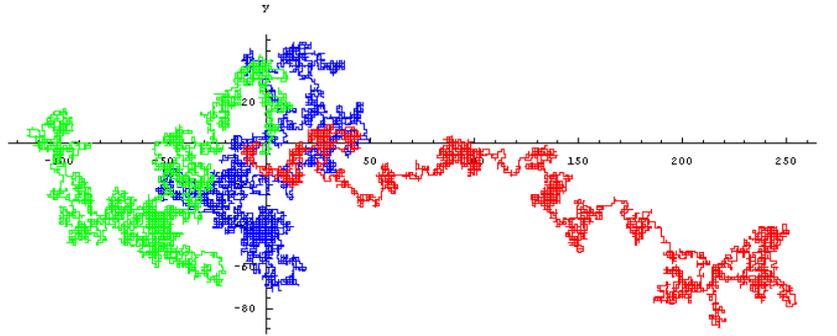


FIGURE 1.25 – **Trajectoire d’une particule suivant une marche aléatoire dans un espace à deux dimensions.** La figure représente 3 marches aléatoires isotropes indépendantes pour une particule démarrant de l’origine. (Figure extraite de http://upload.wikimedia.org/wikipedia/commons/5/53/Walk2d_0.png)

II.1.2 Equation de la diffusion : Stokes-Einstein

L’équation de Stokes-Einstein est la première équation dérivée par Einstein lors de sa thèse de doctorat [Einstein, 1905] afin de calculer le coefficient de diffusion d’une particule se déplaçant selon un mouvement Brownien dans un fluide à température constante.

$$D = \frac{RT}{N6\pi\eta r} \quad (1.3)$$

En 1908, Langevin a établi un modèle théorique qui est une extension de la marche aléatoire dans lequel on analyse l’effet du fluide sur la particule Brownienne. L’équation de Langevin prend en compte deux forces : une force de frottement proportionnelle à la vitesse et une force aléatoire qui regroupe les impacts du fluide sur la particule Brownienne. Marche aléatoire, équation de Langevin et loi de Fick sont trois façons de décrire le mouvement d’une particule à trois échelles de temps et d’espace différentes.

Einstein montre que le mouvement agité des grains dispersés dans le fluide est le même que celui qu’ils auraient s’ils diffusaient dans le fluide en partant tous d’un même endroit. Or la concentration autour de cet endroit suit une loi gaussienne qui dépend du temps et du coefficient de diffusion. Il en déduit ainsi l’expression du déplacement d’un grain ou plus exactement de la racine carré de la moyenne du carré des déplacements d’un grain dans une direction particulière x . Ainsi c’est le carré du déplacement qui est proportionnel au temps et permet de mesurer le mouvement Brownien.

$$\langle x^2(t) \rangle = 2dDt \quad (1.4)$$

avec d le nombre de dimensions.

Einstein [Einstein, 1956] a montré que le coefficient de diffusion D dans une solution diluée infiniment est donné par l’équation :

$$D = \frac{kT}{f} \quad (1.5)$$

avec f le coefficient de friction de la particule, k la constante de Boltzmann et T la température absolue.

Dans la plupart des cas, la valeur de f est inconnue, c'est pourquoi G. Stokes [Stokes, 1851] a montré que dans le cas spécifique d'une particule sphérique de rayon r avec vitesse uniforme dans un fluide continu de viscosité η , le coefficient de friction f_0 est donné par l'équation :

$$f_0 = 6\pi\eta r \quad (1.6)$$

Einstein démontra que le coefficient de diffusion de particules sphériques pouvait être calculé par l'équation 1.3 ci-dessus.

Cette relation est aussi connue sous le nom de relation de Stokes Einstein [Einstein, 1956]. La vitesse de mouvement dépend directement de la taille des particules. Si le coefficient de diffusion est mesuré et que la viscosité du solvant est connue, alors le rayon hydrodynamique peut être calculé à partir de l'équation de Stokes Einstein. Le rayon hydrodynamique est la taille effective de la molécule détectée à partir de son mouvement brownien.

II.2 Caractérisation de la diffusion

Dans la cellule, la diffusion est un processus important car elle concerne aussi bien la localisation de molécules que la cinétique de nombreuses réactions biochimiques. La diffusion des protéines va dépendre de leurs tailles et de la viscosité du milieu mais également des interactions protéine-protéine ou la formation de complexes intermédiaires. Les liaisons ou les collisions avec d'autres molécules vont également ralentir la diffusion.

II.2.1 Marche aléatoire

Pour une solution infiniment diluée avec absence de toute interaction modifiant le mouvement de la molécule, le mouvement de diffusion d'une molécule peut-être décrit par une marche aléatoire : $\langle x^2 \rangle = 2dDt$, $\langle x^2 \rangle$ est le carré moyen de la distance parcourue par la molécule en un temps t , d le nombre de dimensions et D le coefficient de diffusion de la molécule suivant la relation d'Einstein [Einstein, 1905], équation 1.3 page 32.

La présence de nombreuses autres molécules dans l'environnement cytoplasmique diminue le volume libre accessible pour la molécule et crée des barrières. Cela se répercute sur le coefficient de diffusion qui est alors fortement diminué et dépendra alors de la fraction volumique occupée. Ainsi l'observation de la diffusion *in vivo* sur des temps larges indique alors un régime sous-diffusif avec une zone de confinement.

II.2.2 Encombrement moléculaire et diffusion

L'environnement cellulaire est un milieu complexe dont l'une des principales caractéristiques est une concentration très élevée en macromolécules. Cet encombrement moléculaire a un effet non négligeable sur la diffusion des molécules [Luby-Phelps et al., 1987] qui est un mécanisme central pour le fonctionnement cellulaire. Le volume exclu ralentit ainsi la diffusion

en diminuant le volume dans lequel la molécule peut se déplacer et trouver un espace libre [Muramatsu and Minton, 1988] (Cf. section I.1.3 pour la notion de volume exclu).

L'encombrement moléculaire crée des barrières au mouvement des particules. Dans de nombreux cas, il ralentit la diffusion mais l'issue d'une diffusion anormale n'est pas systématique. La réduction de la diffusion est due à une interaction entre les particules et les autres composants cellulaires ou encore la membrane.

Plusieurs modèles récents [McGuffee and Elcock, 2010], [Ando and Skolnick, 2010] ont tenté de modéliser le cytoplasme de la bactérie *E. coli* afin d'expliquer la réduction d'un facteur 10 de la diffusion *in vivo*. Une approche réalisée au niveau atomique ([McGuffee and Elcock, 2010]) incluant la modélisation de 50 des plus abondantes macromolécules présentes dans le cytoplasme de la bactérie *E. coli* a permis d'obtenir des valeurs de diffusion proches de celles obtenues *in vivo* et démontrent que les effets stériques dus à l'encombrement peuvent expliquer à eux seul le facteur 10 observé.

Un autre modèle [Ando and Skolnick, 2010] démontre l'implication de deux facteurs principaux : l'effet de volume exclu et les interactions hydrodynamiques.

II.2.3 Différents types de diffusion

On distingue deux types de transports passifs dans la cellule :

- La diffusion normale dite Brownienne est caractérisée par un déplacement carré moyen (MSD ou Mean Square Displacement) des molécules linéaire avec le temps : $\langle x^2 \rangle = A*t$. Elle peut-être contrainte par un confinement géométrique c'est-à-dire que le volume dans lequel la molécule peut diffuser est limité et l'observation de la diffusion à des temps longs indique une saturation du MSD. Cette saturation permet de calculer la taille du confinement. Nous noterons qu'en milieu confiné, il est possible de prédire le MSD à temps très faible car il est alors non sensible au confinement ($MSD = 2dDt$) et à temps très grands, le MSD étant constant. En revanche, il n'existe pas d'expression mathématique valable à tous les temps (faibles et grands) permettant de prédire le MSD.
- La diffusion anormale qu'il ne faut pas confondre avec le confinement (vu ci-dessus). La diffusion anormale est caractérisée par un déplacement carré moyen proportionnel au temps en suivant une loi de puissance : $\langle x^2 \rangle \sim t^\alpha$ avec α . Si $\alpha > 1$, le régime est sur-diffusif et si $\alpha < 1$, le régime est sous-diffusif. Les expériences *in vivo* montrent que la diffusion dans la plupart des compartiments intracellulaires (membrane, cytoplasmes, noyaux) est anormale et sous-diffusive [Fedotov et al., 2010] [Schmidt and Weiss, 2011]. Bien que les causes de ces comportements ne soient pas encore élucidées à ce jour, la présence d'obstacles [Saxton, 1994] est une hypothèse souvent émise. Pour une concentration modérée d'obstacles, la diffusion est anormale sur des temps courts et normale sur des temps longs [Saxton, 1994]. Quand la concentration des obstacles tend vers la percolation formation d'un amas de taille infinie, la diffusion devient anormale également sur des temps longs. Dans les membranes, on parle souvent de diffusion anormale mais ce n'est pas un phénomène universel et lorsque l'on est dans un cas de diffusion anormale des facteurs autres que l'encombrement moléculaire créé par la présence d'obstacles sont en cause comme des transports actifs ou des interactions protéines-protéines [Dix and Verkman, 2008].

II.3 Mesures de la diffusion

II.3.1 Ordres de grandeur

La diffusion varie en fonction du milieu et de sa viscosité et est plus élevée dans un gaz que dans un liquide en raison des plus faibles interactions à l'échelle moléculaire. Dans un liquide elle est d'autant plus élevée que la masse moléculaire de la molécule est faible. Les coefficients de diffusion ont ainsi des ordres de grandeur qui peuvent être très différents. Le tableau ci-dessous indique les ordres de grandeur de différents coefficients de diffusion.

Molécules	Coefficient de diffusion en $\mu\text{m}^2/\text{s}$
Proton	7000 [Agmon, 1995]
Molécule d'eau dans l'eau	2300 [Agmon, 1995]
GFP dans l'eau	87 [Swaminathan et al., 1997]
Kinase phosphoglycérate chez la levure <i>S. cerevisiae</i>	63.8 [Squire and Himmel, 1979]
GFP dans le cytoplasme de cellules eucaryotes	27 [Elowitz et al., 1997]
Venus YFP dans le cytoplasme de la bactérie <i>E. coli</i>	8 [Elf et al., 2007]
GFP dans le cytoplasme de la bactérie <i>E. coli</i> , DH5 α	7.7 [Elowitz et al., 1999]
ADN dans l'eau	0.81 à 53 [Lukacs et al., 2000]
Facteurs de transcription dans le noyau des eucaryotes	0.5 à 5 [Hager et al., 2009]
Kinase histidine chez la bactérie <i>C. crescentus</i>	0.010 à 0.014 [Deich et al., 2004]

GFP : green fluorescent protein (protéine fluorescente verte)

YFP : yellow fluorescent protein (protéine fluorescente jaune)

II.3.2 Chez la bactérie *E. coli*

Elowitz [Elowitz et al., 1999] fut le premier à quantifier *in vivo* la mobilité de la protéine fluorescente GFP (Green Fluorescent Protein) dans le cytoplasme de la bactérie *E. coli* et il trouva que son mouvement était cohérent avec un mécanisme passif Brownien de diffusion. Cependant, il observe que le coefficient de diffusion apparent de la GFP dans le cytoplasme est inférieur environ de 11 fois ($7.7 \mu\text{m}^2/\text{s}$) à celui mesuré dans l'eau : $87 \mu\text{m}^2/\text{s}$ [Terry et al., 1995], [Swaminathan et al., 1997]. De plus, ce coefficient de diffusion est bien inférieur à celui mesuré dans une cellule eucaryote [Swaminathan et al., 1997] ou encore dans la mitochondrie : $20\text{-}30 \mu\text{m}^2/\text{s}$ ([Partikian et al., 1998]). Cette réduction non négligeable du coefficient de diffusion *in vivo* par rapport à la diffusion de la même protéine dans l'eau est observé pour toutes les protéines.

D'autres études chez la bactérie *C. crescentus* [Deich et al., 2004], ont permis de suivre une protéine de la membrane : l'histidine kinase PleC fusionnée à la protéine fluorescente YFP, qui régule en partie l'asymétrie lors de la division et se localise à un des pôles à des moments spécifiques du cycle cellulaire (Figure 1.26, page 36).

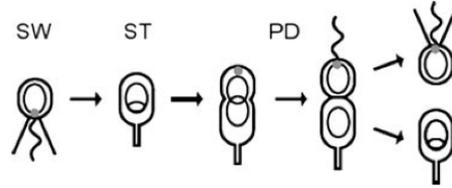


FIGURE 1.26 – Cycle cellulaire de la bactérie *Caulobacter crescentus* et localisation de la protéine PleC. Lorsque la cellule est mobile (SW : swarmer cell), PleC (points gris) se situe au pôle ayant le flagelle. Durant la différenciation en cellule fixée (ST : stalked cell), la localisation de PleC n'est plus possible. Lors de la division (PD : predivisional cells) PleC se situe au pôle où se situera le nouveau flagelle. (Figure extraite de [Deich et al., 2004]).

Le mouvement de la protéine PleC est gouverné par un mouvement brownien dans un espace confiné (saturation du MSD au temps longs) avec une estimation du coefficient de diffusion en 2d de $12 \cdot 10^{-3} \mu\text{m}^2/\text{s}$. Il n'existe pas d'expression mathématique permettant de déterminer $MSD = f(t)$ quand la diffusion est confinée dans un sous-espace. Des simulations basées sur l'individu peuvent alors être utilisées pour estimer les facteurs de correction relatif à la géométrie des cellules et calculer ainsi le coefficient de diffusion (Figure 1.27, page 36).

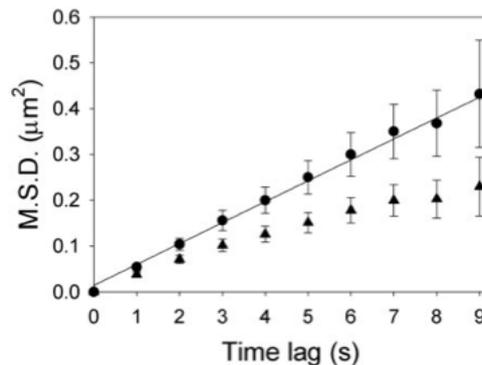


FIGURE 1.27 – Courbe du MSD en fonction du temps pour les données brutes (triangles) et les données corrigées (cercles) par rapport à la projection 2d. Les marqueurs triangulaires représentent le MSD en fonction du temps à partir des trajectoires. Les marqueurs circulaires réfèrent aux données corrigées par les simulations. (Figure extraite de [Deich et al., 2004]).

Aucune variation du coefficient de diffusion (Figure 1.28, page 37) n'a été observée selon la localisation de la protéine PleC dans la cellule ou le type de cellule (Figure 1.26, page 36), à l'exception d'une réduction quand PleC se situe au pôle présentant le flagelle chez la cellule mobile ou en division. Cette diminution de la diffusion peut être expliquée par sa localisation qui impliquerait une viscosité plus importante de la membrane ou encore la topologie de celle-ci. Il est également possible que les molécules étaient mobiles au début du suivi et se sont ensuite immobilisées, ce qui a entraîné une diminution du coefficient de diffusion.

	s coordinate										
	0-0.1	0.1-0.2	0.2-0.3	0.3-0.4	0.4-0.5	0.5-0.6	0.6-0.7	0.7-0.8	0.8-0.9	0.9-1	All
All	4	12	16	19	11	17	15	16	18	14	13
SW	3		17		20		16		14		14
PD	4		16		10		16		19		12
ST	13		17		14		16		22		16

FIGURE 1.28 – **Tableau des coefficients de diffusion en 2d de l’histidine kinase PleC en fonction du type de cellule et de la position selon l’axe de la cellule.** Les coefficients de diffusion sont exprimés en $10^{-3} \mu m^2/s$ et ont été corrigés pour la projection 2d. (Tableau extrait de [Deich et al., 2004]).

Une étude sur le mouvement de fragments d’ADN proche de l’origine de réplication chez la bactérie *E. coli* [Elmore et al., 2005] révèle un mouvement diffusif libre selon l’axe le plus long de la bactérie avec un coefficient de diffusion de $4.3 \cdot 10^{-5} \mu m^2/s$. La diffusion dans la largeur de la bactérie indique un confinement dont le diamètre correspond à celui du noyau avec un coefficient de diffusion de $2.9 \cdot 10^{-5} \mu m^2/s$.

II.3.3 Différences entre *in vivo* et *in vitro*

Les mesures de diffusion effectuées *in vivo* et *in vitro* diffèrent de plusieurs ordres de grandeur et montrent une diffusion réduite *in vivo*. Dans les années 1980, une étude de Muramatsu ([Muramatsu and Minton, 1988]) indiquait que la haute concentration macromoléculaire du cytoplasme réduisait d’un facteur 3 la mobilité des protéines suivies (pour un poids moléculaire compris entre 17 et 150 kDa). D’autres études [Muramatsu and Minton, 1988] [Elowitz et al., 1999] [Konopka et al., 2009] [Verkman, 2002] révèlent un coefficient de diffusion apparent plus lent *in vivo* qu’*in vitro*, avec une différence pouvant atteindre un facteur 10.

Récemment, Nenninger *et al.* [Nenninger et al., 2010] ont suivi et quantifié la diffusion d’oligomères de GFP dont la taille variait de 1 à 6 monomères assemblés. Ils ont montré que le coefficient de diffusion dépendait de la taille des multimères de GFP et suivait ainsi la prédiction de l’équation de Stokes-Einstein en milieu visqueux pour un nombre de monomères variant de 1 à 4 (Figure 1.29, page 38). Le coefficient de diffusion diminue graduellement avec l’augmentation de la taille de la molécule. Cette conclusion peut laisser suggérer que pour des molécules de taille inférieure à une centaine de kDa (4 monomères de GFP), l’encombrement moléculaire ou encore les interactions avec les autres composants cytoplasmiques n’ont que peu d’effets sur la diffusion. Cependant à partir de 5 monomères de GFP, le coefficient de diffusion décroît en-dessous de la prédiction de l’équation de Stokes-Einstein et suggérerait des interactions avec d’autres composants cellulaires ou la membrane pour des molécules de taille supérieur à 111 kDa. Ces résultats sont accord avec ceux obtenus précédemment par Kumar [Kumar et al., 2010] qui indiquent que des molécules à faible poids moléculaire ont une mobilité principalement déterminée par leur taille tandis que des molécules à haut poids moléculaire (à partir de 70 kDa) ont un coefficient de diffusion qui décroît très rapidement avec leur masse moléculaire, démontrant ainsi l’existence de facteurs ou mécanismes ralentissant la diffusion dans le cytoplasme.

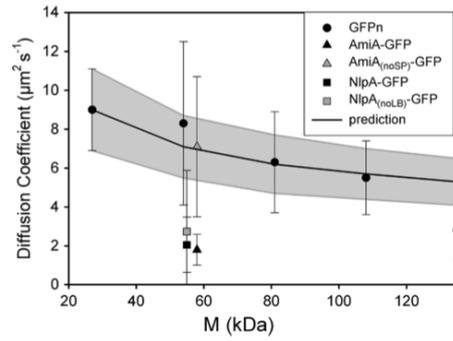


FIGURE 1.29 – Coefficient de diffusion de multimères (1 à 5) de GFP dans le cytoplasme de la bactérie *E. coli*. La ligne indique le coefficient de diffusion prédit par l'équation de Stokes-Einstein. M est la masse molaire. (Figure extraite de [Nenninger et al., 2010])

III Objectifs

Ce projet vise à comprendre la dynamique de l'agrégation protéique chez la bactérie *Escherichia coli*. L'agrégation protéique joue un rôle clé dans la dégénérescence cellulaire et est notamment reliée à de nombreuses maladies humaines en lien avec le vieillissement telles que les maladies d'Alzheimer et Parkinson ou encore la maladie du prion. Chez la bactérie *E. coli*, l'accumulation de dommages sous forme d'agrégats protéiques et leur ségrégation asymétrique aux pôles (les extrémités hémisphériques de la cellule) ont permis de démontrer des signes de vieillissement malgré une division symétrique et l'absence de distinction claire entre la cellule mère et la cellule fille. Nous nous sommes intéressés à l'étude de la dynamique spatiale des agrégats protéiques in vivo chez la bactérie *E. coli*.

Les agrégats protéiques peuvent être classés en deux grandes classes selon leur ordre de structuration : les corps d'inclusions (amorphes ou peu structurés) et les amyloïdes (structuration élevée). Nous nous sommes intéressés à ces deux classes et pour les amyloïdes, nous nous sommes concentrés plus particulièrement sur les prionoides qui sont des prions non infectieux.

Dans ce travail, nous avons mis en évidence que le mouvement des agrégats protéiques (sous forme de corps d'inclusion) localisés au pôle chez la bactérie *E. coli* est gouverné par un mouvement de diffusion Brownien et ne nécessite aucun mécanisme actif.

Au moyen d'une double approche théorique et expérimentale, nous avons suivi, grâce à la microscopie time-lapse, l'apparition d'agrégats protéiques via la protéine chaperonne *ibpA* fusionnée à la protéine fluorescente jaune (YFP). Le développement d'un logiciel d'analyse d'images automatisé en collaboration avec l'équipe de Lionel Moisan a permis d'analyser de manière quantitative et très précise (10ème de pixel) les trajectoires des agrégats protéiques dans l'espace intracellulaire et de comprendre quel mécanisme (passif ou actif) régissait leur localisation. Nos résultats indiquent clairement un mouvement Brownien de diffusion avec un coefficient de diffusion dépendant de la taille de la molécule en accord avec l'équation de Stokes-Einstein. De plus, le confinement observé selon l'axe le plus long de la bactérie semble confirmer que les agrégats protéiques sont localisés dans les régions dépourvues d'ADN. Ainsi l'encombrement moléculaire dû aux nucleoides (ADN bactérien) est un facteur important dans la distribution spatiale des agrégats. De plus, au travers d'un modèle 3d individus basés de diffusion-agrégation, nous avons démontré qu'un mouvement de diffusion-agrégation associé à la présence des nucleoides était suffisant pour expliquer la distribution spatiale des agrégats protéiques. Par ailleurs les différences observées entre conditions natives et conditions de choc thermique, en terme de nombres d'agrégats protéiques semblent uniquement reliées à l'augmentation du nombre total de molécules "agrégables" en condition de choc thermique. L'ensemble des résultats apporte la preuve quantitative que la localisation des agrégats protéiques au pôle chez la bactérie *E. coli* s'effectue par un mécanisme passif de diffusion-agrégation couplé à une distribution spatiale non-homogène de l'encombrement macromoléculaire.

Dans un second temps, en collaboration avec l'équipe de R. Giraldo en Espagne, nous avons également étudié d'autres agrégats protéiques de type amyloïde et prionide chez la bactérie *E. coli*. Leur particularité est de pouvoir être transféré verticalement entre la cellule mère et la cellule fille

lors de la division mais ils ne sont pas infectieux. Grâce à un mutant de la protéine RepA, protéine initiant la réplication, fusionnée à la protéine fluorescente rouge (RFP), nous nous sommes intéressés à l'impact de ce type d'agrégats sur le vieillissement (taux de croissance de la bactérie) au moyen de la microscopie microfluidique afin d'observer à la fois les cellules uniques et leur lignage. Ces agrégats se présentent sous trois formes : solubles, foci et fibrilles amyloïdes. Leur identification et caractérisation automatisées, ont nécessité la mise au point d'un second logiciel d'analyses d'images. Ce modèle synthétique simple permet l'étude de protéinopathies amyloïdes chez la bactérie.

Dans la suite de ce manuscrit, je débiterai par un chapitre évoquant les différents outils méthodologiques que nous avons été amenés à utiliser d'un point de vue expérimental et théorique puis je présenterai les résultats obtenus sous la forme d'articles et je terminerai par une discussion et les perspectives de ce projet.

Méthodologie

Ce projet interdisciplinaire a fait appel à différents domaines tels que la physique, la biologie ou l'informatique. Ce chapitre introduira les différents outils méthodologiques utilisés en microscopie, traitement d'images et modélisation.

Sommaire

I	Microscopie	42
I.1	Microscopie optique	42
I.2	Technologies	45
I.3	Suivi de molécules <i>in vivo</i>	47
II	Analyses d'images	49
II.1	Cellst (Cell Segmentation and Tracking)	50
II.2	Bacteria Analyzer	54
III	La modélisation	58
III.1	Le modèle multi-agents (SMA)	58
III.2	Le modèle basé individus (IBM)	59
III.3	Notre modèle	59

I Microscopie

En 1675, Antonie Philips van Leeuwenhoek, connu comme le père de la microbiologie, a été le premier à observer et décrire des cellules et organismes au moyen d'un microscope très simple composé d'une seule lentille.

I.1 Microscopie optique

I.1.1 Les bases

Un microscope est composé de deux lentilles convergentes qui sont l'objectif et l'oculaire. L'objet à étudier est placé en amont du foyer objet (f) de l'objectif et fournit une image réelle renversée agrandie (Figure 2.1, page 42). L'image formée par la lentille se forme au-delà de son foyer image (f'). Le grandissement dépend de la distance du foyer au centre de la lentille et de la distance de l'objet au foyer objet. Plus la distance objet-foyer objet sera courte, plus fort sera le grossissement.

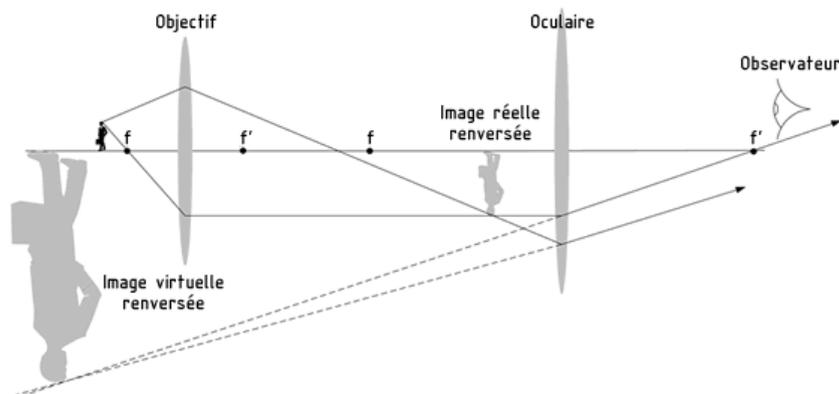


FIGURE 2.1 – La lentille convergente.

Le constituant majeur d'un microscope est l'objectif. C'est de lui que dépend le grossissement et la résolution d'un microscope. L'ouverture numérique (NA) de l'objectif est une caractéristique essentielle car elle détermine la luminosité et la résolution. Elle est définie par :

$$NA = n \sin \alpha \quad (2.1)$$

où n est l'indice de réfraction du milieu et α le demi-angle d'ouverture de l'objectif.

Un objectif de bonne qualité sera caractérisé par une grande ouverture numérique mais aura une distance de travail limitée et une profondeur de champ réduite.

Afin de limiter l'angle de réfraction à la sortie de l'échantillon, on utilise des objectifs à immersion. L'huile est le plus souvent utilisée. L'objectif peut ainsi récolter plus de rayons lumineux.

On notera que les microscopes que nous avons utilisés sont des microscopes inversés. À l'inverse de la microscopie optique classique où la lumière arrive sur l'échantillon par le bas et où l'observation

se fait par le dessus, pour le microscope inversé la source de lumière est placée au-dessus de l'échantillon et les objectifs en dessous. Les microscopes inversés permettent d'observer des boîtes de Pétri ou d'utiliser des systèmes microfluidiques.

I.1.2 Microscope à contraste de phase

Le principe repose sur la formation d'un contraste d'intensité créé par l'interférence de rayons lumineux. Des anneaux de phase sont insérés dans le trajet optique et vont créer des retards de phase différents suivant que les rayons traversent ou non l'objet. Le déphasage entre les rayons créé le contraste et permet de mettre en évidence des structures dans l'objet observé.

Plus précisément, un anneau de phase est placé dans le condensateur et une plaque de phase dans l'objectif du microscope (Figure 2.2, page 43). L'anneau de phase permet de créer l'illumination de forme particulière qui sera stoppée par l'anneau opaque de la plaque de phase. Les rayons lumineux diffractés par les structures fines traversent la plaque de phase soit dans la zone périphérique, soit dans la zone centrale. La zone centrale a pour rôle de retarder les rayons diffractés qui la traversent. Ces rayons retardés vont interférer avec ceux transmis sans retard sur la périphérie. Le déphasage différentiel entre ces rayons crée le contraste et permet de mettre en évidence les structures cellulaires.

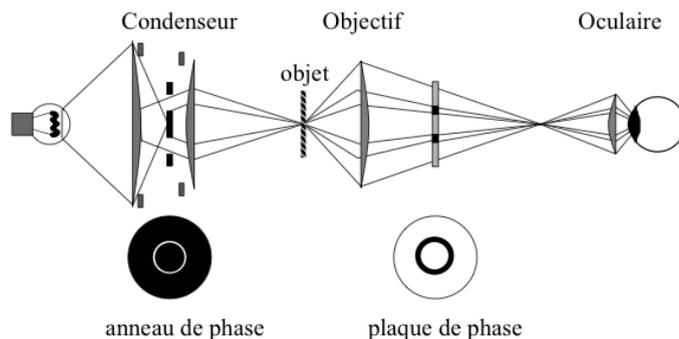


FIGURE 2.2 – Le microscope à contraste de phase.

I.1.3 Microscope à fluorescence

Un microscope à fluorescence est constitué de deux lampes, une pour l'observation par transmission et une pour la fluorescence. Les jeux de filtres sont constitués d'un filtre d'excitation, d'un miroir dichroïque et d'un filtre d'émission (Figure 2.3, page 44). Une molécule ou un atome peut absorber de l'énergie. Il est alors dans un état excité instable et son retour à l'état fondamental peut se produire par photoluminescence. Chaque molécule peut être caractérisée par des spectres d'absorption et d'émission qui lui sont propres et qui reflètent la distribution de probabilité des transitions énergétiques.

La fluorescence peut être naturelle, on parle alors d'autofluorescence (par exemple la chlorophylle fluorescent naturellement en rouge) ou résulter de l'ajout d'un fluorophore qui se fixe à des structures spécifiques comme le DAPI qui se fixe spécifiquement à l'ADN et fluoresce en bleu.

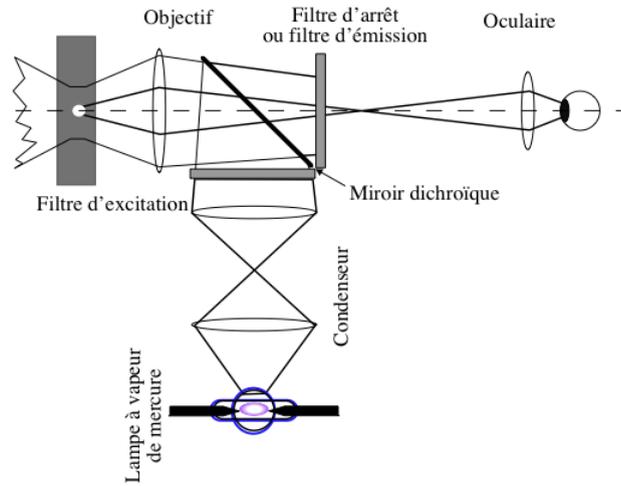


FIGURE 2.3 – Le microscope à fluorescence.

I.1.4 Résolution optique

La formation de l'image fait intervenir les propriétés ondulatoires de la lumière et en particulier la diffraction.

Elle est la résultante de deux fonctions de répartition de la lumière :

- Une fonction d'étalement du point ou Point Spread Function (PSF) dans le plan confocal conditionnant la résolution latérale (en XY). C'est la description mathématique de l'image du disque d'Airy obtenue lorsqu'on observe un point lumineux au travers d'un objectif. Ce disque d'Airy (Figure 2.4, page 44) résulte de la formation d'interférences des ondes lumineuses diffractées le long du trajet des rayons au niveau de l'image. L'image d'une source ponctuelle au travers d'un objectif est une tâche lumineuse entourée d'anneaux concentriques d'intensité décroissante.
- Une fonction de défocalisation conditionnant la résolution axiale (en Z).



FIGURE 2.4 – Disque d'Airy.

La résolution d'un microscope optique se définit comme la plus petite distance entre deux points d'un échantillon qui peuvent être distingués par l'oeil ou un capteur comme deux entités différentes. La résolution latérale (R_{xy}) définie par Rayleigh comme la distance entre deux points objets pour laquelle le maximum d'intensité du disque d'Airy du premier point correspond au premier minimum d'intensité du disque d'Airy du second point. Elle s'exprime par l'équation

suivante :

$$R_{xy} = 0.61\lambda NA \quad (2.2)$$

où λ est la longueur d'onde d'émission et NA, l'ouverture numérique de l'objectif.

Plus l'ouverture numérique de l'objectif sera grande, plus la tâche d'Airy sera étroite et meilleure sera la résolution optique (Figure 2.5, page 45).

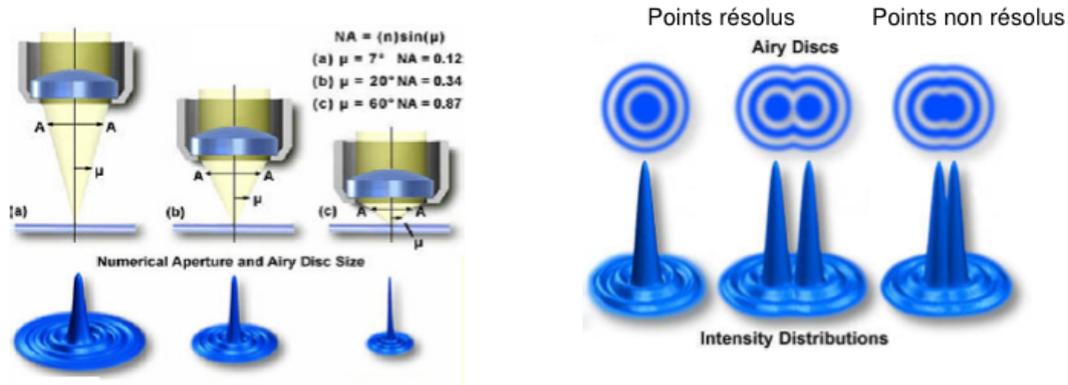


FIGURE 2.5 – Résolution latérale et ouverture numérique.

La résolution axiale (le long de l'axe Z, R_z) est définie selon le même principe. Elle est inversement proportionnelle au carré de l'ouverture numérique de l'objectif.

$$R_z = \frac{2\lambda n}{NA^2} \quad (2.3)$$

où n est l'indice de réfraction du milieu et NA l'ouverture numérique de l'objectif.

I.2 Technologies

Désormais, la microscopie permet l'acquisition automatisée d'images dans le temps (création de séquences d'images). Les films sont ensuite analysés afin d'étudier de nombreux mécanismes biologiques à différentes échelles et au niveau de la cellule unique également.

I.2.1 Microscopie time-lapse

La microscopie time-lapse permet de suivre la croissance de cellules ou la dynamique de protéines dans un environnement déterminé. Les cellules sont fixées sur une fine couche d'agar (1.5% d'agarose dans le milieu qui peut être du LB ou du LB demi-salé ou du milieu minimum selon les conditions d'étude). Une lamelle est posée par dessus afin de fixer le tout. Les cellules se situent ainsi entre l'agar et la lamelle.

I.2.2 Système microfluidique

Cette technologie permet d'étudier des cellules uniques et leur lignage dans un environnement contrôlé à long terme. De plus, cet environnement peut également être changé afin d'étudier

des variations.

Un système microfluidique est un circuit composé de canaux ou chambres. Le circuit est dessiné puis imprimé sur un bloc de PDMS grâce à un moule (Figure 2.6 A., page 46). Il est possible de créer des design différents : variation dans le nombre de canaux d'alimentation ou contenant les bactéries, plusieurs ramifications, création de chambres, etc.).

Le design de notre circuit microfluidique est composé de canaux agencés de la manière suivante (Figure 2.6 B., page 46) :

- Le canal principal permettant d'amener le milieu qui nourrira les bactéries et d'évacuer les bactéries sortant des canaux.
- Des canaux perpendiculaires au canal principal adapté à la taille des bactéries (environ $1.4 \mu\text{m}$ de section).

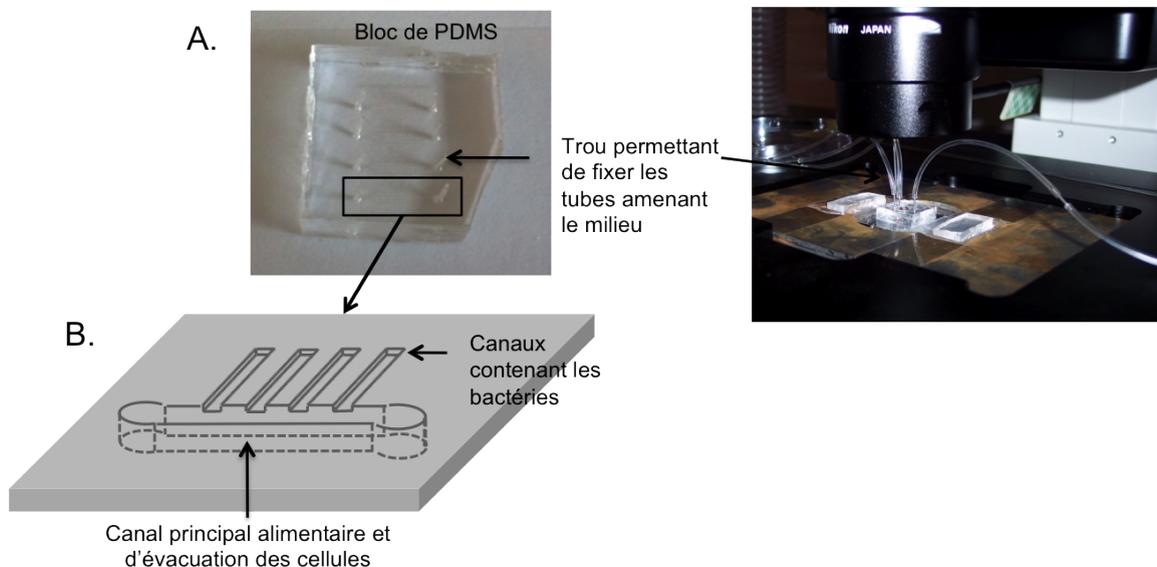


FIGURE 2.6 – **Système de microscopie microfluidique.** **A.** Bloc de PDMS sur lequel sont imprimés les canaux (circuit microfluidique). **B.** Design de notre système microfluidique avec un canal principal permettant d'alimenter les bactéries et de les évacuer et des canaux perpendiculaires contenant les bactéries.

Des tubes sont fixés sur le bloc de PDMS en entrée et en sortie du canal principal afin d'injecter les bactéries en début d'expérience (les bactéries sont insérées dans les canaux par centrifugation du système microfluidique) et ensuite d'amener le milieu nécessaire à la survie des bactéries. Ce milieu circule à flux constant (une pompe permet de contrôler les seringues contenant le milieu) et permet également d'évacuer les bactéries sortant des canaux (Figure 2.7 B., page 47). En effet, quand la cellule mère, qui est au fond du canal, croît puis se divise, elle donne naissance à une cellule fille et pousse les autres cellules vers l'extrémité du canal (à proximité du canal principal).

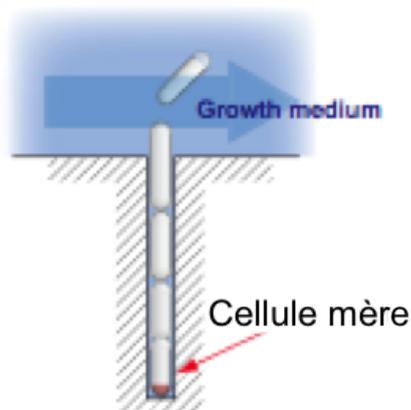


FIGURE 2.7 – Cellule mère et évacuation des cellules.

Au cours d'une même expérience, il est possible d'étudier jusqu'à 4 souches ou 4 conditions de milieu différentes sur plusieurs centaines de générations.

I.3 Suivi de molécules *in vivo*

Localiser et suivre des molécules est une problématique qui a été souvent abordée ces 20 dernières années. Le single particle tracking (SPT) est une technique permettant de suivre le mouvement de molécules individuelles en les labellisant avec des fluorophores [Yildiz et al., 2003]. De nombreux marqueurs fluorescents ont été développés à cette fin.

Depuis le clonage du gène codant la GFP en 1992 [Prasher et al., 1992], les applications ne cessent de se diversifier. Il est désormais possible de visualiser l'expression d'un ou plusieurs gènes directement en microscopie optique en marquant la protéine ciblée avec une protéine de fusion fluorescente (GFP, YFP, RFP [Campbell et al., 2002], etc.). Chaque protéine fluorescente a des propriétés différentes concernant les longueurs d'ondes (émission et excitation), stabilité ou temps de maturation. La review [Shaner et al., 2005] récapitule les différentes caractéristiques des protéines fluorescentes.

Depuis, l'étude la structure tridimensionnelle de la GFP a permis d'apporter des modifications afin d'améliorer la brillance, l'optimisation du repliement par un accroissement de la vitesse de repliement de la protéine et de la formation du chromophore ou encore la stabilité thermique, etc. D'autres protéines telles que la YFP [Nagai et al., 2002] [Shaner et al., 2004] ou la RFP [Shaner et al., 2008] ont également été optimisées.

Cette méthode est de plus en plus utilisée chez les procaryotes, comme chez la bactérie *E. coli* afin d'identifier de faibles niveaux d'expression de gènes [Yu et al., 2006] ou d'étudier la dynamique de facteurs de transcription [Elf et al., 2007] ou encore la stochasticité de l'expression génique [Choi et al., 2008]. Elle est également utilisée pour le suivi de macromolécules afin de caractériser le mouvement d'ARN_m chez la bactérie *E. coli* [Golding and Cox, 2006] ou de protéines chez la bactérie *C. crescentus* [Deich et al., 2004] [Kim et al., 2006] [Bowman et al., 2008].

L'expression de protéines fluorescentes dans la cellule peut être utilisée pour suivre le niveau d'expression d'un gène. Dans ce cas, la protéine fluorescente est insérée par fusion traductionnelle en aval du promoteur du gène d'intérêt. Cela permet de suivre le niveau d'expression du gène mais non sa localisation ou sa dégradation. En revanche une fusion transcriptionnelle (la protéine fluorescente est fusionnée à la protéine d'intérêt) permet de localiser et caractériser le niveau d'expression d'une protéine. Cependant, ce type de fusion peut altérer l'activité et/ou la dynamique de la protéine étudiée (taux de dégradation, localisation, etc.).

De plus, très récemment, Landgraf [Landgraf et al., 2012] a montré la possible interférence des marqueurs fluorescents sur la localisation et l'agrégation de protéines chez la bactérie *E. coli*. L'étude de certaines protéines telles que la protéase ClpX montre une localisation et agrégation très différente selon la protéine fluorescente utilisée (Figure 2.8, page 48). Ils montrent ainsi que la protéine ClpX fusionnée au marqueur fluorescent mTagBFP, Venus, mcherry et mcherry2 ou sfGFP résulte en la formation d'un agrégat protéique alors que la fusion avec le marqueur TagRFP-T entraîne la formation de multiples foci alors que les marqueurs GFP, mGFPmut3 ou mYPet indique la présence de quelques foci mais également d'un état diffus (répartition homogène) de ClpX dans le cytoplasme. Leur outil expérimental est basé sur la comparaison relative de la localisation et agrégation de la protéine cible fusionnée ou non à un marqueur fluorescent. Ces travaux soulignent l'importance du choix de la protéine fluorescente lors du suivi de molécules.

Un autre problème rencontré avec l'utilisation de protéines fluorescentes est le photoblanchiment (photobleaching en anglais). Il se définit par une diminution de la fluorescence émise par le fluorophore à cause de l'excitation précédente. Le photoblanchiment dépend de l'intensité et de la longueur de l'excitation et des propriétés du fluorophore.

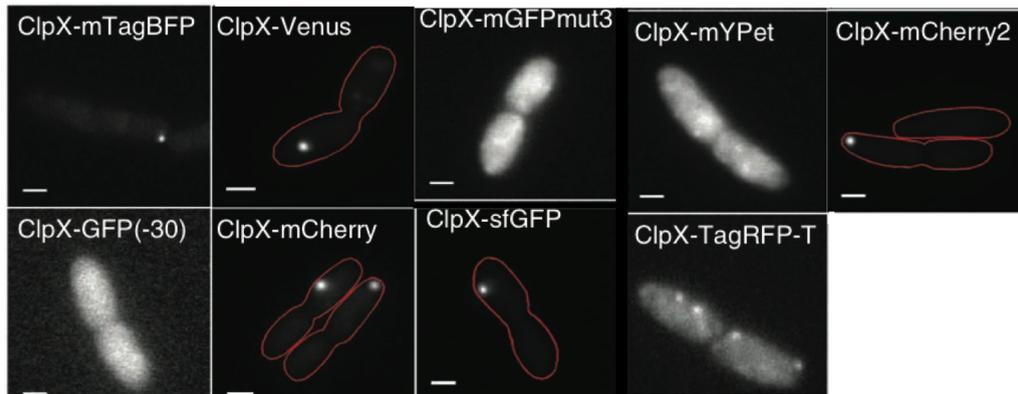


FIGURE 2.8 – Images de fluorescence de la bactérie exprimant la protéine ClpX fusionnées à différents marqueurs fluorescents. Les contours de cellules sont indiqués en rouge pour les cellules qui présentent un faible signal cytoplasmique. (Figure extraite et modifiée de [Landgraf et al., 2012])

II Analyses d'images

L'acquisition automatisée d'images génère des quantités très élevées d'images de différents types. L'extraction de données quantitatives à partir de celles-ci nécessite divers traitements. Le traitement d'images est complexifié par divers paramètres comme la perte d'informations due à la projection en 2d des objets 3d, ou la qualité des images (bruit dû à la qualité du matériel de microscopie mais aussi à l'exposition et au mouvement des objets), ou la nature ou superposition des objets ou encore des images non homogènes au cours du temps (changement du nombre d'objets dans les images, de l'illumination, etc.). Le choix des algorithmes pour le traitement d'images dépendra du type d'images à analyser et des données à extraire.

L'analyse de séquences d'images biologiques nécessite deux étapes principales :

- la segmentation qui permet de définir les frontières spatiales des objets que l'on veut mesurer dans chaque image.
- le suivi d'objets avec le temps afin d'observer leurs évolutions et les variations en terme de mesures.

La segmentation consiste à diviser un objet en plusieurs objets (dans notre cas des bactéries). Il existe plusieurs techniques de segmentation telles que le seuillage, le contour actif, la ligne de partage des eaux, etc. Mais il n'existe aucune méthode universelle car chaque image et/ou problématique demande une technique adaptée. Le but de la segmentation est de partitionner une image en régions correspondant aux objets que l'on souhaite suivre. Les algorithmes les plus simples implémentent un seuil par rapport à l'intensité qui permet de séparer un premier plan plus clair que l'arrière-plan ou le contraire. Cependant, cette méthode ne fonctionne pas dans tous les cas et notamment lorsque deux cellules sont très proches par exemple.

Une fois que les formes des objets ont été identifiées dans les images, l'objectif est de suivre chaque objet à travers le film afin d'effectuer un lien entre les différentes images pour chaque objet. Dans notre cas où nous suivons des cellules bactériennes, une troisième étape consiste à reconstruire le lignage et établir des liens de parenté entre les cellules à partir de la cellule mère.

Afin de pouvoir analyser les trajectoires des agrégats protéiques, deux logiciels d'analyse d'images automatisés ont été développés, Cellst et Bacteria Analyzer, afin de pouvoir segmenter et suivre les cellules à partir des images de contraste de phase et dans un second temps de suivre les agrégats protéiques contenus dans ces cellules à partir des images de fluorescence et des données extraites des images de contraste de phase. Dans la suite de cette section, je détaillerai le fonctionnement des deux logiciels développés.

II.1 Cellst (Cell Segmentation and Tracking)

Cellst est un logiciel d'analyse d'images permettant de segmenter et suivre des bactéries de manière automatisée. Ce logiciel a été développé en collaboration entre l'INSERM U1001, localisé au Centre de Recherche Interdisciplinaires et le laboratoire MAP5, situé à l'Université Paris Descartes par Maël Primet [Primet, 2011] et Alice Demarez [Demarez, 2011] afin d'étudier la croissance de bactéries observées en microscopie time-lapse.

Cellst utilise une approche en deux étapes permettant ainsi de vérifier les résultats intermédiaires obtenus (Figure 2.9, page 50). La première étape est de segmenter partiellement les images de manière statique, ce qui permet de produire une image contenant des éléments appelés graines (dans les sections suivantes, je reviendrai plus en détails sur la définition) qui sont ensuite dilatées pour former des blobs. La seconde étape consiste à segmenter et suivre les cellules. Les blobs sont fusionnés afin de former des cellules et ensuite le suivi des cellules, nommé lignage, au travers du film, est reconstruit. Le lignage consiste à établir les liens entre les cellules, déterminer quelle cellule mère donne naissance à quelle cellule fille, etc. Dans chacune de ces étapes, aucun paramètre ou seuil basé sur l'intensité n'est utilisé. Seules les propriétés géométriques de la cellule (les cellules ont une largeur comprise entre deux réels et leur aire est supérieure à une valeur seuil) sont utilisées ainsi qu'un paramètre lié à son mouvement entre l'image n et $n - 1$. Chaque étape sera détaillée dans les sections suivantes.

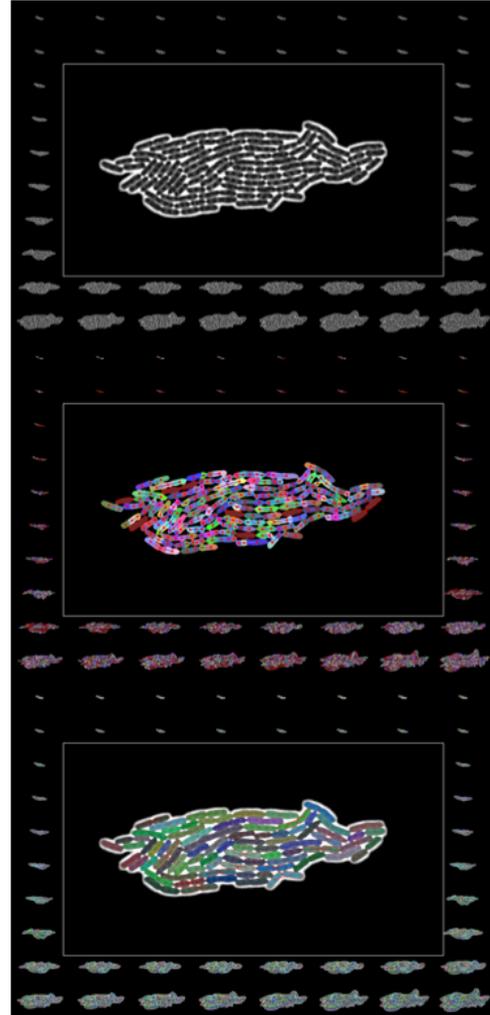


FIGURE 2.9 – Vue d'ensemble de l'algorithme implémenté dans Cellst. Les images sont normalisées afin de définir des graines. Les graines sont dilatées pour obtenir des blobs (les blobs de couleur rouge sont détectés comme cellule). Enfin, les blobs sont connectés afin d'obtenir des cellules.

II.1.1 La segmentation

La sur-segmentation permet de simplifier l'image à un petit nombre d'objets nommés blobs sans avoir à résoudre dans un premier temps des décisions ambiguës. Après un débruitage des images au moyen d'un algorithme de variation totale [Rudin et al., 1992], l'image est normalisée afin d'homogénéiser les niveaux de gris à l'intérieur des cellules. L'idée est de décomposer l'images en éléments connectés, les graines, l'arrière-plan ayant une valeur égale à 0. Pour cela, le niveau de gris est mesuré à différents endroits de l'image près des cellules afin d'estimer le seuil local qui sera utilisé afin de définir les graines et l'espace occupé par les cellules.

Des blobs sont ensuite définis à partir de graines que l'on fait croître par dilatation concurrente (les blobs ne peuvent pas s'interpénétrer) et qui sont définies grâce à un seuillage du niveau de gris (Figure 2.10, page 51). Les régions les plus sombres sont à l'intérieur des cellules.

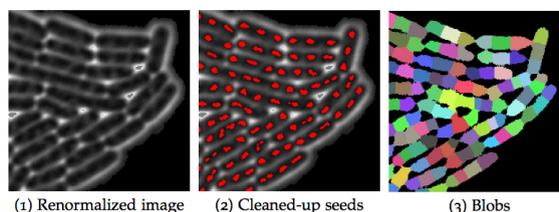


FIGURE 2.10 – **Vue d'ensemble du processus de dilatation.** 1) Image normalisée d'une colonie. 2) Nettoyage des graines détectées en fonction des paramètres morphologiques. 3) Blobs obtenus à partir des graines.

II.1.2 Le suivi des cellules

Les cellules sont segmentées et suivies de manière itérative. Cependant, les images ne sont pas analysées de manière séquentielle afin de pouvoir prendre une décision à n'importe quelle étape.

Les blobs consécutifs sont assemblés en utilisant une approche basée sur la notion de risque (Figure 2.11, page 52). Afin de conserver un algorithme simple, l'algorithme fait des choix locaux en les ordonnant de façon à commencer par les choix les plus évidents et les plus simples et ne comportant aucun risque afin de minimiser les erreurs. Cela permet ensuite de réduire la complexité et le risque pour les choix qui présentaient une complexité élevée au départ. Les ambiguïtés sont levées au fur et à mesure. Par exemple, les choix les plus évidents se font sur les cellules en bordure de colonies qui comportent peu de voisins. Ensuite, il est possible de regarder les cellules plus à l'intérieur mais qui se retrouvent en réalité en bordure car celles qui étaient réellement en bordure ont été segmentées en premier.

Afin de suivre la transition d'une cellule A dans l'image n à la cellule B dans l'image n+1, les coordonnées du centre de masse de la cellule et son orientation sont calculés. Des corrections manuelles permettent de supprimer des cellules ou de corriger le lignage.

L'analyse d'un film de 301 images nécessite environ 4 heures incluant les corrections manuelles qui prennent le plus de temps. Afin d'avoir des données significatives sur le mouvement des agrégats protéiques d'un point de vue statistique, 82 films ont été analysés avec l'aide de ce logiciel.

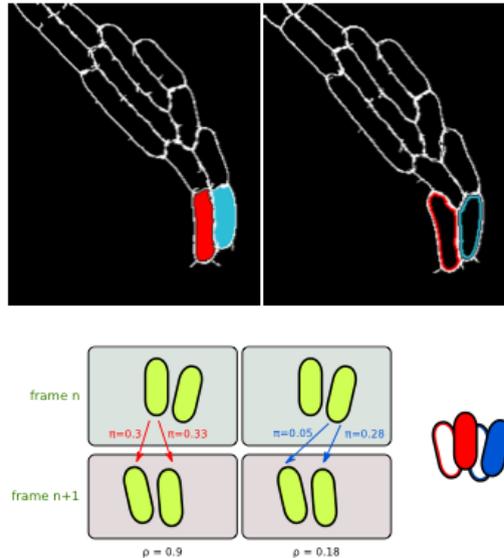


FIGURE 2.11 – **Notion de risque.** Un algorithme utilisant la probabilité de transition prendra la mauvaise décision en commençant par la cellule rouge car les probabilités que la cellule rouge dans la frame n devienne la cellule de gauche dans la frame $n + 1$ ($n = 0.3$) ou la cellule de droite ($n = 0.33$) sont très proches contrairement à la cellule bleu dont les probabilités de transitions ($n = 0.05$ et $n = 0.28$) sont différentes. En revanche, un algorithme classant les cellules en fonction du risque de transition ($p=0.9$ et $p=0.18$ dans notre cas) prendra la bonne décision. La probabilité de risque sera égale à 0 s'il n'y a qu'un successeur possible et le risque sera élevé quand il y aura un doute. Si l'on considère la division d'une cellule A (rouge) en cellule B (gauche) et C (droite), le risque est défini par (probabilité de transition de A vers B soit 0.3)/(probabilité de transition de A vers C soit 0.33), ce qui nous donne une valeur de 0.9.

II.1.3 Caractérisation et suivi des agrégats

Les agrégats de protéines fluorescentes ou spots ont été détectés par un autre logiciel développé par l'équipe de Lionel Moisan (Jacob, Dimiccoli, Moisan 2012).

Pour chaque spot, le pixel d'intensité maximum est détecté par un algorithme basé sur la méthodologie *a contrario*. La détection *a contrario* consiste à mesurer la surprise d'observer une mesure dans l'image, d'un point de vue probabiliste. Elle nécessite ainsi une mesure et une distribution naïve de probabilité. Pour la distribution naïve de probabilité, on choisit la distribution gaussienne. Pour la mesure, on compare les moyennes d'intensité en utilisant deux cercles concentriques (Figure 2.12, page 53) : si le cercle le plus interne est bien centré sur le spot, alors sa moyenne d'intensité est très supérieure à celle dans le cercle le plus externe. Grâce à cette mesure, on optimise le rayon du cercle le plus interne, le rayon optimal étant celui qui donne la plus grande mesure. Cela permet d'obtenir une segmentation optimale (contour circulaire) pour chaque spot détecté.

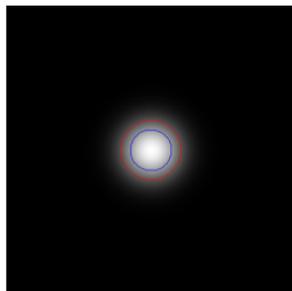


FIGURE 2.12 – **Utilisation de la méthode *a contrario***. Elle utilise deux cercles concentriques dont le plus interne pour la détection du spot et le plus externe pour le contexte local.

Les coordonnées des spots détectés sont calculées et ensuite affinées au sous-pixel de résolution. Ceci est réalisé en calculant une moyenne pondérée des coordonnées des pixels du cercle interne (cf. ci-dessus) du spot détecté. Seuls les pixels ayant une intensité plus grande que l'intensité médiane des pixels du cercle le plus externe, sont pris en compte dans la moyenne pondérée. La précision est de 1/10ème de pixel sur des agrégats peu contrastés.

Après la détection et la localisation dans l'image, les agrégats (spots) sont suivis par un troisième logiciel nommé Aggtracker (Primet, Moisan 2011) qui utilise le lignage des bactéries et les spots détectés précédemment.

Aggtracker permet ainsi de suivre les agrégats contenus dans chaque cellule au cours du temps. Des cas d'ambiguïtés peuvent apparaître à savoir qu'un agrégat se situe entre deux cellules (qui sont accolées) et dans ce cas, une intervention manuelle de l'utilisateur est nécessaire afin de décider à quelle cellule l'agrégat appartient.

Une dernière étape permet de projeter les coordonnées de chaque agrégat sur les axes de la cellule à laquelle il appartient. Afin de déterminer précisément les contours de cellule, la méthode des squelettes actifs a été utilisée (Jacob, Moisan2011). La bactérie est représentée par une ligne médiane principale selon la longueur de la bactérie (Figure 2.13, page 53). Le squelette permet ainsi de déterminer précisément la longueur et la largeur (avec une seconde médiane selon la largeur) de la bactérie et les coordonnées précises des agrégats dans la cellule.

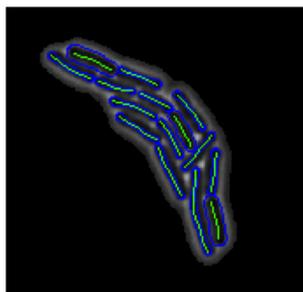


FIGURE 2.13 – **Utilisation de la méthode des squelettes actifs**. Elle permet de déterminer précisément la longueur et la largeur des bactéries afin de projeter ensuite les coordonnées des agrégats sur les axes de la cellule.

II.2 Bacteria Analyzer

Ce logiciel est un plugin ImageJ et a été développé par Xiaohu Song, ingénieur de recherche en développement et déploiement d'applications pour l'analyse d'images à l'INSERM U1001. Il permet d'analyser des images issues de microscopie microfluidique.

II.2.1 Traitement d'images

Les images analysées sont des images de contraste de phase contenant des bactéries dans des canaux et des images de fluorescence contenant des agrégats protéiques de différentes formes (circulaires et allongées). Afin de simplifier l'étude de chaque canal au cours du temps, une nouvelle séquence d'images est créée pour chaque canal en fonction du temps (Figure 2.14, page 54) et ce pour les images de contraste de phase et de fluorescence. L'analyse s'effectue ainsi canal par canal.

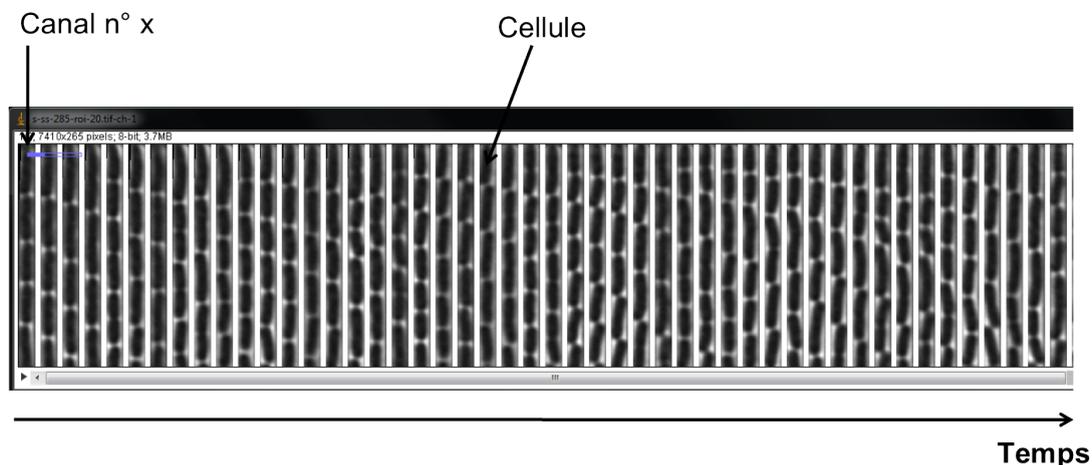


FIGURE 2.14 – Séquence d'images créée pour un canal en fonction du temps pour les images de contraste de phase. Chaque forme noire représente une cellule.

Un filtre moyenneur (low pass) 3x3 (pixels) est d'abord appliqué à l'ensemble des images) avec le logiciel Metamorph (<http://www.moleculardevices.com/Products/Software/Meta-Imaging-Series/MetaMorph.html>). Puis les images sont normalisées en soustrayant l'arrière-plan de l'image. Ensuite, les images sont stabilisées au moyen de l'algorithme Lucas-Kanade développé sous la forme du plugin Image Stabilizer (http://www.cs.cmu.edu/~kangli/code/Image_Stabilizer.html). Les paramètres de stabilisation obtenus, issus des images de contraste de phase sont également appliqués aux images de fluorescence.

Enfin, les images de contraste de phase sont débruitées en utilisant un algorithme de variation totale [Rudin et al., 1992]. Le principe est de réduire le bruit de l'image en conservant les informations de contour en fonction de la variation totale de l'image gradient et l'image originale.

II.2.2 La segmentation et le suivi des cellules

La segmentation est effectuée en employant la technique du seuillage de l'histogramme local automatiquement et global manuellement. Le seuillage de l'histogramme local est calculé une

fois la segmentation de la cellule effectuée (au moyen de l'histogramme global par l'utilisateur). (Figure 2.15, page 55).

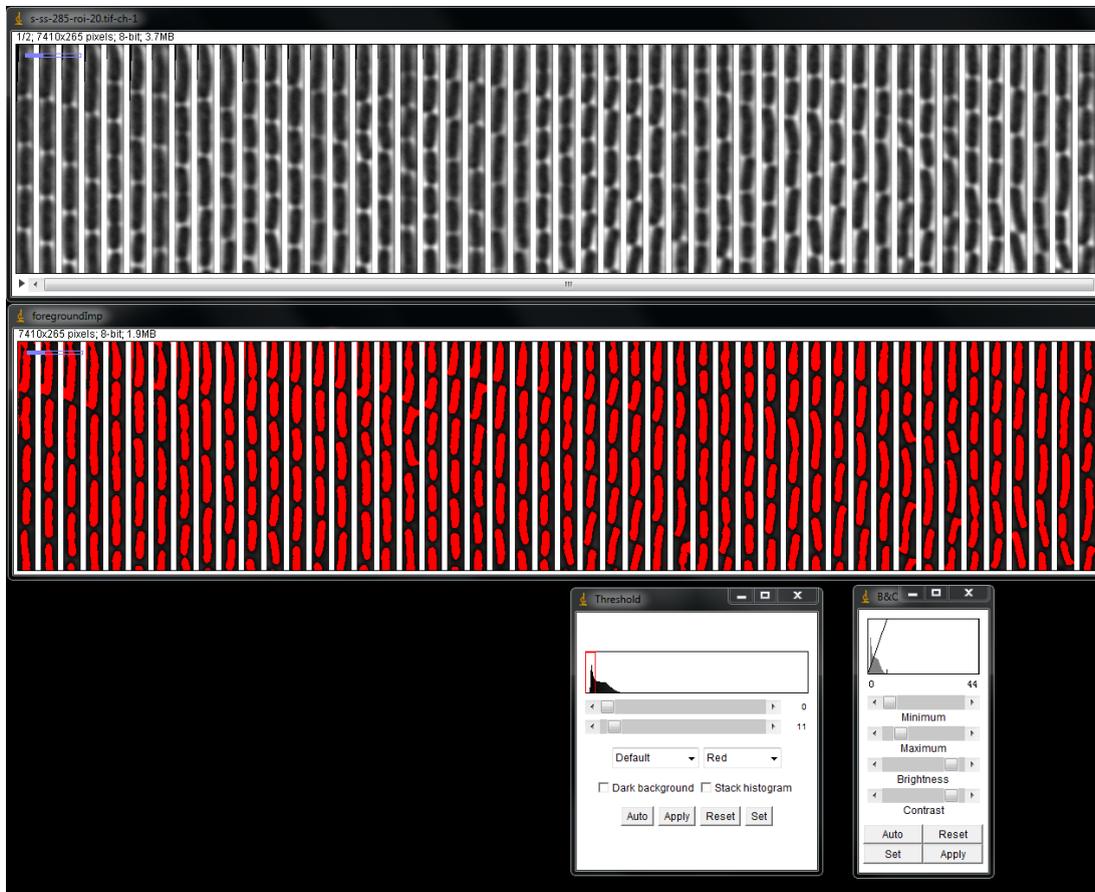


FIGURE 2.15 – Exemple de segmentation pour un canal en fonction du temps. Utilisation de l'histogramme global par l'utilisateur.

Le lignage des cellules (lien entre chaque cellule) (Figure 2.16, page 56) à partir de la cellule mère est basé sur un paramètre morphologique qui est la longueur de la cellule. Quand la cellule se divise (image n), la longueur totale des deux nouvelles cellules (image $n+1$) doit être supérieure à 90% de la longueur de la cellule avant division (image n).

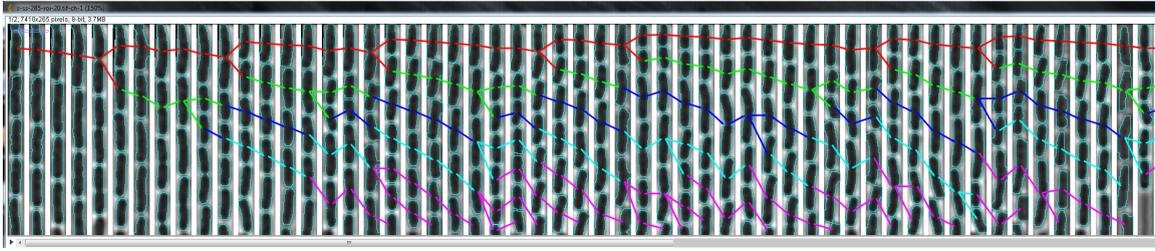


FIGURE 2.16 – Exemple de lignage pour un canal en fonction du temps. Le lignage est effectué à partir de la cellule mère située au fond du canal.

II.2.3 Caractérisation et suivi des agrégats

La séquence d'images de fluorescence est superposée à celle d'images de contraste de phase. Dans notre étude, on distingue deux types d'agrégats protéiques en fonction de leur circularité. Si la circularité (division de la surface réelle et la surface d'un cercle théorique, $4 * \pi(\text{surface}/\text{perimetre}^2)$) est supérieure au seuil déterminé (0.6 définie par les données expérimentales) alors c'est un agrégat de type "foci" (forme plus ou moins ronde). En revanche, si la circularité est inférieure au seuil, c'est un agrégat de type "comet" dont la forme est proche d'une virgule très allongée (Figure 2.17, page 56).

Pour chaque cellule, on identifiera de manière automatisée les paramètres morphologiques de la cellule tels que la longueur et largeur ou la position du centre de masse, et les caractéristiques liées aux agrégats comme le nombre et type d'agrégats contenu dans la cellule ainsi que l'intensité moyenne de chaque agrégat ou sa position.

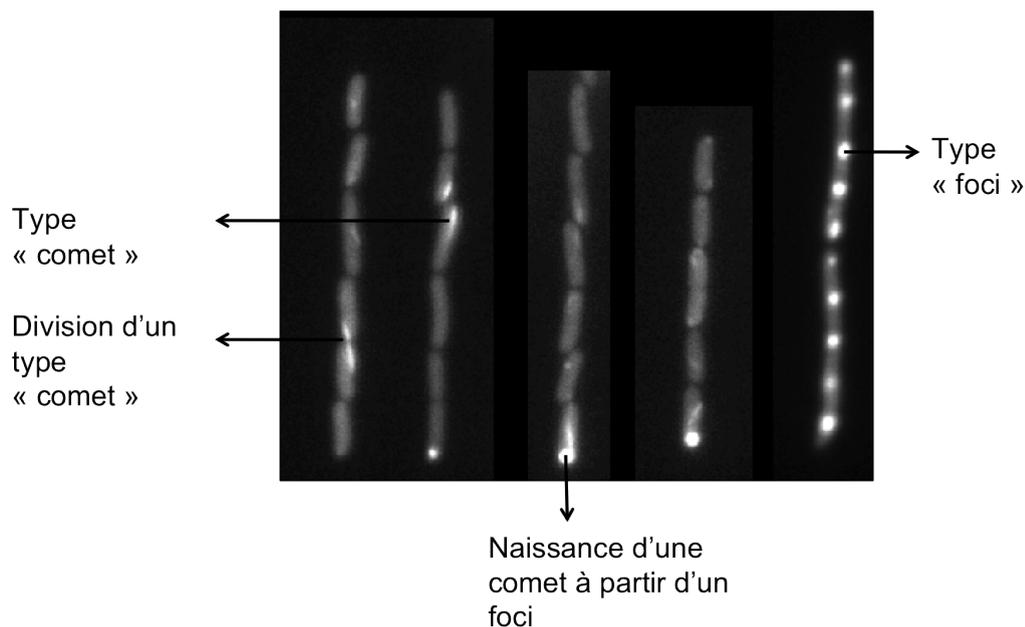


FIGURE 2.17 – Types d'agrégats protéiques : foci et comet. Ce sont des canaux isolés qui n'ont aucun lien entre eux.

II. ANALYSES D'IMAGES

Des corrections manuelles sont parfois nécessaires et essentielles sur les images de contraste de phase afin d'ajouter/séparer ou supprimer des cellules afin de reconstruire un lignage correct, et également sur les images de fluorescence afin de corriger des erreurs de détection ou de type d'agrégat.

III La modélisation

Un modèle est une image simplifiée de la réalité qui nous permet de comprendre le fonctionnement d'un système en fonction d'une question. Tout modèle est constitué d'une part de la description de la structure du système et d'autre part de la description des fonctionnements et dynamiques qui modifient la structure au cours du temps.

La description de la structure d'un système peut aller d'un ensemble de variables quantitatives ou qualitatives jusqu'à un ensemble d'entités munies de propriétés et de relations entre elles.

Face à un système complexe, on réalise des modèles pour :

- contrôler la dynamique et la viabilité d'un système,
- faire apparaître des équilibres et des attracteurs entre les éléments,
- aider à la décision,
- prédire des évolutions possibles.

III.1 Le modèle multi-agents (SMA)

La modélisation multi-agents est une approche apparue dans les années 1980. Elle repose sur l'hypothèse que les propriétés d'un groupe d'individus sont déterminées par les individus qui le composent. Les systèmes multi-agents ont vocation d'être des modèles explicatifs basés sur les entités. C'est une représentation explicite des comportements individuels.

On appelle agent une entité physique ou virtuelle [Ferber, 1995]

- qui est capable d'agir dans un environnement,
- qui peut communiquer directement avec d'autres agents,
- qui est mue par un ensemble de tendances (sous la forme d'objectifs individuels ou d'une fonction de satisfaction, voire de survie, qu'elle cherche à optimiser),
- qui possède des ressources propres,
- qui est capable de percevoir (mais de manière limitée) son environnement,
- qui ne dispose que d'une représentation partielle de cet environnement,
- qui possède des compétences et offre des services,
- qui peut éventuellement se reproduire,
- dont le comportement tend à satisfaire ses objectifs, en tenant compte des ressources et des compétences dont elle dispose, et en fonction de sa perception, de ses représentations et des communications qu'elle reçoit.

On appelle système multi-agent [Ferber, 1995], un système composé des éléments suivants :

- Un environnement E , c'est-à-dire un espace disposant généralement d'une métrique,
- Un ensemble d'objets O . Ces objets sont situés, c'est-à-dire que, pour tout objet, il est possible, à un moment donné, d'associer une position dans E . Ces objets sont passifs, c'est-à-dire qu'ils peuvent être perçus, créés, détruits et modifiés par les agents,
- Un ensemble A d'agents, qui sont des objets particuliers, lesquels représentent les entités actives du système,
- Un ensemble de relations R qui unissent des objets (et donc des agents) entre eux,
- Un ensemble d'opérations Op permettant aux agents de A de percevoir, produire, consommer, transformer et manipuler des objets de O ,

- Des opérateurs chargés de représenter l'application de ces opérations et la réaction du monde à cette tentative de modification, que l'on appellera les lois de l'univers.

Les modèles multi-agents représentent l'ensemble des éléments du système ainsi que leurs interactions ou encore l'ensemble des molécules et leurs déplacements aléatoires.

L'agent peut communiquer avec d'autres agents et se déplacer. Le temps est directement intégré sous la forme d'une horloge qui rythme les séquences d'actions des agents.

On cherche à déterminer l'évolution du système afin de prévoir l'organisation qui en résulte.

L'avantage d'un système multi-agents est d'une part sa modularité que l'on peut définir par l'ajout ou le retrait d'agents mais aussi l'amélioration ou l'affinement des agents. Les SMA permettent la modélisation de situations complexes dont les structures globales émergent des interactions entre individus.

III.2 Le modèle basé individus (IBM)

Les modèles basés individus sont des systèmes multi-agents réactifs. Chaque agent est régi par quelques lois/règles très simples. Ils sont utilisés lorsque l'on s'intéresse à la dynamique de l'ensemble d'une population ou d'un groupe tout en modélisant chaque individu ou entité. Ils représentent les interactions locales entre les individus ou entités tout en gardant une observation globale. Le fonctionnement des IBM repose sur deux dimensions : l'espace et le temps. La dimension spatiale met en relation le comportement d'un agent avec celui d'un autre agent. La dimension temporelle met en relation un comportement apparaissant à un instant donné avec un second comportement apparaissant plus tard.

Le modèle définit les individus, mais le résultat est observé au niveau d'un ensemble d'individus (pour lequel aucune propriété moyenne n'a été définie). On modélise des individus ou entités dont la dynamique de groupe nous intéresse.

Dans un IBM :

- L'état de l'individu est représenté par une ou plusieurs variables.
- Son comportement est décrit par des règles logiques ou des fonctions mathématiques.
- L'état d'un individu change en fonction du temps et de son environnement (interactions avec d'autres individus et avec l'environnement physique).
- Le niveau de détail peut varier énormément d'un modèle à l'autre et dépend de l'objectif.
- Ces modèles permettent l'intégration de paramètres mesurés par les biologistes.

Un IBM [Bonabeau, 2002] a pour objectif de capturer des phénomènes émergents qui sont le résultat des interactions entre individus, d'apporter une description naturelle du système et d'être flexible.

Depuis les années 1990, les modèles basé individus sont de plus en plus utilisés notamment en sciences sociales, économie ou encore en écologie.

III.3 Notre modèle

Afin de simuler le processus de diffusion/agrégation dans des cellules uniques d'*E. coli*, nous avons utilisé un modèle 3d basé individus sans treillis (lattice free).

Chaque monomère (monomère IbpA-YFP en comparaison aux expériences *in vivo*) a été modélisée comme une sphère de rayon r_0 et de coordonnées (x_p, y_p, z_p) dans l'espace 3d intracellulaire de la cellule. La bactérie est représentée par un rectangle 3d dont la largeur et la profondeur sont de $1 \mu\text{m}$ et la longueur de $4 \mu\text{m}$. Nous avons lancé des simulations avec une forme plus réaliste et proche de la bactérie *E. coli* et des formes arrondies au pôle. Les résultats étaient inchangés à l'exception du temps de calcul qui était grandement plus long avec la forme réaliste.

Dans la cellule, nous avons également modélisé l'encombrement moléculaire dans les nucleoides qui est plus élevé que dans le reste du cytoplasme [Foley et al., 2010]. Afin de modéliser l'encombrement moléculaire dans les nucléoides, nous avons placé au hasard 50 000 obstacles immobiles et impénétrables de diamètre 10 nm dans les régions de la cellule où se situent *in vivo* les nucléoides. En phase exponentielle, plus de 75% des cellules contiennent deux nucléoides du fait que le cycle cellulaire et la réplication de l'ADN ne sont pas synchronisés chez la bactérie *E. coli* [Donachie and Begg, 1989]. La position et la taille des nucléoides ont été estimées à partir d'images de fluorescence utilisant des cellules marquées au DAPI [Zimmerman, 2006]. Chaque nucléoïde est représenté par un rectangle 3d de longueur 1220 nm et de largeur et hauteur de 532 nm. Chaque nucléoïde est positionné à 540 nm du pôle (Figure 2.18, page 60) . Le volume occupé par les nucléoides est d'environ 12% ce qui est en accord avec la littérature [Woldringh and Odijk, 1999] [Odijk, 1998].

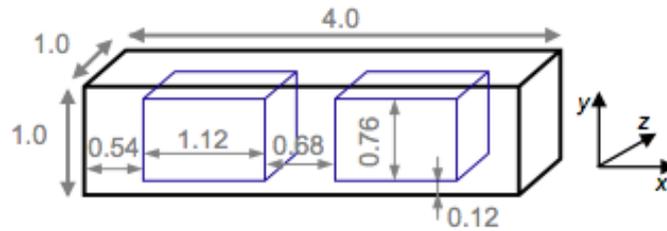


FIGURE 2.18 – Représentation géométrique de notre modèle.

Chaque simulation est initialisée avec un nombre de monomères (représentant des monomères de la protéine chaperonne IbpA fusionnée à la protéine fluorescente jaune YFP, utilisée par Lindner *et al.*) défini variant de 100 à 7000 permettant ainsi de simuler des conditions naturelles ou des conditions de stress telles que le choc thermique. Les monomères sont placés au hasard dans les régions libres et ne peuvent pas se superposer. A chaque pas de temps de 10^{-6} s, chaque molécule se déplace indépendamment sur une distance w qui dépend de la constante de diffusion de la protéine D_p et selon l'équation du mouvement Brownien :

$$w = (6D_p\delta t)^{\frac{1}{2}} \quad (2.4)$$

avec δt le pas de temps. On notera que D_p dépend de la taille de l'agrégat.

Le pas de temps a été choisi afin que qu'une molécule ne puisse pas sauter par dessus une autre durant ce pas de temps. La distance parcourue durant un pas de temps doit donc être inférieure à $4r_0$. En utilisant la définition de w (Equation 2.4, page 60), le pas de temps doit être inférieur

à $5 \mu s$.

La nouvelle position $(x'(t), y'(t), z'(t))$ d'une molécule est calculée au moyen de deux angles choisis de manière aléatoire θ et ϕ et uniformément distribués entre $[0, 2\pi]$ et $[-1, 1]$ respectivement, sachant que la distance d est fixée. Afin de calculer les nouvelles coordonnées de chaque molécule, nous avons utilisé les coordonnées sphériques (Figure 2.19, page 61) :

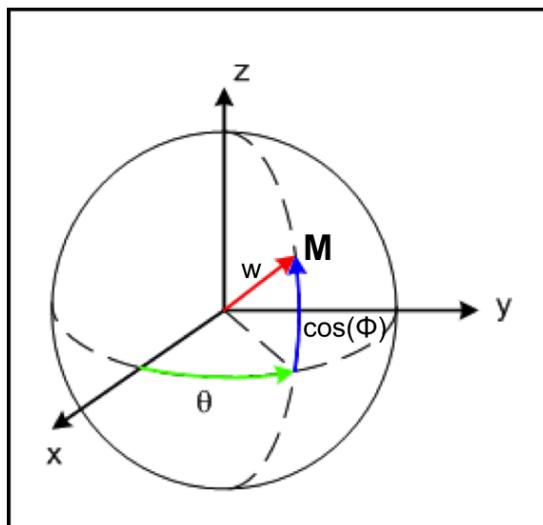


FIGURE 2.19 – Coordonnées sphériques d'un point M.

Les coordonnées sphériques sont calculées à partir des équations suivantes :

$$x' = x(t) + w \sin(\arccos(\phi)) \cos(\theta) \quad (2.5)$$

$$y' = y(t) + w \sin(\arccos(\phi)) \sin(\theta) \quad (2.6)$$

$$z' = z(t) + w\phi \quad (2.7)$$

$x(t)$, $y(t)$ et $z(t)$ sont les coordonnées de la position initiale de la molécule. La distance d est calculée grâce à l'équation 2.4 (page 60) où l'on connaît le pas de temps δt et le coefficient de diffusion de la molécule D .

Si la nouvelle position d'un monomère ou agrégat est occupée par un obstacle immobile, le déplacement est annulé. Cette simplification ne doit pas changer significativement les résultats mais il est certain qu'elle diminue grandement le temps de calcul.

La deuxième étape consiste à analyser les agrégations possibles. Deux monomères ou agrégats peuvent s'agréger si la distance entre leurs centres est plus petite que la somme de leur rayons. Nous avons ajouté une probabilité d'agrégation lors de la rencontre de deux agrégats ou monomères. Cette probabilité varie de 0.1 à 1 et est indépendante de la taille de du monomère ou

agrégat.

Lorsqu'il y a agrégation entre deux molécules A et B, le rayon du nouvel agrégat C suit l'équation :

$$r_C = (r_A^3 + r_B^3)^{\frac{1}{D_f}} \quad (2.8)$$

avec r_A le rayon de la molécule A, r_B le rayon de la molécule B, D_f représente la dimension fractale des agrégats et dans notre cas nous avons choisi une valeur de 3 pour D_f car nous avons émis l'hypothèse d'une forme sphérique de l'agrégat tout au long des simulations avec une densité interne constante.

Le coefficient de diffusion des agrégats est déterminé par l'équation de Stokes-Einstein où le coefficient de diffusion est inversement proportionnel au rayon. Dans notre cas :

$$D_p = D_0 \left(\frac{r_0}{r_p} \right)^\alpha \quad (2.9)$$

avec r_0 et D_0 le rayon et le coefficient de diffusion d'un monomère et $\alpha = 1$ dans notre cas.

On notera que dans notre modèle, l'agrégation est irréversible. Cela est en adéquation avec nos observations *in vivo* où nous n'avons jamais observé de décroissance de la fluorescence des agrégats.

Afin de déterminer le coefficient de diffusion de notre monomère IbpA-YFP ainsi que son rayon, nous avons utilisé les données suivantes. Le coefficient de diffusion de la GFP, dont le diamètre est de 2 nm et le poids moléculaire de 26 kDa, dans le cytoplasme de la bactérie *E. coli* est d'environ $7 \mu m^2/s$ [Elowitz et al., 1999]. Celui de la GFP-MBP de poids moléculaire 72 kDa est d'environ $2.5 \mu m^2/s$. En utilisant ces données et l'équation de Stokes-Einstein, combinées à l'hypothèse d'une forme sphérique, nous avons estimé que le rayon d'un monomère IbpA-YFP, r_0 , dont le poids moléculaire est de 39 kDa, est de 3 nm et son coefficient de diffusion de $4.4 \mu m^2/s$.

Chaque simulation a été lancée sur une durée de $2 \cdot 10^6$ pas de temps (2 s temps réel). Afin d'obtenir des résultats significatifs, pour chaque paramètre et conditions (probabilité d'agrégation, nombre de molécules initiales), 1000 simulations ont été effectuées. Pour chacune, les monomères sont placés au hasard au début de la simulation.

Résultats

Sommaire

I	Article : Dynamique de l'agrégation protéique chez <i>E. coli</i>	64
I.1	Résumé	64
I.2	Article	64
II	Article : Effet de la protéine chaperonne DnaK sur la transmission du prionoïde RepA-WH1	104
II.1	Résumé	104
II.2	Article	104

I Article : Dynamique de l'agrégation protéique chez *E. coli*

I.1 Résumé

Le vieillissement chez la bactérie *Escherichia coli*, comme dans beaucoup de maladies liées au vieillissement, est lié à l'accumulation d'agrégats protéiques à des localisations spécifiques, le pôle. La bactérie *E. coli* étant un organisme modèle avec un fort potentiel, comprendre le vieillissement et l'agrégation protéique chez cette bactérie pourrait apporter des avancées significatives afin de comprendre le vieillissement chez l'Homme. Le premier obstacle est de comprendre les mécanismes par lesquels la cellule peut adresser un agrégat protéique à une localisation spécifique dans l'espace intracellulaire. En utilisant une approche intégrée basée sur la microscopie de fluorescence en time-lapse de cellules d'*E. coli in vivo*, un logiciel d'analyse d'images automatisée innovant, et des modèles basés individus, nous montrons que le mouvement des agrégats liés au vieillissement chez la bactérie *E. coli* est purement diffusif (mouvement Brownien). En suivant dans chaque cellule l'agrégat protéique "en live", nous avons pu estimer la taille moyenne des agrégats et leurs coefficients de diffusion ainsi que la distance à partir de laquelle leur mouvement est confiné. Nos résultats montrent sans ambiguïté que les agrégats se déplacent par un mouvement Brownien avec un confinement (diffusion passive) avec un coefficient de diffusion qui dépend de leur taille, ce qui est en accord avec la loi de Stokes-Einstein. Cependant le déplacement des agrégats selon l'axe long de la cellule fait apparaître un confinement dont la taille correspond à l'espace dépourvu d'ADN au pôle. Ceci confirme l'importance d'augmenter l'encombrement moléculaire dans les nucléoides. Enfin, nous avons utilisé des modèles 3D basés individus afin de montrer que la diffusion associée à une agrégation irréversible et un encombrement important dans les nucléoides, sont suffisants afin de reproduire les données expérimentales obtenues sur la localisation des agrégats dans la cellule. L'ensemble de nos résultats supporte l'hypothèse que la localisation polaire des agrégats liés au vieillissement chez la bactérie *E. coli* résulte d'une diffusion et agrégation passives dans un espace non homogène doté d'un encombrement moléculaire élevé.

I.2 Article

Localization of protein aggregation in Escherichia coli is governed by
diffusion and nucleoid macromolecular crowding effect.

A.-S. Coquel^{1,3,4}, J.-P. Jacob⁵, M. Primet⁵, A. Demarez¹, M. Dimiccoli⁵, T. Julou⁶, L.
Moisan⁵, A.B. Lindner^{1,2*}, H. Berry^{3,4*}

1. Institut National de la Santé et de la Recherche Médicale, Unité 1001, 75015 Paris, France
2. Faculty of Medicine, Paris Descartes University, 75015 Paris, France
3. EPI Beagle, INRIA Rhone-Alpes, 69603 Villeurbanne, France.
4. University of Lyon, LIRIS UMR5205 CNRS, 69621 Villeurbanne, France
5. University Paris Descartes, MAP5 - CNRS UMR 8145, 75270 Paris, France
6. Laboratoire de Physique Statistique de l'École Normale Supérieure, UMR 8550 CNRS, 24 rue Lhomond, 75005 Paris, France

* .Equal contribution; corresponding authors: ariel.lindner@inserm.fr; hugues.berry@inria.fr;

ABSTRACT

Aggregates of misfolded proteins are a hallmark of many age-related diseases. Recently, they have been linked to aging of *Escherichia coli* (*E. coli*) where protein aggregates accumulate at the old pole region of the ageing bacterium. Because of the huge potential of *E. coli* as a model organism, elucidating aging and protein aggregation in this bacterium may pave the way to significant advances in our global understanding of aging. A first obstacle along this path is to decipher the mechanisms by which protein aggregates are targeted to specific intercellular locations.

Here, using an integrated approach based on individual-based modeling, time-lapse fluorescence microscopy *and* innovative automated image analysis, we show that the movement of aging-related protein aggregates in *E. coli* is purely diffusive (Brownian). Using single-particle tracking of protein aggregates in live *E. coli* cells, we estimated the average size and diffusion constant of the aggregates. Our results unambiguously evidence that the aggregates passively diffuse within the cell, with diffusion constants that depend on their size in plain agreement with the Stokes-Einstein law. However, the aggregate displacements along the cell long axis are confined to a region that roughly corresponds to the nucleoid-free space in the cell pole, thus confirming the importance of increased macromolecular crowding in the nucleoids. We thus used 3d individual-based modeling to show that these three ingredients (diffusion, aggregation and diffusion hindrance in the nucleoids) are sufficient and necessary to reproduce the available experimental data on aggregate localization in the cells. Taken together, our results strongly support the hypothesis that the localization of aging-related protein aggregates in the poles of *E. coli* results from the coupling of passive diffusion-aggregation with spatially nonhomogeneous macromolecular crowding. They furthermore uncover the importance of "soft" intracellular structuration (based on macromolecular crowding) in diffusion-based protein localization in *E. coli*.

INTRODUCTION

While aging is usually obvious in metazoans, its identification in single-cell organisms is more delicate. For those with frank asymmetrical division, such as the budding yeast *Saccharomyces cerevisiae* or the bacterium *Caulobacter crescentus*, aging usually goes with the presence of a clearly identifiable juvenile (the bud in *S. cerevisiae*) [Jazwinski *et al.* 2002; Ackerman *et al.* 2003]. Aging nevertheless also exists in bacteria that divide through a process symmetrical from the morphology viewpoint, such as *Escherichia coli* and *Bacillus subtilis*. Although they are morphologically identical, the two daughter cells left after the division of *E. coli* are *functionally* different and can be distinguished on the basis of the age of their poles (the hemispherical cell caps): the sibling that inherited the older cell pole in the division grows slower than the other [Stewart *et al.* 2005; Lindner *et al.* 2008; Veening *et al.* 2008 (for *B. subtilis*); Lele *et al.* 2011]. Lindner *et al.* 2008; Rokney *et al.* 2009; Winkler *et al.* 2010

This replicative aging phenomenon can be interpreted as resulting from an asymmetric inheritance of damages accumulating with time. Indeed, like in many age-related diseases in animals [Lindner *et al.* 2009], aging in *E. coli* is related to accumulating protein aggregation and preferential segregation of the large protein aggregates to the sibling that inherits the older cell pole [Lindner *et al.* 2008; Rokney *et al.* 2009; Winkler *et al.* 2010]. Preferential sequestration of damaged proteins is also observed in *S. cerevisiae* between the bud and the mother cell [Aguilaniu *et al.* 2003; Erjavec *et al.* 2008; Zhou *et al.* 2011] and between specific intracellular compartments in yeast and mammalian cell [Kaganovich *et al.* 2008; Rujano *et al.* 2006]. Therefore spatial localization, i.e. the non-homogeneous distribution of damaged protein aggregates in the intracellular space has been postulated to be an optimized strategy allowing cell populations to maintain large growth rates in the face of the accumulation of damages that accompany metabolism during cell life [Watve *et al.* 2011; Erjavec *et al.* 2008]. These results suggest that aging presents conserved processes along the tree of life from bacteria to human, including spatial localization of damaged protein aggregates. Given the formidable advantage of *E. coli* as a model organism for biology, elucidating aging and protein aggregation in *E. coli* raises high hopes for our understanding of aging in humans.

A first obstacle along this path is to understand the mechanisms by which cells can localize protein aggregates at specific locations within their intracellular space. Generally, thermal agitation and the resulting diffusion (Brownian movement) of proteins forbids localization in space on the long run, since diffusion is a mixing process that renders every accessible position equiprobable at long time. Inside eukaryotic cells, active mechanisms such as directed transport or sub-compartmentalization by internal membranes permit to counteract

the uniformization effects of diffusion. It is however known since the 1952 seminal paper by Alan Turing [Turing 1952] that subtle interactions between chemical reactions and diffusion can spontaneously lead to steady states with non-uniform spatial extension. This is also true for bacteria, as exemplified by the spatial oscillations in the minCDE system [Kruse *et al.* 2007] or in the case of diffusion-trapping coupling [Wang *et al.* 2008]. In fact, recent years have witnessed the emergence of the idea that precise sub-cellular localization of proteins is significant in bacteria [Shapiro *et al.* 2009; Iñiesta *et al.* 2010] Since bacteria are traditionally not expected to possess cytoskeleton-based directed transport nor internal membranes, this would favor diffusion-reaction based localization (see however [Sun *et al.* 2011; Domínguez-Escobar *et al.* 2011; Landgraf *et al.* 2012]).

In the specific case of the preferential localization of damaged proteins in single-cell organisms, it is still not clear if the localization mechanism is based on some active directed transport or passive Brownian diffusion. In *S. cerevisiae*, initial reports incriminated a role for active directed transport (actin cytoskeleton) or sub-compartmentalization (membrane tethering) in the segregation of molecular damages (damaged proteins, episomal DNA) in the mother cell [Aguilaniu *et al.* 2003; Shcheprova *et al.* 2008]. Yet, more recent reports contradict the need for directed transport, e.g. on the actin cable, and favor diffusion-based localization [Zhou *et al.* 2011; Gehlen *et al.* 2011].

In *E. coli*, the aging-related protein aggregates have consistently been reported to localize in the cell poles and in the middle of the cell [Lindner *et al.* 2008; Rokney *et al.* 2009; Winkler *et al.* 2010]. The number of distinct aggregates per cell seems to depend on the cellular environment: in non-stressed conditions, at most one aggregate per cell is observed [Lindner *et al.* 2008], whereas under heat shock, most cells contain two or three aggregates [Rokney *et al.* 2009; Winkler *et al.* 2010]. In heat-shock conditions, Winkler *et al.* [Winkler *et al.* 2010] conclude in favor of a Brownian passive motion of the aging-related protein aggregates. This study also points out that one simple possible passive mechanism to emphasize the localization of the aggregates may be spatially non-homogeneous macromolecular crowding. Indeed, in healthy cells, the bacterial chromosome spontaneously condensates [Mondal *et al.* 2011] thus delineating a restricted sub-region of the cell called “nucleoid”, where molecular crowding is much larger than in the rest of the cytoplasm [Foley *et al.* 2010]. Macromolecular crowding then alternates along the cell long axis between low intensity zones (cytosol) and large intensity ones (nucleoid). Monte-Carlo dynamics modeling suggests that such non-homogeneous spatial distribution of the molecular crowding may be sufficient to localize large proteins to the cell poles [Saber *et al.* 2010]. In line with this proposal are the experimental reports that the observed aggregates preferentially localize in the nucleoid-free regions of the cell [Lindner *et al.* 2008; Rokney *et al.* 2009], i.e. precisely in the regions of alleged lower macromolecular crowding. In spite of these hints though, whether the transport

of the aging-related protein aggregates in *E. coli* is of a directed active nature or purely passive Brownian origin remains elusive, since contradictory results indicate that this process would include ATP-dependent stages [Rokney *et al.* 2009].

Here, our aim is to determine whether the movement of aging-related protein aggregates in *E. coli* is purely diffusive (Brownian) or includes some active process (ATP-dependent, directed transport or membrane tethering). To this aim, we devised an integrated approach combining time-lapse fluorescence microscopy of *E. coli* cells *in vivo*, innovative automated image analysis, and individual-based modeling. Our results strongly indicate that purely diffusive pattern of aggregates' mobility combined with nucleoid occlusion, underlie their accumulation in polar and mid-cell positions.

RESULTS

Trajectory analysis of single protein aggregates

In vivo analysis of individual trajectories of proteins of interest (or aggregates thereof) is a powerful method to determine whether the movement of the target protein is of Brownian nature or additionally exhibits further ingredients (active directed transport, caging or corralling effects, transient trapping, anomalous sub-diffusion) [Qian *et al.* 1991; Platani *et al.* 2002; Bronstein *et al.* 2009; Zhou *et al.* 2011; Bakshi *et al.* 2011; English *et al.* 2011; Sliusarenko *et al.* 2011]. Here, we focused on naturally forming protein aggregates tethered with the small heat-shock protein IbpA in *E. coli* whose spatio-temporal dynamics have been implicated in ageing of the bacteria [Lindner *et al.* 2008; Lindner *et al.* 2009].

To characterize the motion of IbpA aggregates in single *E. coli* cells, we monitored intracellular trajectories of single aggregates of IbpA-YFP fusion proteins [Allen *et al.* 1992; Lindner *et al.* 2008] in non-stressed conditions (37°C, in LB medium without streptomycin, see Materials and Methods). For the automatic quantification of the resulting time-lapse fluorescence microscopy movies, we developed dedicated image analysis and tracking software tools (see Materials and Methods). This software suite performs automatic segmentation and tracking of the cells (Figure 1A). Moreover, it automatically detects the fluorescence aggregates' foci and monitors their movements relative to the cell in which they are located in, with subpixel resolution.

Localization of protein aggregates is non-homogeneous along the cells

Detectable protein aggregates (in the form of localized fluorescence foci) were observed in only half of the cells monitored (54%), in agreement with previous experimental reports [Lindner *et al.* 2008; Landgraf *et al.* 2012]. Cells in non-stressed conditions tend to exhibit smaller copy numbers of distinct protein aggregates than in heat-shocked conditions

(compare e.g. [Lindner *et al.* 2008] with [Rokney *et al.* 2009; Winkler *et al.* 2010]). Accordingly, in our hands, nearly all the foci-containing cells (98.8%) displayed a single fluorescence focus within the imaging time, while the other ones showed at most two foci.

The distribution of the (initial) spatial localization of the aggregates is displayed in Figure 1. As a convention, we denote the long axis of the bacteria cell as its x -axis and the short one as its y -axis (Figure 1B). In this figure, we express the aggregate position relative to the cell center of mass, and rescale to $[-1, +1]^2$ range. Thus the two bacterial poles correspond to $x = -1$ and $x = +1$ in this relative scale.

The histogram of the location of the aggregate at the starting point of each trajectory is shown in Figure 1C. The distribution is highly non-homogeneous, with most of the aggregates predominantly localized at one of the two poles, and the others mainly around the middle of the cell. This distribution is similar to the results obtained in [Lindner *et al.* 2008] (Fig 2-E-F in [Lindner *et al.* 2008]), except that, since we do not differentiate between old and new poles, the amplitude of the polar modes in Figure 1C are roughly symmetrical. Similar distributions were also obtained in heat-shock conditions [Rokney *et al.* 2009; Winkler *et al.* 2010].

The marked localization of the aggregates suggests they might be tethered to the membrane at the poles (and center) at some point of the aggregation process, thus restricting their motion. Figure 1D shows the location of the polar aggregates at first detection (both poles were pooled). Because these experimental results are two-dimensional projections of three-dimensional positions, one cannot directly determine whether the aggregates are bound to the cell membrane or spread in the three dimensional cytoplasmic bulk. To this aim, we generated 10^4 aggregate positions (uniformly) at random in a volume of the same shape and dimensions than the cell pole. Fig. 1E shows the two-dimensional projection of these positions when the proteins were randomly located in the three-dimensional bulk whereas Fig. 1F shows the two-dimensional projections when the proteins were randomly located on the cytoplasmic membrane enclosing the bulk.

To quantify these plots, we analyzed the local density of protein positions in the two-dimensional projections. Assuming the 2d projection of the pole is a semi-ellipse of radii a_x and a_y (green dashed shapes in Fig. 1 D-F), its surface area is $1/2\pi a_x a_y$. The surface area of the elementary semi-elliptic disk D_s , delimited by the semi-ellipse of radii sa_x and sa_y ($0 < s < 1$) and that of radii $(s+ds)a_x$ and $(s+ds)a_y$ is thus $A(D_s) = \pi a_x a_y ds(s+ds/2)$. Varying s between 0 and 1, we counted the number of aggregate positions n_s found within the semi-elliptic disk D_s and computed the corresponding density $\rho(s) = n_s / A(D_s)$. Therefore, when s approaches 1, $\rho(s)$ represents the local density of aggregate positions close to the external boundary of the pole (green dashed shapes in Fig. 1 D-F) whereas for small s values, $\rho(s)$ gives the local

density of aggregate positions close to center of the pole (i.e. the center of the semi-ellipse). To validate the approach, we generated random aggregate positions directly in two dimensions (uniform distribution inside the 2D semi-ellipse) and checked that the density $\rho(s)$ exhibit a constant value with s (Fig. 1G, dashed black line).

When $\rho(s)$ was used to quantify the data of Fig. 1E (aggregates in the bulk), the density was roughly constant up to $s \sim 0.3$ then decayed smoothly for larger values (Fig. 1G, blue line), in agreement with the slight decay of the aggregate density close to the semi-ellipse boundary that is already visible in Fig. 1E. By contrast, with the data shown in Fig. 1F (aggregates in the membrane around the bulk), $\rho(s)$ increased first slowly with increasing s , then very abruptly close to 1 (Fig. 1G, red line), in agreement with the accumulation of aggregate locations close to the semi-ellipse boundary that is already visible in Fig. 1F. Therefore, these simulated data show that the distortions due to the projection in two dimensions are expected to manifest as strongly increasing $\rho(s)$ values close to $s=1$ if the aggregates are tethered to the membrane versus smoothly decreasing $\rho(s)$ if the aggregates are randomly spread in the 3d bulk. Finally, we used this approach to quantify the experimental data described above (Fig. 1H). The local density $\rho(s)$ exhibits a bimodal behavior. Close to the pole boundary, $\rho(s)$ clearly shows an almost linear decay, which is very similar to the behavior observed at large s for membrane-bound simulations s (Fig. 1F). For small s values however, $\rho(s)$ increases with s , thus indicating a phenomenon that hinders aggregate location close to the cell center. To conclude, these results unambiguously plead in favor of a 3d bulk distribution of the aggregates in the poles, thus rejecting the hypothesis that they would be tethered or attached at the cell membrane.

Polar protein aggregates exhibit Brownian motion

We next quantified the displacements of the polar aggregates. To this end we used two temporal regimes in our time-lapse fluorescence microscopy. Low sampling frequency (LF; 0.33 Hz) at long time scale (LT; 5 minutes) was used to assure coverage of the saturation of the measured trajectories ($n=1149$), whereas, high frequency (HF; 1.67 Hz) was used to increase the temporal resolution of the observed linear mean displacement regime observed in the former regime. These conditions were optimized considering the trade-off between extensive exposure leading to bleaching of the foci and satisfactory temporal resolution

The resulting mean displacements along the x or y -axis are shown in Supplementary Figure 1. For both sampling frequencies, the mean displacement along the y -axis is roughly stationary and fluctuates around a close-to-zero value. We therefore applied a minor correction and rescaled the mean displacement along the y -axis so that it fluctuates around exactly zero value.

The mean displacement along the x -axis displays a close-to-linear increase with time (Supplementary Figure 1) with a slope of around $0.05 \mu\text{m}/\text{min}$, a value that corresponds to the linear approximation of the cell elongation rate under our experimental conditions (doubling time around 30 minutes, during which the cell half-length grows from $\approx 1.0\text{-}2.0$ to $\approx 2.0\text{-}3.5 \mu\text{m}$). Therefore, the raw movement of the aggregates along the x -axis is dominated by ballistic displacement toward the pole under the effect of cell elongation. We corrected for this passive transport by subtracting the increase rate of the cell elongation.

The corrected mean displacements (Figure 2A) are stationary and slightly fluctuate around zero, for both HF and LF trajectories.

This is a typical characteristic of Brownian motion ('random walk'). To confirm this hypothesis one has to study higher-order moments of the displacement, in particular the second one.

The resulting mean-squared displacement $\langle (u(t) - u(0))^2 \rangle$ is displayed in Figure 2B. The LF and HF data here again are in very good agreement, with the HF data nicely aligned on the LF ones. This agreement is an important test of the coherence and quality of our measurement and analysis methodology. The inset of the figure shows a magnification of the HF data until time $t = 30$ sec. For the first 10 to 15 seconds, the HF data exhibits a clear linear behavior. As expected from an unbounded Brownian motion the same slope was observed for both the x - and y -axis. The non-zero intercept with the y -axis is typically due to the noise in the experimental determination of the aggregate position [Savin & Doyle 2005]. Such a linear dependence of the mean-squared displacement (MSD) is a further indication that the movement of the aggregates is Brownian diffusion, as one expects

$\langle (u(t) - u(0))^2 \rangle = 2D_u t$ in the case of a random walk (where $D_u = D_x$ or D_y the diffusion constant in the x - or y -direction, respectively). Using the first 15 seconds of the HF data, our estimates yield $D_x \approx 5.1 \times 10^{-4} \mu\text{m}^2/\text{s}$ and $D_y \approx 4.0 \times 10^{-4} \mu\text{m}^2/\text{s}$. Note that these values are at best rough estimates as the initial size of the aggregates averaged in these data (thus their expected mobility) has a large variation range (see above). Nevertheless, the fact that the values for the x - and y -axes are similar is another indication of the isotropy of the Brownian motion that seems to govern the movement of the aggregates. These values are compatible with previous experimental reports of the diffusion constants of large multi-protein assemblies in bacteria, such the origin of replication in *E. coli* (around $10^{-4} \mu\text{m}^2/\text{s}$ [Sliusarenko *et al.* 2001]) but are significantly smaller than the values reported for single fluorescent proteins such as mEos2 or GFP (1 to $10 \mu\text{m}^2/\text{s}$ [Elowitz *et al.* 1999; English *et al.* 2011]).

Altogether the analysis of the first 15 seconds of the HF data thus clearly pleads in favor of the hypothesis that the aggregates motion is due to simple passive Brownian motion, thus excluding directed transport due to some active mechanisms. After 15 seconds however, the

MSD from HF data show hints of saturation. As observed in the LF data, full saturation is reached after about 40 seconds. For both x - and y -data the MSD saturates at values 0.030 and 0.040 μm^2 . This corresponds to a restriction of the movements of the aggregates in a roughly volume subregion of the bacterial interior, with characteristic size ≈ 400 nm. To understand these results, one has to take into account the variability of the monitored aggregates.

Protein aggregates exhibit passive confinement within bacterial polar region

To study the saturation behavior of the aggregate movement at the longer time scale, we divided the LF data into 5 classes of increasing initial median fluorescence (table 1), so as to average data over aggregates of more homogeneous sizes. Class 1 groups together aggregates with low initial fluorescence (<1459), class 2 groups aggregates with larger initial fluorescence (ranging in $[1459,2015[$) etc. The MSD in the x - and y -direction are then averaged class by class. Figure 3A and B show the obtained values (full lines). These curves lead to two significant conclusions: (i) the initial slope of the curves increases when the average intensity in the class decreases. Assuming that the initial median fluorescence is proportional to the initial size of the aggregate, this result further supports a purely diffusive behavior, for which the diffusion constant is expected to decrease as the molecular size increases; (ii) the observed saturation levels are not constant but clearly decrease when the initial total fluorescence increase. This suggests that the size of the aggregates is of the same order than the cell dimension, so that the intracellular space available for an aggregate of size r is in fact $L-r \ll L$ (where L is the cell size). In turn, this enables us to estimate the actual size of the foci, independent of the microscope resolution.

To this end, we used the data in Figure 3A&B to obtain reliable estimates of the sizes of the aggregates and that of the intracellular subregion in which they are confined. We used an automatized procedure based on the optimization by an evolutionary strategy of the parameters of an individual-based model (see Materials and Methods for full description). In short, our strategy can be seen as an automatized fit of the values of 12 parameters: the dimensions L_x and $L_y = L_z$ of the intracellular subregion in which the aggregates are confined and the average radius r_i and diffusion constant D_i of the aggregates belonging to class i ($i = 1..5$). We start by setting these parameters to initial guess values. We then simulate the diffusion of 5,000 spherical molecules per class that are endowed with the corresponding values of r_i and D_i ($i = 1..5$) and diffuse in a spatial domain which size is given by the initial guess values of L_x and $L_y = L_z$. The values of the MSD (averaged over each class) in the x and y directions are then sampled during the simulation at the same frequency (0.33 Hz) as

in the LF data of Figure 3A&B. We then compute the distance between the resulting curves for the simulated data and those for the experimental LF data. This distance represents the error between the MSD predicted by the simulations using the initial values of the parameters. To minimize this error automatically, we used an evolutionary strategy called CMA-ES (see [Hansen *et al.* 2001] and Material and Methods). In short, CMA-ES automatically finds the values of the parameters that yield the smallest distance between experimental and simulation data. Note that because of the correction procedure (that removes the effect of cell elongation) for the x-data is strong (i.e. the aggregate motion in this direction is originally dominated by cell elongation), it is difficult to apply this fitting procedure on the experimental MSD data for the x-and y-axis simultaneously. This obliged us to fit the parameters separately on the x and y-data. The numerical values indicated in table 1 below represent the average of these two fits (the "±" values indicate the total variation range).

The class-averaged MSD corresponding to the optimized parameters are shown in Figure 3A&B as dashed lines. It is obvious from these data that optimization procedures found parameter values that lead to behaviors that are in very good agreement with the experimental data. The best-fit values for the cell dimension parameters are $L_x = 650$ nm and $L_y = L_z = 750$ nm. Considering our pixel size (64 nm), we conclude that the value of the space available to the aggregate in the y and z directions is close to the whole space available in this direction (around 900 nm). Therefore the aggregate motion in the y and z directions seem unrestrained except by the inner cell membrane. In strong contrast, while the total intracellular space in the x-direction ranges from 2 to 5 μ m (depending on the cell elongation), our fitting procedure indicates that the aggregates do not move beyond 650 nm in this direction. As they are initially located in the cell pole, this shows that the aggregates remain in the pole, where a passive mechanism hinders their free movement and keep them from leaving a subregion spanning $\approx 1/4$ th to $1/9$ th of the total cell length. This is coherent with the positioning of the nucleoids, hindering the diffusion of the aggregates (*see below*).

The experimental relation between the aggregate radius and diffusion constant, as determined by the fitting procedure above (listed in table 1). shows a good fit to a Stokes-Einstein relation: $D(r) = C_0 / r$ where one expects $C_0 = kT / (6\pi\eta)$ (describing simple sphere diffusion in a liquid of viscosity η) (Fig. 3C). We obtain a remarkable fit to the experimental data with $C_0 = 47230$ nm³/s. Note that a power-law fitting $D(r) = C_0 / r^b$ gives $b = 0.89 \pm 0.14$ (thus that does not exclude $b = 1.0$) with a similar quality of fit (chi-square values). In the inset of Figure 3C, the experimentally determined values of D are plotted as a function of $1/r$. The relation is linear except for very large values of r , thus emphasizing the very good agreement with the Stokes-Einstein law.

These results thus confirm that some passive mechanism hinders the free movement of the

aggregates along the long axis. They also show that the aggregates follow the Stokes-Einstein law even for large sizes since the diffusion constant of the aggregates simply decays as the inverse of its radius.

Table 1. Comparison of the estimated radius and diffusion constants of protein aggregates depending on their initial median fluorescence. Data are averages of the fits on the MSDx- and MSDy-data and +/- values locate min-max.

Intensity classes	# aggregates	r (nm)	D (nm²/s)
I < 1459	128	54 ± 2	837 ± 38
1459 < I < 2015	115	105 ± 11	550 ± 2
2015 < I < 2905	128	141 ± 33	437 ± 26
2905 < I < 4727	129	174 ± 10	261 ± 4
I > 4727	128	273 ± 12	92 ± 26

Simulation of aggregation, Brownian diffusion, and molecular crowding reconstitute

If the spatial dynamics of protein aggregates in *E coli* is indeed based exclusively on Brownian motion and aggregation, we should be able to reproduce by computer simulations the observed spatial patterns of the aggregates described here and in previous experimental reports [Lindner *et al.* 2008; Rokney *et al.* 2009; Winkler *et al.* 2010]. In particular, given the relative simplicity of these two ingredients (diffusion + aggregation), individual-based modeling is a very convenient tool in this framework. Here, we used an individual-based model in which proteins diffuse via lattice-free random walks in a 3D domain that has the size of a typical *E coli* cell (see Material and Methods and Figure 4A). When two molecules meet along their respective random walks, they form a single aggregate with probability p_{ag} . The size of the resulting aggregate is updated according to the size of its constituting molecules and its diffusion constant is updated as a function of the aggregate size using a simple Stokes-Einstein law. In turn, such aggregates can combine to form bigger aggregates.

The above experimental finding shows that the aggregate Brownian motion is restricted in the x-direction over a distance that is comprised between 1/4th to 1/9th of the cell long axis. These numbers are in agreement with the hypothesis that the physical structures that hinders aggregate diffusion are the nucleoids, i.e. the subregion of the bacterial cytoplasm where the chromosome condensates, thus increasing molecular crowding and diffusional hindrance [Foley *et al.* 2010]. In a typical *E coli* culture in exponential phase, most cells

contain two nucleoids [Donachie *et al.* 1989] and the free space between each nucleoid and the closest pole end is around 500-600 nm (deduced from Ref [Zimmerman 2006]). We thus tested the hypothesis (already pointed out in [Lindner *et al.* 2008] or [Winkler *et al.* 2010]) that the restriction to the free movement of the obstacles along the long axis of the bacteria is caused by an increased molecular crowding in the nucleoids. To simulate this increased molecular crowding in our individual-based model, we added immobile non-reactive obstacles to represent the densely packed nucleoids (see Materials and Methods).

Each simulation is initiated by positioning N_p single proteins at random inside the cells. The simulation then proceeds by moving the proteins via random walks and letting them aggregate (with resulting update of the aggregate radius and diffusion constant) when they encounter. The initial number of proteins N_p was varied in order to account for different experimental conditions. We used $N_p=100$ to emulate experiments in non-stressed conditions (as in this work) where one does not expect the presence of large numbers of proteins in the aggregates. For experiments where aggregation is intensified using e.g. heat shocks [Rokney *et al.* 2009; Winkler *et al.* 2010], one generally expects to recover more proteins in the aggregates. In such conditions, the number of proteins in the aggregates was evaluated to be between 2,000 and roughly 20,000 molecules [Winkler *et al.* 2010]. Here, to account for these data, we used at most a total of $N_p = 7,000$ molecules in the simulations.

In the experiments, protein aggregates are detected by fluorescence microscopy when their size is large enough; though we have currently no means to quantify this threshold size, we estimate our detection limit to contain about 10-50 YFP copies, based on the detection of slow mobile particles [Xie, 2010]. Therefore, in the simulations, we had to arbitrarily fix a threshold for the number of proteins per aggregate above which the aggregate is considered large enough to be detected. In non-stressed conditions ($N_p = 100$), we used a detection threshold of 30 proteins per aggregate, whereas with $N_p = 7,000$ total molecules per cell, we varied the detection threshold between 30 and 1750 monomeric proteins per aggregates.

We first focused on the distribution of the position of the aggregates at their first detection along the bacterial long axis. The corresponding distribution is shown in Figure 4B. In non-stressed conditions ($N_p = 100$), the distribution of the aggregate position is qualitatively very similar to our experimental results (see Figure 1C): most of the aggregates are located in the poles, the rest being mainly located in the cell center. The distribution of the aggregate location at first detection in our simulations in these conditions does not depend much on the detection threshold: although the full lines in the figure correspond to aggregate detection thresholds ranging from 5 to 50 proteins per aggregate, the resulting spatial distributions are similar and match all well with the experimental result (dashed line histogram). Importantly, in agreement with experimental reports, the nucleoids in the simulations are not

fully impermeable to the aggregates, since the probability to observe large aggregates in the nucleoids is not null. Rather, the presence of the largest molecular crowding in the nucleoids reduces the probability that large aggregates form in the nucleoids.

Another intriguing experimental result is that the number of aggregates simultaneously present in a cell depends on the experimental conditions. In non-stressed conditions only 1.2% of the observed cells presented more than a single aggregate within the observation time [Lindner *et al.* 2008] whereas in heat-shocked conditions a majority of the cells present two distinct aggregates, and the rest is roughly equally distributed between one and three protein aggregates per cell. These observations are faithfully reproduced by diffusion-aggregation mechanisms. Indeed, in our simulations emulating non-stressed conditions (100 proteins), a vast majority of the simulations develop a unique detectable aggregate within the simulation time (Fig. 4C). Despite the increase with time of the probability for a simulation to feature two detectable aggregates per cell, the probability to observe only one aggregate remains largely higher, even at the end of the simulations.

In simulated heat-shock conditions (see Figure 4D, total of 7,000 proteins) with a very low aggregate detection threshold, virtually all simulations display at least 4 distinct detectable protein aggregates per cell from the very beginning of the simulation. These results are comparable with the experimental data obtained in [Lindner *et al.* 2008], where the addition of streptomycin was shown to strongly increase the size and number of detected aggregates per cell (to 5 or more). When we increased the aggregate detection threshold (illustrated by a detection threshold of 1750 aggregates in Fig 4C3) the simulations showed very different behavior. For short simulation times, we mainly observed cells with a unique detected aggregate. As simulation time increases, the probability to detect a unique aggregate decays in favor of the probability of detect 2 and 3 aggregates simultaneously in the cells. Eventually, most of the simulations (around 60%) display two aggregates, and the others are roughly evenly shared between 1 and 3 detected aggregates per cell. These proportions are in a very good agreement with the experimental results reported in [Winkler *et al.* 2010] that employs heat-shock triggered aggregation.

These simulations therefore predict that the number of detected aggregates in the cell is crucially dependent on two main factors: the aggregate detection threshold and the total number of aggregation-prone proteins. Therefore they suggest that the discrepancy observed concerning the number of aggregates per cell between non-stressed and heat shock conditions is due to the larger quantity of aggregate-prone proteins resulting from the heat shock. Taken together, our results show that the three basic ingredients we considered in our simulations (passive Brownian motion, aggregation, increased molecular crowding in

the nucleoids) are sufficient to reproduce several experimental observations on the spatial distribution and number of protein aggregates in *E coli*. Therefore, they confirm the conclusion drawn from our experimental results above that the movement of the chaperone protein aggregates in *E coli* is driven by passive diffusion (Brownian motion). They moreover indicate that observed non-homogeneous spatial distribution is not due to active or directed aggregate movement but is a mere result of the interplay between Brownian diffusion and molecular crowding.

DISCUSSION

Our objective in this work was to decipher the mechanisms by which protein aggregates in *E. coli* localize to specific intercellular regions. Using single-particle tracking of protein aggregates marked with the small heat shock chaperone IbpA (inclusion-proteins Binding Protein A) translationally-fused to the yellow fluorescence protein (YFP), our results indicate that the aggregate movement is purely diffusive, with coefficient constants of the order of 500 nm²/s, depending on their size.

Applying evolutionary strategy to the parameters of individual-based simulations of confined diffusion, we used these experimental data to estimate the average size and diffusion constant of the aggregates and the distances over which their movement is confined. As expected, the aggregate diffusion constant decreases with increasing aggregate sizes, but, more surprisingly, we find that the relation between the aggregate diffusion constant and their size is in very good agreement with the Stokes-Einstein law, thus strengthening the demonstration of pure Brownian motion. The agreement with the Stokes-Einstein law, that predicts a decay of the diffusion constant as the inverse of the radius, $D \sim 1/r$, was found valid for all the estimated aggregate radii, even as large as 250-270 nm.

Recent experimental tests of the validity of this law in *E. coli* were more ambiguous. [Kumar *et al.* 2010] quantified the diffusion of a series of 30–250 kDa fusion proteins (some of which contained native cytoplasmic *E. coli* proteins) in *E. coli* cytoplasm and found very strong deviation from the Stokes-Einstein law -even for small proteins- with very sharp decay of the diffusion constant $D \sim 1/r^6$. However, using GFP multimers of increasing sizes, [Nenninger *et al.* 2010] found very good agreement with Stokes-Einstein law from 20 to 110 kDa, i.e. up to quadri-mers, while the diffusion constant for penta-mers (138 kDa) was found much smaller than Stokes-Einstein prediction. Moreover, deviations from the Stokes-Einstein law was suggested an indication of specific interactions of the diffusing protein with other cell components. A tentative interpretation of our observation that even large cytoplasmic protein aggregates in *E. coli* do follow Stokes-Einstein law, would be that these aggregates actually have very few interactions with other cell components. This hypothesis would match very well with the putative protective function of the aggregates as scavengers of harmful misfolded proteins, allowing their retention within large, stable objects [Lindner and Demarez, 2009].

A second major finding of our study is the demonstration that the Brownian motion of the

aggregates is restricted by the cell membrane in the section plane of the cell, while, along the cell long axis, the aggregates are confined to a region that roughly corresponds to the nucleoid-free space in the pole, thus confirming the importance of hindered diffusion in the nucleoids. In further support to this hypothesis, we used 3D individual-based modeling to show that these three ingredients are sufficient to explain the most salient experimental observations. Our simulations exhibit spatial distributions of the aggregates that are similar to those observed in non-stressed as well as heat-shock conditions. They also explain the differences in the number of distinct aggregates per cell as a mere difference in the total number of aggregation-prone (misfolded) proteins. Therefore, our results strongly support the hypothesis that the localization of aging-related protein aggregates in the center and poles of *E. coli* is due to the coupling of passive diffusion-aggregation with the spatially nonhomogeneous macromolecular crowding resulting from the localization of the nucleoid(s).

A recent study has shown that tagging of native proteins by fluorescent proteins via translational fusions can have deep impact on the localization and aggregation patterns [Landgraf *et al.*, 2012]. Some of the commonly used fluorescent proteins were found to cause severe mislocalization and greatly exaggerate protein aggregation when fused to native *E. coli* proteins, presumably through fluorescent protein-mediated clustering. We do not think that our results suffer from this experimental bias. Indeed the localization and aggregation pattern for the lbpA-YFP fusion reported in [Landgraf *et al.*, 2012] are similar to those reported here, with protein aggregates mostly at poles and mid-cell position. Moreover, our earlier work [Lindner *et al.*, 2008] has shown that the localization of lbpA-YFP aggregates on fluorescence images overlaps with the localization of the inclusion bodies on phase contrast images, thus excluding mislocalization of lbpA due to tagging by YFP.

As a whole, our results uncover the importance for diffusion-based protein localization of the "soft" intracellular structuration of *E. coli* along the large axis due to increased macromolecular crowding in the nucleoids. In addition to this implication in the localization of aging-related protein aggregates, the structuring effect of the nucleoids has very recently been evidenced in the accurate and robust positioning of the divisome proteins (that mediate bacterial cytokinesis) [Männik *et al.*, 2012] or the nonhomogeneous spatial distribution of the transcription factor LacI [Kuhlman and Cox, 2012]. . The polar localization pattern of low diffusive elements in bacteria is not limited to protein or protein aggregates. Our recent work demonstrates that large engineered RNA assemblies accumulate as well in the cells' poles [Delebecque *et al.*, 2011, (electron microscope images therein)]. More intriguingly, the two poles of *E. coli* cells might be unequally attractive for the accumulation of damages. A recent article [Lloyd-Price *et al.*, 2012] shows that once produced, RNA-MS2-GFP complexes are

asymmetrically partitioned between the two poles.. While the localization of these macromolecules to the cell pole could be due to the same nucleoid structuring as the phenomena reported above, the mechanism under this bias toward one pole over the other remains to be elucidated.

MATERIALS AND METHODS

Bacterial strains.

The sequenced wild-type *E. coli* strain, MG1655 [Blattner *et al.* 1997] was modified to express an improved version of the YFP fluorescent protein fused to the C terminus of IbpA [Allen *et al.* 1992] under the control of the endogenous chromosomal *ibpA* promoter resulting in the MGAY strain. *E. coli* strains were grown in Luria-Bertani (LB) broth medium half salt at 37°C. For more information about the cloning, see [Lindner *et al.* 2008] *S.I.*

Fluorescence time-lapse microscopy setup.

After an overnight growth at 37°C, MGAY cultures were diluted 200 times. When the cells reached an absorbance 0.2 (600 nm), they were placed on microscope slide that was layered with a 1.5% agarose pad containing LB half salt medium. The agarose pad was covered with a cover-slide, the border of which was then sealed with nail polish oil.

Cells were let to recover for 1 hour before observation using Nikon automated microscope (ECLIPSE Ti, Nikon INTENSILIGHT C-HGFIE, 100x objective) and the Metamorph software (Molecular Devices, Roper Scientific), at 37°C. Phase contrast and fluorescence images (25% lamp energy, 1 second illumination LF movies and 600 milliseconds for HF movies) were sampled at two different time-scales. For low-frequency (LF) movies, images were taken every 3 seconds for a total of 5 minutes, while for high-frequency (HF) movies, fluorescence images were taken about every 0.60 seconds for a total of 2 min (and phase contrast images were sampled about every 7 fluorescence images).

Image analysis and aggregate tracking.

Phase contrast images were analyzed by customized software “Cellst” [Primet *et al.* 2008; M. Primet, PhD manuscript November 25th 2011] for cell segmentation and single cell lineage reconstruction. Phase contrast images have been denoised using the flatten background filter of Metamorph software for long movies or a mixed denoising algorithm [Louchet & Moisan 2011] for fast movies. The mixed denoising algorithm combines two famous image denoising methods: NL-means denoising [Rudin *et al.* 1992], which is patch-based and Total Variation denoising [Buades *et al.* 2005], which is used as regularization. The Cellst software was used to automatically segment the cells on most of the images, albeit when necessary, manual corrections were applied. At the end of the whole segmentation & tracking process, Cellst also calculates the lineage of every cell in the movie.

The fluorescent protein aggregates were detected by another customized software (Jacob, Dimiccoli, Moisan 2012). Detection of the maximal intensity pixel of each spot was realized using the a contrario methodology based on a circular patch model with a central zone of detection and an external zone of context. The patch radius is then optimized to optimally match the spot. The energy of each spot is then computed in the following way: the total image is modeled as a sum of a constant background and 2D circular Gaussian curves, centered on the maximal intensity pixel of the detected spots with a deviation of 3 pixels. The quadratic minimum deviation between the image and the model enables to calculate the Gaussian coefficients. These coefficients are considered the energy values of each spot. The coordinates of the detected spots are then refined to subpixel resolution. This is achieved by computing a weighted average of the coordinates of the pixels in a circular neighborhood of the detected spot. The weights are given by the intensity values of the pixels to which the local background is subtracted. The local background is computed as the median value of the pixel intensity in the neighborhood. Only pixels having an intensity bigger than the median value are considered in the weighted average. This algorithm has been tested on both synthetic and real image and it has been shown a precision of 1/10 of a pixel on very poor contrasted spots.

After detection and localization, the movements of the fluorescent aggregates were tracked and quantified by a third customized software named "aggtracker" (Primet, Moisan 2011) based on the cell lineage and the detected spots. The algorithm uses the lineage and cell information to ensure that an aggregate is consistently tracked through points with points that are inside the same cell.

The output of this software are time-series for the coordinates x and y (in pixels) of each fluorescence spot as well as the affiliation of the spot to the cell it is in. The last step consisted in the projection of the coordinates of the fluorescence spot from their initial absolute values in the image (in pixels) to their value along the long and short axes of the 2d image of the cell. To this aim, we used active skeletons (Jacob, Moisan 2011). A skeleton represents an object by a median line (the center line in the case of a tubular bacteria). Here we used one active skeleton for the long axis of the cell image and a second one for its short axis. Active skeletons are adapted to bacteria in order to optimize the position of the skeleton in the image of the cell. The coordinates of the fluorescence spots are then expressed as the coordinate of the center of the fluorescence spot in the basis composed of the two active skeletons that localize the cell long and short axes. We exploited the simple shape of the skeleton to estimate the total cell width and length as that of the respective skeleton. As a convention, we refer below to the aggregate coordinate along the long axis as the x -

coordinate and that along the short axis as the y-coordinate.

In order to improve precision, aggregate trajectories made of less than 10 successive images in the movies were not further taken into account. In total, we obtained 1644 aggregate trajectories.

Individual-based modeling of YFP aggregation.

To simulate the diffusion and aggregation process of proteins in a single cell, we used a 3d individual-based lattice-free model. Each protein p was explicitly modeled as a sphere of radius r_p centered at coordinates (x_p, y_p, z_p) in the 3d intracellular space of the cell. We simulated protein diffusion in the cell and aggregation as they encounter using as realistic conditions as possible. In particular, the radius and diffusion coefficient of the protein aggregates explicitly increased as they grow. Moreover, we explicitly modeled the larger molecular crowding in the nucleoids. Details of the simulations are as follows.

The bacterial cell was simulated as a 3D square cylinder with width and depth 1.0 μm [Volkmer *et al.* 2011] and length 4.0 μm (chosen to correspond to a bacterial cell just before division) and reflective boundaries. Note that we also ran simulations with more realistic cell shapes (i.e. spherical caps at cell ends) and did not find significant differences compared to square cylinders (except for the much higher computation cost with spherical caps).

Within each cells, we also explicitly modeled the larger molecular crowding found in the nucleoids. Indeed, in healthy cells, the bacterial chromosome condensates into a restricted sub-region of the cell called "nucleoid", where molecular crowding is much larger than in the rest of the cytoplasm [Foley *et al.* 2010]. To model this increased molecular crowding in the nucleoids, we placed at random (with uniform probability) 50,000 bulky immobile, impenetrable and unreactive obstacles (radius 10 nm) in the region of the cell where a nucleoid is expected. Because cell cycle and DNA replication in *E. coli* are not synchronized, roughly 75% of the cells in exponential phase contain two nucleoids [Donachie *et al.* 1989]. We thus explicitly positioned two nucleoids within the cell. The location and size of the two nucleoids were estimated from DAPI-stained inverted phase contrast images of the nucleoids found in Ref [Zimmerman 2006]. Both nucleoids were 3d square cylinders of length 1220 nm (along the cell long axis) and width and height 532 nm. Each nucleoid started at 540 nm from each cell pole and was centered on the cell long axis. The volume occupied by the two nuclei area thus formed is about 12%, which is consistent with literature [Woldringh 1999; Odijk 1998].

Each simulation was initialized by positioning N individual IbpA-YFP proteins (monomers) at non-overlapping randomly chosen (with uniform probability) locations in the free intracellular space of the cell (i.e. the whole interior of the cell minus the space occupied by the obstacles in the nucleoids). At each time step, each molecule is independently allowed to diffuse over a

distance d that depends on the protein diffusion constant D_p , according to $d = (6 D_p \Delta t)^{1/2}$, where Δt is the time step, in agreement with basic Brownian motion. Note that D_p itself depends on the aggregate size r_p (see below). The new position of the protein (x', y', z') was then computed by drawing two random real numbers, θ and c , uniformly distributed in $[0, 2\pi]$ and $[-1, 1]$, respectively, and spherical coordinates: $x' = x(t) + d \sin(\text{acos}(c)) \cos(\theta)$; $y' = y(t) + d \sin(\text{acos}(c)) \sin(\theta)$ and $z' = z(t) + d c$ where $(x(t), y(t), z(t))$ is the initial position of the protein. If the protein in this new position (x', y', z') overlaps with any of the immobile obstacles (i.e. if there exists at least one obstacle such that the distance between the obstacle center and (x', y', z') is smaller than the sum of their radii) the attempted movement is rejected $(x(t+\Delta t), y(t+\Delta t), z(t+\Delta t)) = (x(t), y(t), z(t))$. This classical approximation of the aggregate reflection by the static obstacles is not expected to change the simulation results significantly, but it drastically reduces the computation load. If no obstacle overlaps, the movement is accepted, i.e. $(x(t+\Delta t), y(t+\Delta t), z(t+\Delta t)) = (x', y', z')$.

After each molecule has moved once, the algorithm searches for overlaps between proteins. Two proteins are overlapping whenever the distance between their centers is smaller than the sum of their radii. Each overlapping pair is allowed to aggregate with (uniform) probability p_{ag} (irrespective of their size). In our simulations, p_{ag} was varied between 0.1 and 1.0. To model the aggregation from two overlapping proteins, we could not, for computation time reasons, keep track of the shape of the aggregates (i.e. the individual location of each protein in the aggregates). Instead, we used the simplifying hypothesis that all along the simulation, the aggregates maintain a spherical shape with constant internal density. It follows that the radius of an aggregate C, born out of the aggregation of two aggregates A and B of respective size r_A and r_B is $r_C = (r_A^3 + r_B^3)^{1/3}$. Upon aggregation, we thus remove the aggregates A and B from the cell, and add a new aggregate with size r_C , centered at the center of mass of the two former aggregates A and B.

Finally, to set the diffusion constant of the aggregates, we used the classical Stokes-Einstein relation for a Newtonian fluid, where the diffusion constant is inversely proportional to its radius. In our case, this leads to $D_p = D_0 r_0 / r_p$ where r_0 and D_0 are the radius and diffusion constant, respectively, of individual (monomeric) IbpA-YFP molecules (see below). Note that this relation seems to be violated for large molecules in the cytoplasm of *E. coli* [Nenninger *et al.* 2010; Kumar *et al.* 2010]. In a subset of simulations, we used power law relations, such as $D_p \propto r_p^{-6}$, as suggested in [Kumar *et al.* 2010], without noticeable change in our results (except that the time needed to reach a given threshold aggregate size was increased). Note that aggregation was considered irreversible in our model (i.e. aggregates do never breakdown into smaller pieces). This is in agreement with our experimental observations, where we never measured decay of foci fluorescence.

The diffusion constant of the 26-kDa GFP (radius 2 nm) in *E. coli* cytoplasm is around $7.0 \mu\text{m}^2/\text{s}$ and that of the GFP-MBP fusion (72 kDa) around $2.5 \mu\text{m}^2/\text{s}$ [1]. Using this data and the Stokes-Einstein relation combined to our constant spherical hypothesis, led to estimates of the radius and diffusion coefficient of the individual (monomeric) 39 kDa IbpA-YFP fusion of $r_0 = 3 \text{ nm}$ and $D_0 = 4.4 \mu\text{m}^2/\text{s}$.

The value of the time step Δt has to be small enough so that proteins cannot jump over each other during a single time step, meaning that the distance diffused during a single time step is limited by $d < 4r_0$. Using the definition for d above, one then has $\Delta t < 8/3 r_0^2/D_0 \approx 5 \mu\text{s}$. Here, we used $\Delta t = 1 \mu\text{s}$ yielding $d = 5 \text{ nm}$ for monomeric proteins.

Every simulation was run for a total of 2×10^6 time steps. The translation of this value into real time is hardly possible since we have no indication of the experimental value of the aggregation probability per encounter p_{ag} (see above) even less so of its dependence on the aggregate size. A lower bound can be estimated to 2 seconds real time (for 2×10^6 time steps) if the aggregation is always diffusion limited (i.e. $p_{ag} = 1$). On general grounds however, the experimental value of p_{ag} can be expected to be smaller, so that the 2×10^6 simulation time steps would correspond to more than this 2 seconds real time minimal value. For the results to be statistically significant, we ran n_{run} simulations for each parameter and condition, with different realization of the random processes (initial location, random choice of the positions or of the aggregation events) and averaged the results over these n_{run} simulations. In the results presented here we used $n_{run} = 10^3$.

Fitting procedure for the aggregate radius, diffusion constant and cell dimensions

The data from the LF movies were partitioned into 5 classes based on the aggregate fluorescence intensity at the beginning of the measured trajectory, leading 5 pairs of experimental curves for the mean-squared displacement, $\langle (x_i^{exp}(t) - x_i^{exp}(0))^2 \rangle = f_x(t)$ and $\langle (y_i^{exp}(t) - y_i^{exp}(0))^2 \rangle = f_y(t)$ where $i = \{1, \dots, 5\}$ is the intensity class. Corresponding theoretical values were obtained by individual-based simulations comparable as those described above but modified as follows: the cells, of dimensions LX (length), $LY = LZ = LYZ$ (height and width) were devoid of nucleoids or aggregation (aggregation probability $p_{ag} = 0$) and we used $N = 5,000$ IbpA-YFP proteins. Each 12-uplet of parameters $\{LX, LYZ, r_i, D_i\}$ yields two theoretical curves $\langle (x_i^{the}(t) - x_i^{the}(0))^2 \rangle = g_x(t)$ and $\langle (y_i^{the}(t) - y_i^{the}(0))^2 \rangle = f_y(t)$. The aim of the fitting procedure is to minimize the distance between the experimental and theoretical curves, in a least-square sense, ie to minimize the cost function:

$$F = \sum_{i=1}^5 \sum_{j=1}^N \left(\frac{\langle (x_i^{exp}(t) - x_i^{exp}(0))^2 \rangle - \langle (x_i^{the}(t) - x_i^{the}(0))^2 \rangle}{\langle (x_i^{the}(t) - x_i^{the}(0))^2 \rangle} \right)^2 \dots$$

$$+ \left(\frac{\langle (y_i^{exp}(t) - y_i^{exp}(0))^2 \rangle - \langle (y_i^{the}(t) - y_i^{the}(0))^2 \rangle}{\langle (y_i^{the}(t) - y_i^{the}(0))^2 \rangle} \right)^2$$

where the indices j are over the N time steps. To minimize automatically the cost function F , thus adjusting the theoretical to the experimental curves, we used the C++ implementation of the state-of-the-art evolutionary strategy algorithm CMA-ES [Hansen *et al.* 2001] with population size 12 and 400 generations.

ACKNOWLEDGEMENTS

We thank N. Hansen, for providing the source code of CMA-ES for various programming languages (downloadable at http://www.lri.fr/~hansen/cmaes_inmatlab.html). We also thank the CNRS-IN2P3 Computing Center (cc.in2p3.fr) for providing computer resources.

REFERENCES

1. Ackermann M. *et al.*, *Senescence in a bacterium with asymmetric division*. *Science*. 2003. 300: p. 1920.
2. Aguilaniu H. *et al.*, *Asymmetric inheritance of oxidatively damaged proteins during cytokinesis*. *Science*. 2003. 299: p. 1751–1753.
3. Allen S.P. *et al.*, *Two novel heat shock genes encoding proteins produced in response to heterologous protein expression in Escherichia coli*. *Journal of Bacteriology*. 1992. 174: p. 6938-6947.
4. Bakshi S. *et al.*, *Subdiffraction-limit study of Kaede diffusion and spatial distribution in live Escherichia coli*. *Biophysical Journal*. 2011. 101: p. 2535-2544.
5. Blattner F.R. *et al.*, *The complete genome sequence of Escherichia coli K-12*. *Science*. 1997. 277: p. 1453-1474.
6. Bowman G. R. *et al.*, *A polymeric anchors the chromosomal origin/ParB complex at bacterial cell pole*. *Cell*. 2008. 134(6): p. 945-955.
7. Bronstein I. *et al.*, *Transient Anomalous Diffusion of Telomeres in the Nucleus of Mammalian Cells*. *Physical Review Letters*. 2009. 103(1): e018102.
8. Buades A. *et al.*, *A review of image denoising algorithms, with a new one*. *SIAM Multiscale Modeling and Simulation*. 2005. 4(2): p. 490-530.
9. Delebecque C.J. *et al.*, *Organization of intracellular reactions with rationally designed RNA assemblies*. *Science*. 2011. 333(6041): p. 470-474.
10. Domínguez-Escobar J. *et al.*, *Processive movement of MreB-Associated cell wall biosynthetic complexes in bacteria*. *Science*. 2011. 333 (6039): p. 225-228.
11. Donachie W.D. *et al.*, *Chromosome partition in Escherichia coli requires*

- postreplication protein synthesis*. Journal of bacteriology. 1989. 171: p. 5405-5409.
12. Elowitz M.B. *et al.*, *Protein mobility in the cytoplasm of Escherichia coli*. Journal of Bacteriology. 1999. 181(1): p. 197-203.
 13. English B.P. *et al.*, *Single-molecule investigations of the stringent response machinery in living bacterial cells*. 2011. Proceedings of the National Academy of Sciences of the United States of America. 108: p 365-373.
 14. Erjavec N. *et al.*, *Selective benefits of damage partitioning in unicellular systems and its effects on aging*. Proceedings of the National Academy of Sciences of the United States of America. 2008. 105(48): p. 18764-18769.
 15. Foley P.L. *et al.*, *Macromolecular crowding can account for RNase-sensitive constraint of bacterial nucleoid structure*. 2010. Biochemical and Biophysical Research Communication 345: p .42-47.
 16. Gehlen L.R. *et al.*, *Nuclear geometry and rapid mitosis ensure asymmetric episome segregation in yeast*. Current Biology 2011. 21 : p. 25–33.
 17. Hansen N. *et al.*, *Completely derandomized self-adaptation in evolution strategies*. 2001. Evolutionary Computation. 9: p. 159-195.
 18. Iniesta A.A. *et al.*, *Cell pole-specific activation of a critical bacterial cell cycle kinase*. Proceedings of the National Academy of Sciences of the United States of America. 107(15): p. 7012-7017.
 19. Jazwinski S.M., *Growing old: Metabolic control and yeast aging*. Annual Review Microbiology. 2002. 56: p. 769–792.
 20. Kaganovich D. *et al.*, *Misfolded proteins partition between two distinct quality control compartments*. Nature. 2008. 454 : p. 1088–1095.
 21. Kruse K. *et al.*, *An experimentalist's guide to computational modelling of the Min system*. 2007. Molecular Microbiology. 63:1p. 279-128.
 22. Kuhlman, T. & Cox, E. *Gene location and DNA density determine transcription factor*

distributions in Escherichia coli. Molecular Systems Biology, 2012. 8:610

23. Kumar M. *et al.*, *Mobility of cytoplasmic membrane, and DNA-binding proteins in Escherichia coli*. Biophysical Journal. 2010. 98: p. 552-559.
24. Landgraf D. *et al.*, *Segregation of molecules at cell division reveals native protein localization*. Nature Methods. 2012. 9: p. 480-485.
25. Lele U.N. *et al.*, *Phenotypic plasticity and effects of selection on cell division symmetry in Escherichia coli*. PLoS ONE, Public Library of Science. 2011. 6: e. 14516.
26. Lindner A.B. *et al.*, *Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation*. Proceedings of the National Academy of Sciences of the United States of America. 2008. 105(8): p. 3076-3081.
27. Lindner A.B. & Demarez A., *Protein aggregation as a paradigm of aging*. Biochimica. Biophysica. Acta. 2009. 1790(10): p. 980-96.
28. Lloyd-Price J. *et al.*, *Asymmetric disposal of individual protein aggregates in escherichia coli, one aggregate at a time*. Journal of Bacteriology. 2012. 194(7): p. 1747– 52.
29. Louchet C. & Moisan L., *Total Variation as a local filter*. SIAM Journal on Imaging Sciences. 2011. 4 (2): p. 651-694.
30. Männik J. *et al.*, *Robustness and accuracy of cell division in Escherichia coli in diverse cell shapes*. Proceedings of the National Academy of Sciences of the United States of America. 2012. 109(18): p. 6957-6962.
31. Mondal J. *et al.*, *Entropy-based mechanism of ribosome-nucleoid segregation in E. coli cells*. Biophysical Journal. 2011. 100: p. 2605-2613.
32. Nenninger A. *et al.*, *Size dependence of protein diffusion in the cytoplasm of Escherichia coli*. Journal of bacteriology. 2010. 192(18): p. 4535-4540.
33. Odijk T., *Osmotic compaction of supercoiled dna into a bacterial nucleoid*. Biophysical Chemistry. 1998. 73: p. 23-29.

34. Platani M. *et al.*, *Cajal body dynamics and association with chromatin are ATP-dependent*. Nature Cell Biology. 2002. 4: p. 502-508.
35. Primet M. *et al.*, *Tracking of cells in a sequence of images using a low-dimension image representation*. 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI). 2008. p 995-998.
36. Qian H. *et al.*, *Single particle tracking. Analysis of diffusion and flow in two-dimensional systems*. Biophysical Journal. 1991. 60: p. 910-921.
37. Rokney A. *et al.*, *E. coli Transports Aggregated Proteins to the Poles by a Specific and Energy-Dependent Process*. Journal of Molecular Biology, 2009. 392(3): p. 589-601.
38. Rudin L. *et al.*, *Nonlinear total variation based noise removal algorithms*, Physica D. 1992. 60: p. 259-268.
39. Rujano M.A. *et al.*, *Polarised asymmetric inheritance of accumulated protein damage in higher eukaryotes*. PloS Biology. 2006. 4(12): e417.
40. Saberi S. *et al.*, *Chromosome driven spatial patterning of proteins in bacteria*. PLOS Computational Biology. 2010. 6(11): e10000986.
41. Savin T. & Doyle P.S., *Static and dynamic errors in particle tracking microrheology*. Biophysical Journal. 2005. 88: p. 623-638.
42. Shapiro L. *et al.*, *Why and How Bacteria Localize Proteins*. Science. 2009. 326(5957): p. 1225-1228.
43. Shcheprova Z. *et al.*, *A mechanism for asymmetric segregation of age during yeast budding*. Nature. 2008. 454: p. 728–734.
44. Sliusarenko O. *et al.*, *High-throughput, subpixel precision analysis of bacterial morphogenesis and intracellular spatio-temporal dynamics*. Molecular Microbiology. 2011. 80(3): p. 612-627.

45. Stewart E.J. *et al.*, *Aging and death in an organism that reproduces by morphologically symmetric division*. Plos Biology. 2005. 3(2): p. 295-300.
46. Sun M. *et al.*, *Motor-driven intracellular transport powers bacterial gliding motility*. Proceedings of the National Academy of Sciences of the United States of America. 2011. 108(18): p. 7559-7564.
47. Turing A.M., *The chemical basis of morphogenesis*. Philosophical Transactions of The Royal Society. Lond. B. 1952. 237: p. 37–72.
48. Veening K.W. *et al.*, *Bistability, epigenetics, and bet-hedging in bacteria*. Annual review of Microbiology. 2008. 62: p.193-210.
49. Volkmer B. *et al.*, *Condition-dependent cell volume and concentration of Escherichia coli to facilitate data conversion for systems biology modeling*. PloS ONE. 2011. 6(7): e23126. doi: 10.1371/journal.pone.0023126.
50. Wang H. *et al.*, *Self-organized periodicity of protein clusters in growing bacteria*. Physical Review Letter. 2008. 101(21): p218101.
51. Watve M. *et al.*, *Aging may be a conditional strategic choice and not an inevitable outcome for bacteria*. Proceedings of the National Academy of Sciences of the United States of America, 2006. 103: p. 14831-14835.
52. Winkler J. *et al.*, *Quantitative and spatio-temporal features of protein aggregation in Escherichia coli and consequences on protein quality control and cellular ageing*. EMBO Journal. 2010. 29(5): p. 910-923.
53. Woldringh C.L. *et al.*, *Structure of dna within the bacterial cell: physics and physiology*. Charlesbos R, ed. Organization for the prokaryotic genome, American Society of Microbiolgy Press. 1999. 171 p.
54. Xie X. Sunnay, *Enzymology and life and the single molecule level*. Single Molecule Spectroscopy in Chemistry, Physics and Biology Nobel Symposium. Edited by Astrid Graslund, *et al.* p. 435-448, Springer Series in Chemical Physics, Springer-Verlag Berlin Heidelberg (2010).

55. Zimmerman S.B. *et al.*, *Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of Escherichia coli*. Journal Molecular Biology. 1991. 222: p. 599-620.
56. Zimmerman S.B., *Shape and compaction of Escherichia coli nucleoids*. Journal of Structural biology. 2006. 156: p. 255-261.
57. Zhou C. *et al.*, *Motility and segregation of Hsp104-associated protein aggregates in budding yeast*. Cell. 2011. 147: p. 1186-1196.

FIGURE LEGENDS

Figure 1: Localization of the detected aggregates in the cells

(A) In each image on the time-lapse fluorescence movies, the bacterial cells are automatically isolated (each individual cell is given a unique random color). The aggregates appearing during the movie are automatically detected and their trajectory within the cell quantified (internal trajectories). (B) By convention, we referred to the projection of the aggregate location on the long axis of the cell as the x -component and that along the short axis as the y -component. (C) Histogram of the x -component of the initial position of the trajectories. Since the cell length at the start of the trajectory is highly variable, the x -component was rescaled by division by the cell half-length. Thus computed, the cell poles are located at locations -1.0 and 1.0 , respectively, for every trajectory. Total of 1644 trajectories. (D) Experimentally measured positions of the aggregates detected in the poles (both poles pooled, $n = 9,242$ points). The green-dashed curves in (D-F) locate the 2d projection of the 3d semi-ellipsoid that was used to approximate the cell pole. (E) Synthetic data for bulk positions: 10,000 3d positions were drawn uniformly at random in the 3d semi-ellipsoid pole. The figure shows the corresponding 2d projections. (F) Synthetic data of membranary positions: 10,000 3d positions were drawn uniformly at random in the external boundary (membrane) of the 3d semi-ellipsoid pole. The figure shows the corresponding 2d projections. (G-H) Local density of aggregate positions $\rho(s)$ in the synthetic (G) and experimental (H) data shown in E (bulk, blue), F (membranary, red) and D (experimental, orange). The dashed black line shows the local density computed for 10,000 synthetic 2d positions that were drawn uniformly at random in the 2d semi-ellipse resulting from the 2d projection of the 3d pole ellipsoid (green dashed curve in D-F). See text for definitions of the local density $\rho(s)$ and length scale s .

Figure 2: Single-aggregate tracking analysis inside *E. coli* cells.

Coordinates along the x and y -axis are shown in red and black, respectively. Low frequency sampling trajectories (LF) are displayed using full lines and high frequency ones (HF) using open symbols. (A) Corrected mean displacement $\langle u(t) - u(0) - u_c(t) \rangle$ where u_c is the applied correction. For the y -component, the correction is the time-average of the y -coordinate: $u_y = \bar{y}$ where the bar denotes time-averaging of the y -component along the trajectory. For the x -component, the applied correction is the time-average of the y -coordinate: $u_x = [L(t) - L(0)]\Delta t$ where $L(t)$ is the cell half-length at time t and Δt is the time interval between two consecutive images. (B) Corresponding mean squared displacements

$MSD_c = \langle (u(t) - u(0) - u_c(t))^2 \rangle$. The inset shows a magnification of the HF results and their close-to-linear behavior for the first 10 - 15 seconds.

Figure3 : Trajectory analysis with more homogeneous aggregate-size classes.

Trajectories from the LF movies (Figure 2) were clustered into 5 classes of increasing initial median fluorescence (Table 1) and the corresponding were MSD averaged in each class. Full lines show the MSD for the x - (A) and y -axes (B) for each class. The dashed lines show the results of the fitting procedure for each class (see text and Material and Methods). The resulting values of the diffusion constant D are plotted against the radius r in (C). Full circles indicate the values determined from fitting the MSD in the x -direction, while full squares show the values from the fit in the y -direction.. The full line is a fit to a Stock-Einstein law with size-limitation $D(r) = C_0 / r$, yielding $C_0 = 47.23 \times 10^3 \text{ nm}^3/\text{s}$. The inset replots these data as a function of $1/r$.

Figure 4 : Individual-based models of chaperone protein diffusion and aggregation (A)

Geometry of the 3D model used in individual-based models for *E coli* intracellular space. Numbers indicate distances in μm . The blue boxes inside the bacteria locate the nucleoids, where increased molecular crowding is modeled by the insertion of bulk immobile obstacles. (B) Comparison between simulations and experiments of the localization at first detection of the protein aggregates along the long axis (x -axis). The full lines show the spatial distributions extracted from the simulations with different detection thresholds (an aggregate must contain at least 5, 10, 20 or 50 monomeric proteins, respectively, to be detected). Total number of proteins in the simulations $N_p = 100$. The dashed line is an histogram showing the distribution of the experimental data. (C-E) Time-evolution of the probabilities to observe exactly 1 (red), 2 (green), 3 (blue) or 4 (black) distinct aggregates simultaneously in the simulations. The simulations emulated non-stressed conditions (C), with $N_p=100$ total proteins and aggregate detection threshold = 30 or heat-shock triggered aggregation, with $N_p=7,000$ total proteins and aggregate detection threshold = 30 (D) or 1750 (E).

Supplementary figure 1: Mean displacements of single-aggregates.

The figure shows the time evolution of the mean displacement, $\langle u(t) - u(0) \rangle$, where $u = x$ or y (brackets denote averaging over the trajectories). Coordinates along the x and y -axis are shown in red and black, respectively. Low frequency sampling trajectories (LF) are displayed using full lines and high frequency ones (HF) using open symbols. The inset schematizes the increase of the cell half-length during growth that dominates the movement along the x -axis.

FIGURES

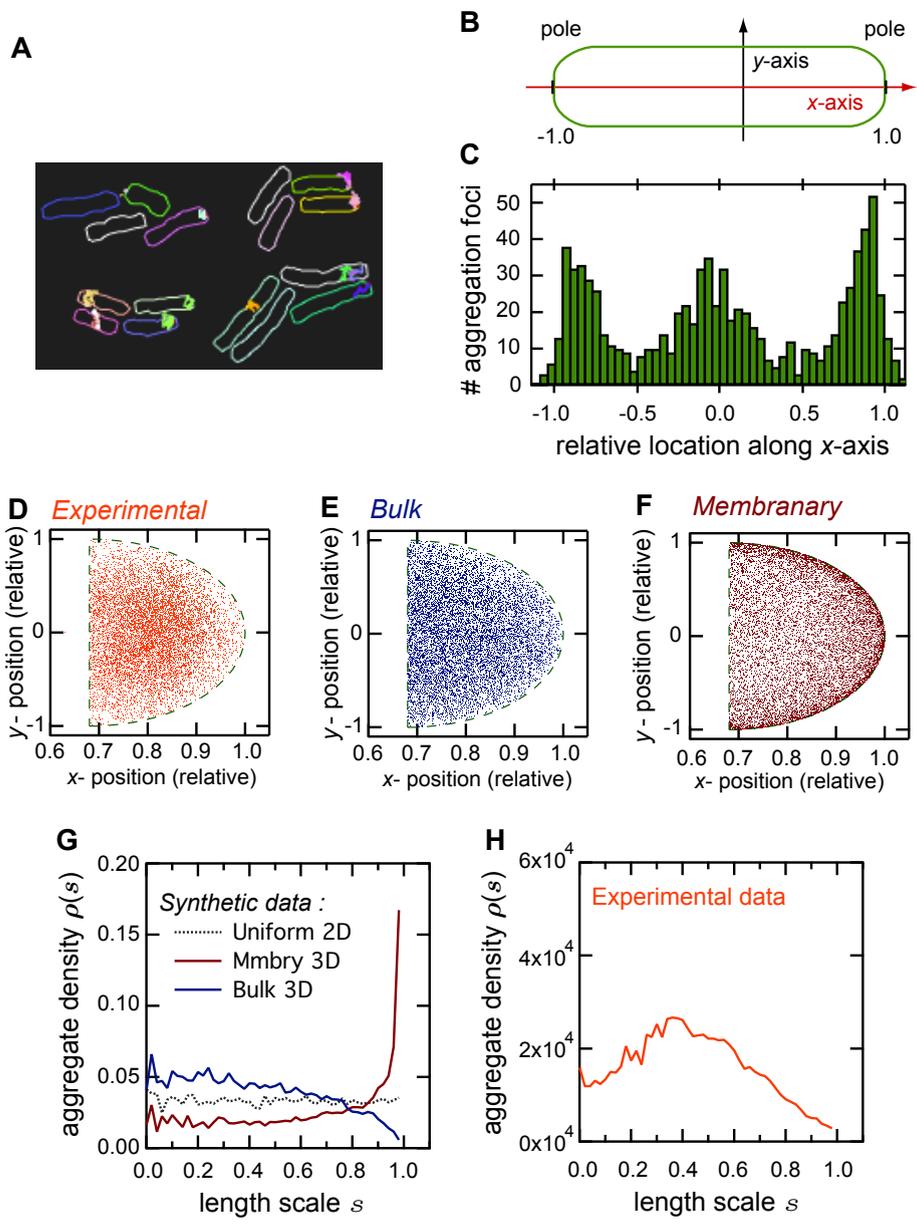


Figure 1

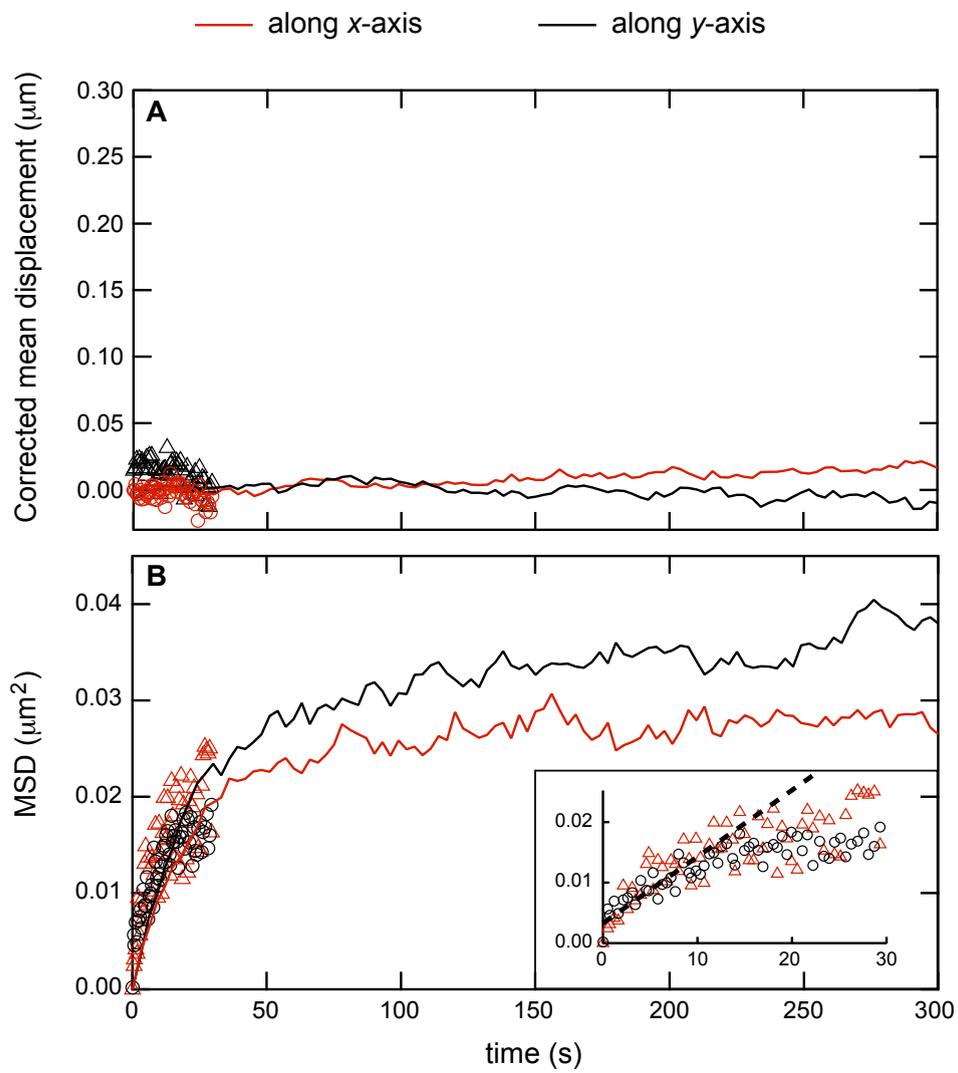


Figure 2

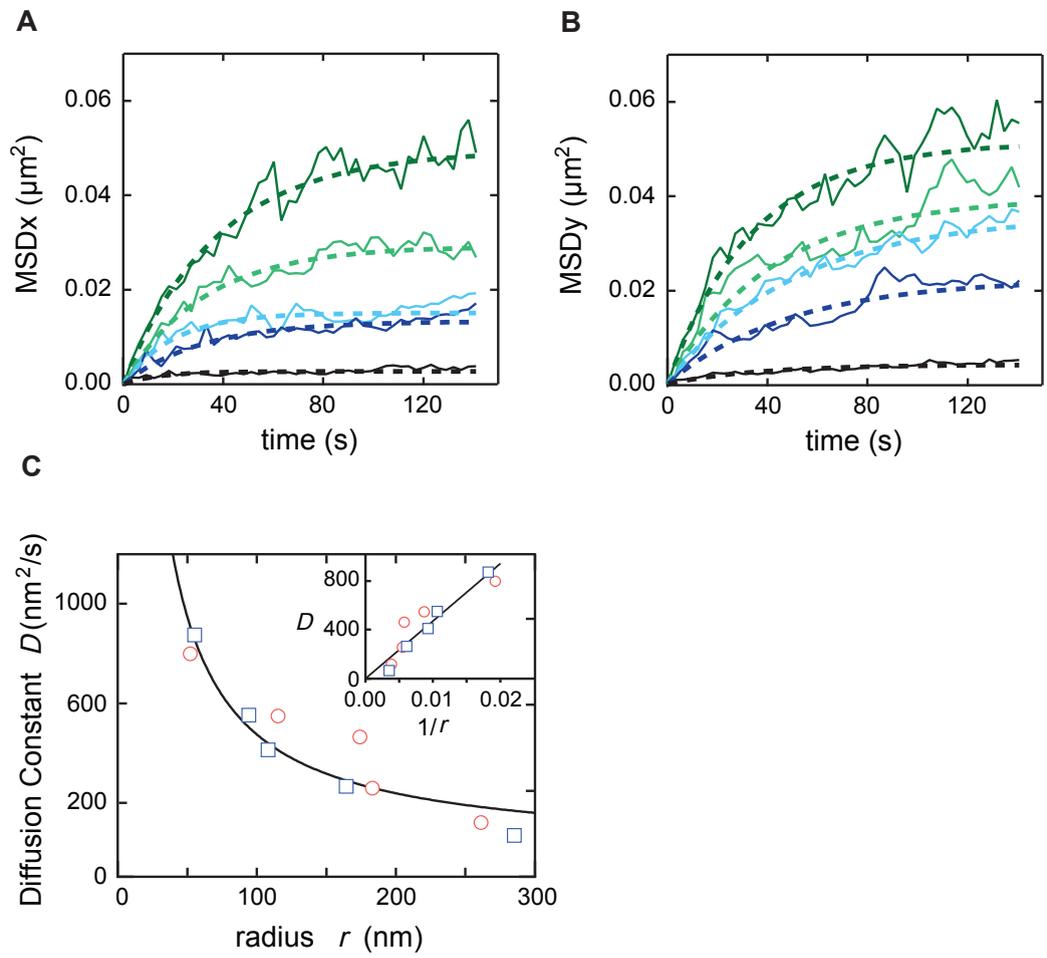


Figure 3

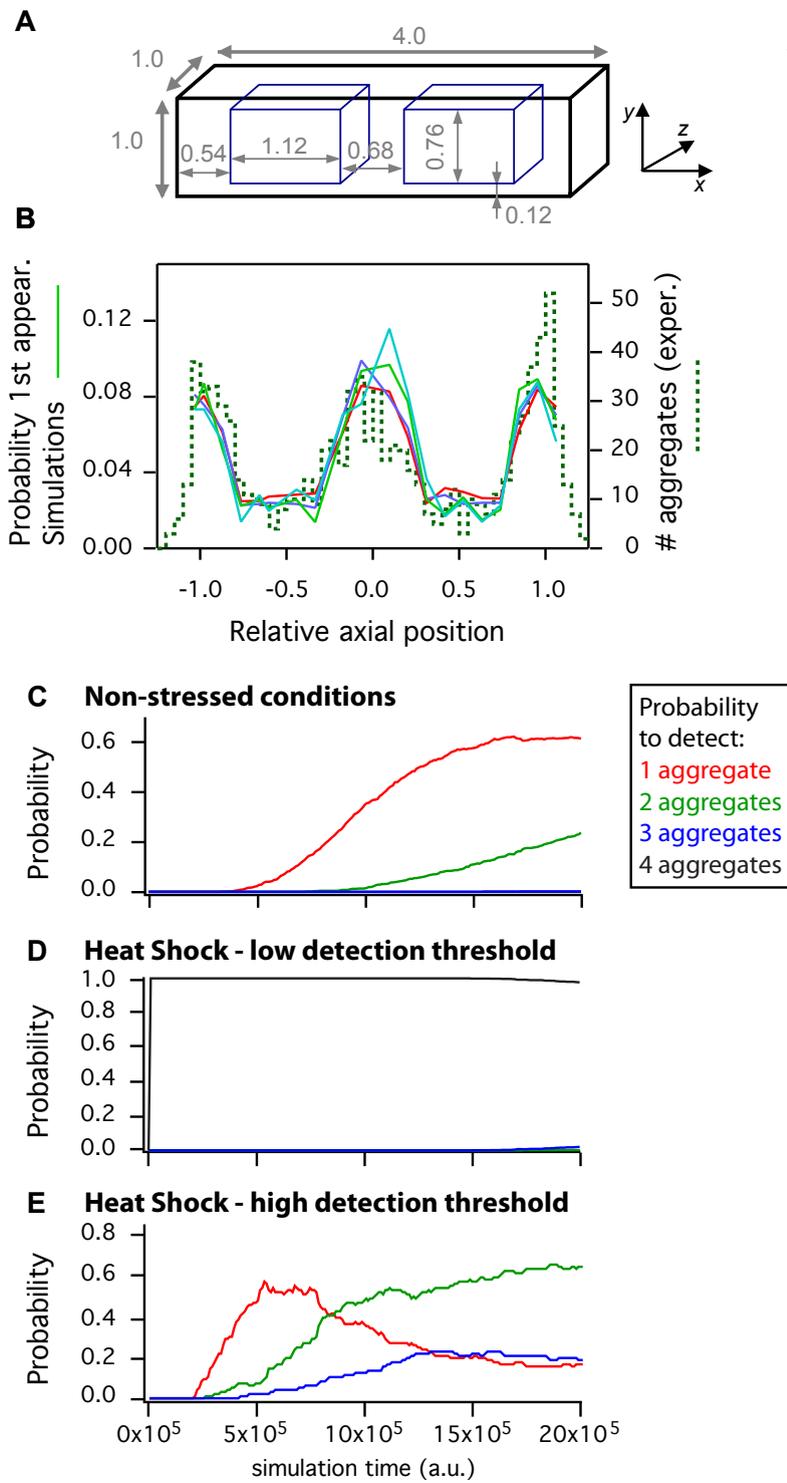
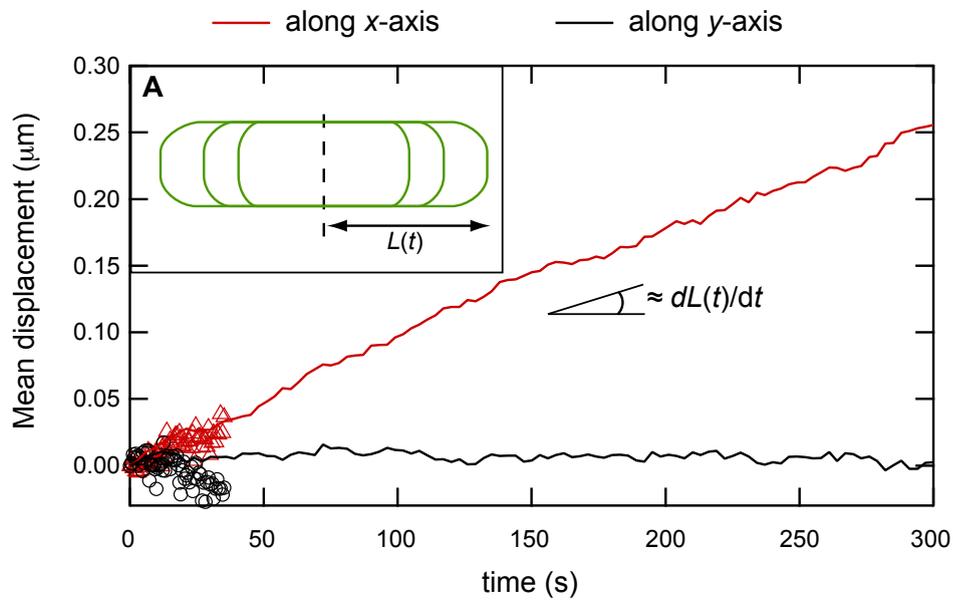


Figure 4



Supplementary Figure 1

II Article : Effet de la protéine chaperonne DnaK sur la transmission du prionoïde RepA-WH1

Les analyses quantitatives étant en cours de finalisation, cette version d'article n'est pas encore finalisée.

II.1 Résumé

Les agrégats protéiques amyloïdes détermine de manière épigénétique le gain de phénotype : soit avantageux, soit protéinopathique. Au moyen de la microscopie microfluidique, nous avons observé pour la première fois une transmission verticale à long terme (de la cellule mère à la cellule fille) d'un agrégat de type amyloïde grâce au prionoïde synthétique RepA-WH1, chez la bactérie *Escherichia coli*. L'observation des lignées bactériennes distinctes montre l'existence de deux types de cellules contenant des agrégats RepA-WH1, distincts des corps d'inclusion : des agrégats globulaires dont la taille augmente avec le temps jusqu'à inhiber la division, ou des agrégats de type "comet" et de forme très allongée qui sont peu néfaste pour la croissance. A faible fréquence, nous avons observé dans une même lignée cellulaire des changements de phénotype. A partir d'un agrégat globulaire, un agrégat de type "comet" est né. La transmission des agrégats de type "comet" est compromise dans des bactéries traitées avec de la myricétine, un inhibiteur de la protéine chaperonne DnaK (Hsp70), ou porteuses de l'allèle *dnaK756(ts)* dans lesquelles les agrégats deviennent tous de type "comet" lors de la sur-expression de DnaK. La protéine chaperonne DnaK semble moduler la conversion en ciblant préférentiellement les agrégats de type "comet", cela rappelant la façon dont sont désassemblées les fibres amyloïdes RepA-WH1 *in vitro*. Cependant, la protéase ClpB, orthologue de la protéine chaperonne Hsp104 chez la levure impliquée dans la propagation des prions, n'a aucun effet significatif.

II.2 Article

Research Article

Title:

Direct assessment of prionoid propagation, toxicity and Hsp70-dependent strain conversion in bacteria

Running Title: Dynamics of a bacterial prionoid *in vivo*

Authors:

Fátima Gasset-Rosa¹, María Moreno-del Álamo^{1,3}, Anne-Sophie Coquel^{2,3}, Ana M. Serrano¹, M. Elena Fernández-Tresguerres¹, Susana Moreno-Díaz de la Espina¹, Ariel B. Lindner^{2,*} and Rafael Giraldo^{1,*}

1. Department of Cell & Synthetic Biology, Centro de Investigaciones Biológicas – CSIC, C/ Ramiro de Maeztu 9, E-28040 Madrid, Spain.
2. Center for Research and Interdisciplinarity (CRI), Faculty of Medicine, Descartes University – INSERM U571, 24 rue du faubourg St Jacques, F-75014 Paris, France.
3. These authors contributed equally to this work.

* Corresponding authors. **ABL:** Tel.: + 33 144412525; Fax: + 33 144412529; E-mail: ariel.lindner@inserm.fr; or **RG:** + 34 918373112; Fax: + 34 915360432; E-mail: rgiraldo@cib.csic.es

Keywords: amyloid proteinopathy / *Escherichia coli* / Hsp70 chaperone / microfluidics / RepA-WH1 prionoid

Number of figures: 9 (plus 4 figures and 6 movies as Supplementary Information)

Character count: 53,788 (spaces included)

Author contributions: FGR (gene constructs, solubility assays/SDD-AGE, fluorescence microscopy, microfluidics); MMdA (TEM and iEM on fibres); AMS (gene constructs); MEFT and SMDE (fluorescence and confocal microscopies, iEM on cells) performed research. ASC analyzed data from microfluidics. ABL and RG conceived the project, directed research and analyzed the data. RG wrote the paper.

Conflict of interest: The authors declare that they have no conflict of interest.

Abstract

Protein amyloid aggregates epigenetically determine gain of function phenotypes, either advantageous or proteinopathic. Through microfluidics, we have directly assessed here for the first time the long-term vertical (mother-to-daughter cells) transmission of an amyloidosis: that caused by the synthetic prionoid RepA-WH1 in *Escherichia coli*. Segregated bacterial lineages maintain two mutually exclusive types (or amyloid strains) of RepA-WH1 aggregates, distinct from inclusion bodies: either multiple globular particles, which get larger until they inhibit cell division, or a single worm-shaped aggregate, mildly detrimental to growth. With low frequency, a lineage carrying one type of aggregate can switch to generate the other one. Propagation of the wormlike strain is compromised in bacteria treated with myricetin, an inhibitor of the Hsp70 chaperone DnaK, or bearing the *dnaK756(ts)* allele, whereas the aggregates in all lineages become wormlike upon DnaK overexpression. DnaK seems to modulate such strain conversion by preferentially targeting the elongated aggregates, recalling how it disassembles RepA-WH1 amyloid fibres *in vitro*. However, ClpB, the orthologue of the chaperone key for the spread of yeast prions, has no substantial effect.

Introduction

The amyloid cross- β assembly (Chiti & Dobson, 2006; Toyama & Weissman, 2011; Eisenberg & Jucker, 2012) is the most stable thermodynamic state than can be achieved by any polypeptide chain, being its native three-dimensional structure a metastable fold kinetically favoured by natural selection in order to achieve a particular function (Baldwin *et al*, 2011). In addition to its intrinsic physical-chemical interest, the amyloid state of some human proteins is at the basis of debilitating, often fatal, neurological disorders (e.g.: β -amyloid peptides in Alzheimer, α -synuclein in Parkinson, poly-Q repeats proteins in Huntington and spinocerebellar ataxias, or SOD1 in amyotrophic lateral sclerosis) and systemic diseases (such as β 2-microglobulin in dialysis-related amyloidosis or amylin in type II diabetes) (Chiti & Dobson, 2006). Furthermore, amyloid aggregates behave as truly epigenetic determinants of selectable traits inherited following a non-Mendelian segregation. They propagate *vertically*, coupled to cytokinesis, from mother cell-to-daughter cells where, as seeds, they template the amyloid conformation on soluble molecules of the same protein (Eichner & Radford, 2011). Prions are a subgroup of amyloids that are also *horizontally* transmissible (infectious) and have contributed much to our current understanding of amyloidosis spread (Tuite & Serio, 2010). They are pathogenic in mammals (PrP in transmissible spongiform encephalopathies; Cobb & Surewicz, 2009), but confer epigenetic, advantageous selectable phenotypes to yeast (Halfmann *et al*, 2012).

Prions have provided evidence for the existence of amyloid *strains*, slightly different three-dimensional structures for a given amyloid assembly that result in distinct phenotypes (Colby & Prusiner, 2011; Toyama & Weissman, 2011; Weissmann *et al*, 2011), and have fuelled the quest for ligand co-factors, including macromolecules directly involved in promoting protein amyloidogenesis (Silva *et al*, 2010; Giraldo, 2010), or small molecule inhibitors (Ehrnhoefer *et al*, 2008; Gasset-Rosa *et al*, 2008; Roberts *et al*, 2009; Ladiwala *et al*, 2011; Woods *et al*, 2011). Among the former, protein chaperones are key players because they are able to assist the refolding and solubilization of misfolded and aggregated proteins, or to bind and stabilize intermediates in the pathway to the assembly of macromolecular complexes, thus being central components of the protein quality triage cellular toolbox (Mayer, 2010; Tyedmers *et al*, 2010a). To deal with aggregated proteins, including amyloids, two sets of chaperones act in concert: the Hsp70-Hsp40-NEF triad and the Hsp104 hexameric disaggregase (Glover & Lindquist, 1998; Goloubinoff *et al*, 1999; Kumar & Sourjik, 2012). Through ATP binding and hydrolysis, both chaperones manage to disentangle amyloids into soluble, less toxic protein monomers. The current view is that Hsp70, assisted by its two chaperone co-factors, extracts protein oligomers from large aggregates and presents any unfolded peptide stretch to Hsp104 (Zietkiewicz *et al*, 2006; Haslberger *et al*, 2010). In yeast prions (Wickner *et al*, 2007) such as [*PSI*⁺] and [*URE3*⁺], there is divergent, sometimes conflicting, evidence on the role of molecular chaperones in prion generation and maintenance (Newnam *et al*, 1999). This is in part due

to the fact that yeast genome bears two sets of paralogous *HSP70* genes (*SSA1-4* and *SSB1-2*), encoding six proteins with high structural similarity but with subtle functional differences (Allen *et al*, 2005; Albanèse *et al*, 2006). Yeast Hsp104p by itself can disrupt Sup35p/[*PSI*⁺] amyloid fibres generating transmissible seed propagons (Shorter & Lindquist, 2004; 2006), although it does so more efficiently together with Hsp70 and its co-chaperones Sis1p or Ydj1p (Shorter & Lindquist, 2008; Newnam *et al*, 2011). However, it is striking that no *HSP104* homologue has been found in metazoa. Having this chaperone a central role in prion propagation in yeast, how the proteotoxic prions in higher eukaryotes manage to efficiently generate transmissible seeds is still an open question.

The quantitative study of individual cell lineages for extensive number or generations has recently become amenable thanks to the coupling between microfluidic devices, fabricated through soft lithography procedures (Vyawahare *et al*, 2010), and epi-fluorescence microscopes. A *mother cell machine* is a microfluidic device recently crafted to monitor bacterial growth in which a founder cell stays at the bottom of a micrometer-wide channel, while the daughter cells are pushed towards the main stream of culture medium (Wang *et al*, 2010). Such setup is actually inspired by the pioneering design by Helmstetter (1967) of a *baby cell machine*, in which the newly born bacteria are recovered through a narrow filter. Until microfluidic devices became available, growth and cell cycle parameters were extracted in a more parsimonious way from time-lapse observations of the growth of microcolonies on agar, with strict limitations in time span, or from flow cytometry experiments, which allow for quantitative measurements but without the direct microscopic observation of the cells. These approaches opened a way to the study of the *aging* phenomenon in bacteria (Stewart *et al*, 2005), which was linked to protein quality control through the preferential inheritance by one of the daughter cells of a single subpolar protein aggregate (inclusion body, IB) (Lindner *et al*, 2008; Winkler *et al*, 2010). Yeast have analogue perivacuolar and juxtannuclear locations for the deposition of aggregated proteins (Kaganovich *et al*, 2008). Although yeast have the pro of being eukaryotic organisms, and thus presumably closer model systems to untangle amyloidosis of clinical relevance, the limited number of divisions of a given mother yeast cell before senescence reduces its applicability in high throughput microfluidic assays. A possibility was recently opened by the successful propagation in *E. coli* of chimaeras between the yeast prion [*PSI*⁺] and GFP, but it still required the help of yeast co-factors, such as seeds of the [*PIN*⁺] prion (Garrity *et al*, 2010). Therefore, the study through microfluidics of the long-term vertical transmission of an amyloidosis using a fully bacterial model system would be of utmost interest.

We have recently reported that the WH1 domain of the RepA plasmid replication protein (Giraldo & Fernández-Tresguerres, 2004; Giraldo *et al*, 2003) is converted from its soluble, dimeric native structure to an amyloid state through a conformational change exerted by transient binding to short, specific DNA sequences *in vitro* (Díaz-López *et al*, 2003; 2006; Giraldo, 2007; Gasset-Rosa *et al*, 2008). When RepA-WH1 fused to a fluorescent protein was expressed in *Escherichia coli*, it aggregated as intracellular amyloid particles (Fernández-Tresguerres *et al*, 2010). Because these

were vertically transmissible to the progeny by coupling to cell division (Fernández-Tresguerres *et al.*, 2010) but lacked microbiological (horizontal) infectivity, RepA-WH1 can be considered a *prionoid* (Aguzzi, 2009). RepA-WH1 has the additional attractiveness of exerting a proteinopathy that severely reduces bacterial fitness (Fernández-Tresguerres *et al.*, 2010), making this prionoid to closer resemble mammalian rather than yeast amyloidosis. RepA-WH1 thus constitutes a unique synthetic model system to study the molecular basis for protein amyloidosis in a minimalist cellular host (Giraldo *et al.*, 2011).

In this article, we have addressed, for the first time, the vertical transmission of an amyloidosis (the RepA-WH1 prionoid) in *E. coli* along many generations by means of microfluidics. This approach has allowed us to define the existence of segregated populations of bacteria bearing one of two different species (amyloid strains) of RepA-WH1 aggregates with characteristic numbers and a distinguishable morphology: either multiple and globular, or single and elongated, which are highly cytotoxic or mildly detrimental for cell proliferation, respectively. Interestingly, these two strains interconvert with a reduced frequency. Since *E. coli* bears single Hsp70 (*dnaK*) and Hsp104 (*clpB*) genes, this has enabled us to directly tackle the roles of both chaperones in amyloidosis, either resorting to defective alleles, complementation assays or through a small molecule inhibitor. We have found that the ClpB disaggregase has no major effect on amyloidosis, whereas DnaK enhances the stable inheritance of the prionoid by promoting the elongated RepA-WH1 oligomers, which readily diffuse to, and are distributed between, the daughter cells during cell division.

Results

Microfluidic studies on E. coli cells reveal the existence of two interconvertible strains of the RepA-WH1 prionoid

In order to characterize the vertical transmission *in vivo* of the bacterial prionoid RepA-WH1 (Giraldo *et al.*, 2011) along tens of generations, we used a microfluidic device in which single *E. coli* MG1655 mother cells stay at the bottom of comb-shaped microchannels (Wang *et al.*, 2010) (Figure 1). Bacteria expressed the hyperamyloidogenic variant His₆-RepA-WH1(A31V) fused to the fluorescent protein mCherry (mRFP) from a *P_{tac}-lacI^q* cassette (Fernández-Tresguerres *et al.*, 2010) inserted in the chromosome (Supplementary Figures S1 and S2A). RepA-WH1 expression revealed the existence of two morphologically distinct species of aggregates: either globular particles distributed in various numbers across the cytoplasm (Fernández-Tresguerres *et al.*, 2010), or a single elongated (worm or comet-shaped) aggregate that originated from a small basal focus in the mother cell (Figure 1A; Supplementary Movie 1). In the newborn cells, globular aggregates grew by the coalescence of smaller foci and by accretion of RepA-WH1 molecules until they reach a size matching cell cross-section, severely reducing the division rate of the cells, inducing their filamentation and, ultimately, leading to stop growth and cell death. On the other hand, the elongated aggregates did not significantly affect cell proliferation, nor impose any apparent hindrance to cell

division [Ariel & Anne-Sophie: please extend further by providing numbers (a table?) on the effects of both types of aggregates]. Starting from a genetically homogenous population in the bulk bacterial culture inoculum, once loaded into the microchannels, bacteria gave birth, along generations, to cells exclusively bearing one or the other type of aggregate. This strict spatio-temporal segregation is expected for any epigenetic determinant based in protein aggregation (Tuite & Serio, 2010).

To check if the two types (globular and wormlike) of RepA-WH1(A31V) aggregates were of amyloid nature, fixed cells were incubated with BTA-1, a fluorophore specific for amyloid cross- β (Wu *et al*, 2008): both kinds of aggregates were stained (Figure 2A), indicative of their amyloid character. As a control, in bacteria bearing RepA-WH1(Δ N37), a metastable mutant deprived of the amyloidogenic peptide stretch (Giraldo *et al*, 1998), a single subpolar focus appeared in most cells barely hampering their growth rates (Figure 1B; Supplementary Movie 2) [Ariel & Anne-Sophie, please provide numbers]. Such RepA-WH1(Δ N37) aggregates behave as expected for conventional IBs (Gatti-Lafranconi *et al*, 2011), including in their heterogeneous composition a nucleus of amyloid-like nature (Carrió *et al*, 2005; Wang *et al*, 2008), which is reflected by their ability to be stained by BTA-1 to some extent (Figure 2A). To check if there were any differences in the affinity of BTA-1 for each type of RepA-WH1 aggregate, the null hypothesis was contrasted by statistical analysis of the variance (one-way ANOVA) of the ratios of fluorescence intensities in the cyan (BTA-1) and red (mCherry) channels in multiple, well-defined particles belonging to those three classes (Figure 2B). The differences in the intensity ratios (affinities for BTA-1) were significant ($p < 0.001$), indicative of a gradation in the amyloidosis of the distinct aggregates: globular WH1(A31V) > elongated WH1(A31V) >> WH1(Δ N37). This result, together with the facts that the two kinds of RepA-WH1(A31V) aggregates were phenotypically distinct (in terms of morphology and relative toxicity) and were stably inherited during many generations, support that they can be truly regarded as different strains/variants of the same prionoid, in a similar sense to prion strains in mammals and yeast (Colby & Prusiner, 2011; Toyama & Weissman, 2011; Weissmann *et al*, 2011). With low frequency, a mother cell in a channel carrying one type of aggregate spontaneously starts to generate the other one, being more common the switch from the wormlike to the globular aggregates than the reversal [Ariel & Anne-Sophie, please provide numbers]. This is reminiscent of the strain conversion phenomenon in the PrP prion (Colby & Prusiner, 2011; Weissmann *et al*, 2011).

The small heat-shock chaperone IbpA is a marker of *bona fide* IBs where, together with its co-chaperone IbpB, targets aggregated proteins for refolding (Ratajczak *et al*, 2009; Kumar & Sourjik, 2012). In cells in which *ibpA-GFP* fusions were transduced (Supplementary Figure S1), tracer green fluorescence colocalized with the basal globular RepA-WH1 aggregates, but not with those in other locations across the cytoplasm or with the wormlike aggregates (Figure 1C; Supplementary Movie 3). This result underlines that RepA-WH1 aggregates are not conventional IBs.

The Hsp70 chaperone DnaK is a major determinant for strain conversion and transmission of the RepA-WH1 prionoid

Since Hsp70 chaperones are central players in cellular protein quality control (Mayer, 2010; Tyedmers *et al*, 2010a), the effect of myricetin, a specific inhibitor of the ATPase activity in DnaK (Chang *et al*, 2011), was tested (Figure 3A; Supplementary Movie 4). Bacterial cells almost exclusively exhibited the globular aggregates, which is compatible with DnaK being involved in the generation of the wormlike RepA-WH1 aggregates. When DnaK was overexpressed from an arabinose (Ara)-inducible promoter cloned in a moderate copy-number (13-30 per cell) p15A plasmid replicon (Supplementary Figure S2B), cells in all the channels exclusively bore the elongated aggregates (Figure 3B; Supplementary Movie 5). This confirmed the results from the inhibition by myricetin pointing to DnaK as a key factor in their generation. On the contrary, expression of the ClpB chaperone, whose Hsp104 homologue has a major role in propagation of yeast prions (Shorter & Lindquist, 2004; 2006; 2008), from an analogous construct (Supplementary Figure S2B) did not alter the normal balance between both types of RepA-WH1 aggregates (Figure 3C; Supplementary Movie 6). Overall, these results indicate that DnaK chaperone, rather than ClpB, is responsible for driving the balance between the globular and wormlike aggregated amyloid species of RepA-WH1 towards the latter, which readily propagates to the daughter cells.

In order to get further evidence, in addition to that provided by microfluidics, of a role of DnaK chaperone in the dynamics of the RepA-WH1 prionoid, cell biology and biochemical studies were carried out with cells grown in bulk cultures. We had previously shown (Fernández-Tresguerres *et al*, 2010) that the RepA-WH1(A31V)-mRFP prionoid formed globular amyloid inclusions in the strain MC4100, similar to those described here (Figures 1-3) in the MG1655 background, in spite of being the former expressed from a multicopy (up to 250 plasmids/cell) pUC18 derivative and the latter from a single copy inserted in the chromosome. A low copy-number vector seemed a good intermediate choice for the new experimental set and thus we cloned the same *Ptac-lacI^f His₆-repA-WH1(A31V)-mRFP* expression cassette used before in a RK2 plasmid replicon (5-8 copies per cell) (Supplementary Figure S2A). The plasmid was transformed into *E. coli* MC4100 and its isogenic JGT20 (*dnaK756*) and JGT3 (Δ *clpB*) strains (Bredèche *et al*, 2001). Because *dnaK* is an essential gene under a number of stress conditions, a thermosensitive rather than a null allele was chosen: DnaK756 allows cell growth at the semi-permissive temperature at which the assays were performed (37 °C) but already has, even at 30 °C, impaired chaperone activities in protein refolding *in vitro* (Buchberger *et al*, 1999) and DNA replication *in vitro* and *in vivo* (Giraldo-Suárez *et al*, 1993).

Expression of RepA-WH1 resulted in the accumulation, fully evident 30 min after the addition of the IPTG inducer, of large fluorescent foci with heterogeneous numbers and distributions, essentially identical in the three strains tested (Figure 4) and alike those observed in microfluidics (Figures 1 and 3), or when expressed from higher plasmid copy numbers (Fernández-Tresguerres *et al*, 2010). The most significant observation was that the expression of RepA-WH1 in the *dnaK756* background,

but not in the $\Delta clpB$ strain, led to a progressive loss of the fluorescent foci from the cell population (40 and 50% foci-free cells after 2 and 4 h, respectively; Supplementary Figure S3A). Since ribosome biogenesis is dependent on DnaK (Alix & Guérin, 1993) it is not surprising that the global expression level achieved for RepA-WH1 in the *dnak756* background is about 60% of that in the two other strains (Supplementary Figure S2A). However, these lower RepA-WH1 levels in *dnak756* must not be the cause of the drop in the number of cells carrying the fluorescent foci, because even lower levels were expected when the same expression cassette was inserted in the chromosome as a single copy, and yet all the cells bore foci (Figures 1-3).

We then attempted complementation by expressing the *dnak* or *clpB* genes cloned in the same compatible vector used in microfluidics (Figure 3B and C; Supplementary Figure S2B). IPTG and Ara concentrations were calibrated, through quantitative Western-blotting of a His₆ tag located at the N-termini of the three proteins, to achieve an equal expression of the RepA-WH1 prionoid and either chaperone during the first half of the experiment (0.5-2 h), and up to two-fold excess of chaperone molecules during the second half (2-4 h) (Supplementary Figure S2B). The order in the sequential addition of each inducer was also tested at 30 min intervals, thus allowing for the accumulation of half the maximum level of the protein induced first before expressing the second one. This would eventually allow distinguishing a possible effect of the chaperones on competing protein aggregation (Ara → IPTG) or on dismantling preformed RepA-WH1 aggregates (IPTG → Ara). An increase in the average cell size and filamentation under heat-shock conditions are other phenotypes characteristic of the *dnak756* allele (Bredèche *et al*, 2001). Expression of DnaK reduced the average cell size by 40% (Figure 4; Supplementary Figure S3B) indicating a successful complementation. Interestingly, the population of cells carrying aggregates increased to $\geq 90\%$, which is compatible with a promoting role for DnaK in the propagation of the RepA-WH1 prionoid. However, in the $\Delta clpB$ genetic background the fraction of cells bearing foci and the average cell size increased just slightly upon ClpB complementation (Supplementary Figure S3A, and B), suggesting that both phenotypes are not dependent on this chaperone. There was no significant effect in MC4100 (WT) cells when either chaperone was expressed under the same conditions.

A major effect of DnaK complementation was a change in the appearance of the RepA-WH1 aggregates (Figure 5), which acquired the elongated, wormlike shape already observed in the microfluidic experiments upon DnaK overexpression (Figure 3B) in 70-95% of the cells (Supplementary Figure S3C). This suggests remodelling of the aggregates by the direct action of the Hsp70 chaperone. On the contrary, ClpB overexpression resulted in highly compacted foci (Figure 5) in 75-95% of the cells, a fraction close to that observed in both *clpB^{wt}* and $\Delta clpB$ backgrounds (Supplementary Figure S3C) and in microfluidics (Figure 3C). On the issue of the number of aggregates per cell (Supplementary Figure S3D), the characteristic two foci (at $\frac{1}{4}$ and $\frac{3}{4}$ axial positions) (Fernández-Tresguerres *et al*, 2010) was, as expected, the most abundant class (50-60%) in the MC4100 and $\Delta clpB$ cells. On the contrary, cells carrying multiple (3 or more) aggregates were

most common in the *dnaK756* background, especially after 4 h of induction, reflecting the larger average cell size of this strain (Supplementary Figure S3B). The expression of either chaperone increased the proportion of cells with 3 or more aggregates to 50-70%. There was no clear dependence of any of these phenotypes on the relative order of induction of the prionoid and each chaperone, except for ClpB on the number of aggregates. In the $\Delta clpB$ background there was a clear inversion in the relative abundance of cells with 1-2 or more RepA-WH1 aggregates after 2 or 4 h of induction. Expressing RepA-WH1 before ClpB made 1-2 aggregates per cell more abundant along the whole time course, whereas the opposite order of induction resulted in multiple aggregates as the dominant group. This suggests that high pre-existing levels of ClpB were required to create multiple amyloid nuclei, while in the case of DnaK complementation cells with multiple aggregates were the most abundant class irrespective of the order of expression, as expected if the Hsp70 chaperone were more effective than the Hsp104 disaggregase in generating amyloid primordia.

DnaK generates small oligomeric particles of the RepA-WH1 prionoid

To explore with further structural detail the action of the DnaK and ClpB chaperones on the RepA-WH1 amyloid aggregates, we carried out immuno-localization studies with cells from bulk cultures expressing one of the two chaperones and the prionoid in conditions close to those used in microfluidics (Figure 3). Bacteria were visualized by confocal laser fluorescence microscopy (Figure 6) and transmission electron microscopy (TEM) (Figure 7). As expected from their large sizes, both the nucleoid and the RepA-WH1 particles mutually excluded. Neither DnaK nor ClpB significantly colocalized with the nucleoid or the large globular RepA-WH1 aggregates, labelling either the periphery of the globular aggregates (ClpB) or the tails of the wormlike aggregates (DnaK) (Figure 6B). TEM confirmed these distinct patterns and revealed that inside the large electron-dense RepA-WH1 aggregates (previously characterized in Fernández-Tresguerres *et al.*, 2010) there were ellipsoid sectors, which probably reflect distinctive protein packing, where most of the DnaK label concentrated (dashed red circles in Figure 7A). Such elongated particles resemble the wormlike species of the RepA-WH1 prionoid seen in bacteria growing in the microfluidic channels (Figures 1 and 3). In addition, some DnaK molecules also located on smaller particles (dashed green circles in Figure 7A), suggesting either DnaK-promoted fragmentation of the RepA-WH1 aggregates or its role in the genesis of the amyloid nuclei.

In order to survey a possible active role for DnaK in promoting disassembly of the prionoid, the purified chaperone was incubated, at physiological temperature and in the presence of its ATP cofactor, with RepA-WH1(A31V) amyloid fibres pre-assembled under standard conditions *in vitro* (Giraldo, 2007). DnaK was able to untangle the fibres into roughly homogeneous globular particles that, from their size (≈ 20 nm ϕ), must be composed of a discrete number of RepA-WH1 and chaperone molecules (Figure 8). In addition, polyclonal antibodies against RepA-WH1 were also labelling the detached oligomers and the periphery of the fibres, suggesting that the major epitope must be either excluded from (or buried into) their axial core.

A biochemical analysis was also performed, under the conditions assayed in bulk cultures, on the distribution of RepA-WH1 between the soluble and aggregated fractions in whole cell extracts. These were separated by centrifugation and then blotted with anti-His tag monoclonal antibodies, which recognize both the RepA-WH1 protein and the DnaK/ClpB chaperones (Supplementary Figure 4A). As expected, most ($\geq 60\%$) of the prionoid was found at the insoluble fraction. This was especially evident in the Hsp70-deficient (*dnaK756*) background (up to 80%; Supplementary Figure 4B). When either DnaK or ClpB were expressed, substantial amounts of the chaperones (50% at 2 h, and $\geq 75\%$ at 4 h) were targeted to the aggregated fraction, whose contents in RepA-WH1 did not significantly vary (Supplementary Figure 4). Semidenaturing-detergent agarose gel electrophoresis (SDD-AGE) allows for separating any protein sample into its monomeric and oligomeric/aggregated fractions up to molecular weights well in the MDa range (Bagriantsev *et al*, 2006). When total extracts from the *E. coli* cells were separated by SDD-AGE (Figure 9), the existence of a major fraction of detergent-resistant RepA-WH1 oligomers became evident in the WT background. These oligomers completely disappeared in the *dnaK756* strain, while they just reduced their presence in the $\Delta clpB$ extracts (Figure 9A). Complementation of *dnaK756* through the expression of DnaK led to the recovery of the detergent-resistant oligomeric RepA-WH1 fraction (Figure 9A), which increased slightly upon ClpB complementation. Although complexes between RepA-WH1 oligomers and DnaK were visualized at the electron microscope (Figures 7A and 8), they seemed to be disrupted by the detergent during electrophoresis (Figure 9B), which should also be the case for ClpB (Figure 9C). It is likely that the DnaK-dependent, detergent-resistant RepA-WH1 oligomeric species detected by SDD-AGE corresponds to the worm-shaped aggregated particles visualized in the microfluidic experiments (Figures 1 and 3) and through TEM (Figure 7A).

Overall, the studies reported here on the transmission of the RepA-WH1 prionoid in *E. coli* cells point to a major, direct role of the Hsp70 chaperone DnaK in transforming large globular proteotoxic RepA-WH1 aggregates into a mixed pool of elongated and relatively small oligomers, which are mildly toxic and readily diffuse from mother-to-daughter cells. These would be the bacterial counterparts of the transmissible amyloid propagons previously characterized in yeast prions (Tuite & Serio, 2010). Furthermore, the two distinct species of the RepA-WH1 prionoid provide the first evidence on the existence in bacteria of temporally stable but interconvertible amyloid strains with clearly different associated phenotypes, as previously defined in yeast and mammalian prions (Colby & Prusiner, 2011; Toyama & Weissman, 2011; Weissmann *et al*, 2011). These findings validate the suitability of the RepA-WH1 prionoid as a minimalist model system to study the universal basis of protein amyloidosis.

Discussion

One of the unsolved questions in protein science is the nature of the co-factors, besides mutations and extreme physicochemical conditions, that trigger the conformational changes enabling a folded soluble protein to self-assemble into amyloid aggregates *in vivo* (Giraldo, 2010). This conundrum is at the core of the biology of prions, a particular class of amyloidogenic proteins found in yeast and vertebrates with the ability to act as infectious microbiological particles, since Prusiner proposed for the PrP prion that a protein co-factor should be involved in facilitating the acquisition of an amyloid state (Prusiner, 1998). A number of biomolecules, most of them polyanions, have been recently found to promote prion amyloidogenesis (Silva *et al*, 2008; Soto, 2011). Molecular chaperones lead the race to become that missing protein cofactor in the case of yeast prions, either by assisting the conformational conversion of native amyloidogenic proteins (Hsp70-40) or by actively catalyzing the shearing of amyloid fibres into small, and thus easily transmissible, particles (Hsp104) (Mayer, 2010; Tyedmers *et al*, 2010a). Many amyloidogenic proteins share with prions the ability to template the conversion of soluble protein molecules into the amyloid state and to be spread through vertical (mother-to-daughter cells) transmission, albeit lacking horizontal infectivity, thus being termed prionoids (Aguzzi, 2009). These include proteins involved in neurodegenerative diseases, in which Hsp70s have a protective role against amyloid aggregation (Muchowski *et al*, 2000; Luk *et al*, 2008; Hoshino *et al*, 2011; Patterson *et al*, 2011). We have recently engineered a synthetic, minimal system to study protein amyloidosis, both *in vitro* and *in vivo*, based in the bacterial protein RepA-WH1 (Giraldo *et al*, 2011). RepA-WH1 behaves as a prionoid in *E. coli* (Fernández-Tresguerres *et al*, 2010) and shares with PrP having nucleic acids as co-factors promoting amyloidogenesis (Giraldo, 2007; Gasset-Rosa *et al*, 2008).

In this report we have characterized the vertical, mother cell-to-daughter cells transmission of the RepA-WH1 prionoid in *E. coli* along tens of generations by using a microfluidic device coupled to a fluorescence microscope. Microfluidics is an emergent approach to study biological macromolecules (Vyawahare *et al*, 2010) and individual cells (Yin & Marshall, 2012) in quantitative terms. A wealth of biologically sound information can be potentially extracted upon altering physico-chemical variables, as temperature, the composition of the buffer or culture media, including the presence of natural ligands, inhibitors or protein factors on the behaviour of macromolecules or on the growth rates and fitness of cells. In the particular case of the *mother cell machine* (Wang *et al*, 2010) that we have used, single founder *E. coli* cells colonize the bottom of cell-wide microchannels that are continuously fed by culture medium under a laminar flow regime. Bacteria in the microchannels readily uptake labelled metabolic precursors and grow with a rate close to that in bulk liquid cultures (Wang *et al*, 2010), underlining the physiological relevance of the studies performed. Besides following the fate of the old pole in a founder mother cell, the microfluidic setup has allowed us to survey, while bacteria are still inside the microchannels, the intracellular dynamics of fluorescence-labelled RepA-WH1 across 4-5 additional cells resultant from subsequent cell division rounds. The experiments presented here constitute the first application of microfluidics to the study of protein

amyloidosis *in vivo*, with the single recent precedent of modelling the nucleation and growth on insulin amyloid aggregates *in vitro* (Knowles *et al*, 2011). At the whole organism level, protein amyloidoses have been usually studied in the case of prions both through genetics in yeast and after successive passages of infectious protein particles in mice, being the former an indirect approach and the latter having the drawback of a low throughput. Studying the minimal, synthetic bacterial RepA-WH1 prionoid through microfluidics has provided, at the single cell level and for lineages spanning many generations, a direct visual assessment of the existence, stability and interconversion of two amyloid strains, that is: distinct aggregated species of a given protein with characteristic morphologies (globular *vs.* elongated) and numbers (multiple *vs.* single) and distinct affinities for BTA-1 and cytotoxicities (acute *vs.* mild). The existence of inter-convertible globular and elongated aggregates was also described for the yeast prion $[PSI^+]$ (Tyedmers *et al*, 2010b; Kawai-Noma *et al*, 2010).

Combining microfluidics *in vivo* with complementary microscopy and *in vitro* studies have also allowed us to address the question of how the Hsp70 (DnaK) and Hsp104 (ClpB) bacterial chaperones influence the propagation and the structural dynamics of the RepA-WH1 prionoid. Cooperation between Hsp70 and Hsp104 is instrumental in prion propagation in yeast (Shorter & Lindquist, 2008; Haslberger *et al*, 2010). Two noticeable exceptions are $[GAR^+]$, a yeast prion independent of Hsp104 but cured in $\Delta ssa1-2$ backgrounds, which is not linked to any amyloidogenic protein (Brown & Lindquist, 2009); and New1p, which contributes to $[PSI^+]$ propagation by enhancing nucleation of Sup35p (through their homologous N-terminal domains) and then fragmenting amyloid fibrils (through its C-terminal ATPase), thus becoming independent of Hsp104 (Inoue *et al*, 2011). In the case of RepA-WH1, when assessed individually, DnaK clearly has a central role in promoting prionoid propagation whereas ClpB has a marginal influence. Unlike prions in yeast, the spread of the RepA-WH1 prionoid in bacteria does not rely mainly on the Hsp104 disaggregase activity, which correlates with the fact that the amyloidogenic sequences in both kinds of proteins are quite different: they are either based in Q/N repeats (yeast prions) or in hydrophobic stretches (RepA-WH1 and many other proteinopathic amyloids). The existence of several paralogue *HSP70* genes in eukaryotes (Albanèse *et al*, 2006) poses an additional challenge to understanding their contribution to prion amyloidogenesis. However, by having a single *HSP70* gene, *E. coli* has allowed us to uncover that DnaK enables RepA-WH1 prionoid transmissibility by generating, from compact globular amyloid aggregates, small oligomeric particles and worm-shaped aggregates that are readily propagated to the offspring cells. Protein dimers of the Rep family of plasmid DNA replication initiators are remodelled by DnaK to become active monomers, through binding to their WH1 domains (Wickner *et al*, 1992; Kim *et al*, 2002). Thus it makes sense that the same chaperone leads the disassembly of the RepA-WH1 amyloid aggregates into discrete oligomers, as reported here.

The bimodal (globular *vs.* wormlike) shape distribution of the amyloid aggregates, their higher proteo-toxicity, poorer co-localization with IbpA and higher avidity for an amyloidotropic agent clearly set the RepA-WH1 prionoid apart from conventional IBs. It has been shown that ClpB disassembles proteins that aggregate as IBs upon overexpression, in a pathway towards their solubilization and/or proteolytic degradation (Doyle & Wickner, 2009). For the RepA-WH1 prionoid, we have found in its relative independence on ClpB for propagation another difference with conventional IBs. It is noteworthy that DnaK was previously reported to be distributed around the surface of IBs (Carrió & Villaverde, 2005) not, as found here, selectively clustered in the elongated RepA-WH1 aggregates.

Considering the apparent absence of any *HSP104* gene in higher eukaryotes, the results reported here underline recent findings (Shorter, 2011) on Hsp70s chaperones as potential modulators of protein amyloidosis in mammalian cells, underline the potential of up-regulating Hsp70s expression to reduce amyloid toxicity, and illustrate the power of synthetic minimal bacterial model systems in addressing relevant biological questions.

Materials and methods

Gene constructs

The chromosomal insertion of the *His6-RepA-WH1(A31V/ΔN37)-mCherry* expression cassettes are described in Supplementary Figure S1. *pRK2-WH1(A31V)-mCherry* and *p15A-dnaK/clpB* plasmid vectors were constructed as indicated in Supplementary Figure S2.

Microfluidics

Microfluidic cells were fabricated of poly-dimethylsiloxane by using standard soft lithography techniques, as described elsewhere (Wang *et al.*, 2010). MG1655 strain derivatives (Supplementary Figure S1) were grown overnight at 37 °C in minimal medium supplemented with glucose and casamino acids (M9-CAA). Then 50 μL of each cell culture were diluted in 20 mL of fresh medium and were grown to an $OD_{600nm} = 0.2$. Cultures were centrifuged, washed twice with M9-CAA supplemented with PEG4000 to 1.5% and the final cell pellets were resuspended in 200 μL. Cells were then loaded into the microfluidic chip by means of a centrifugation pulse and, immediately, the device was fixed on the microscope and a continuous (2 mL.h⁻¹) laminar flow of filtered M9-CAA was passed through the setting at a constant temperature of 37 °C. When required, chloramphenicol (30 μg.mL⁻¹), myricetin (15 μg.mL⁻¹; Sigma), IPTG (0.05 mM) or L-arabinose (0.15%) were injected by means of syringes coupled to valves placed at the inlet of the flow cell. Bacterial growth was followed for 8-72 h, taking each 6 min up to 15 image frames across the chip. Correlative images (in TIFF format) were edited as movies using ImageJ (Collins, 2007).

Data analysis was carried out by a plugin ImageJ developed at INSERM U1001 (Song 2012). For each channel, a new frame is created in function of time for the phase contrast images and one for the fluorescence images. So each channel is analyzed separately.

A low pass filter (3by3 pixels) is applied to the frame and then the images are normalized by subtracting the background. The images are stabilized by using the plugin Image Stabilizer and the parameters of stabilization are applied to fluorescence images.

The phase contrast images are finally denoised using a total variation algorithm (Rudin *et al*, 1992). Cells are segmented using the threshold method and the local histogram in a automatic way and globally by hand.

The lineage from the mother the cell is based on a morphological parameter that is the cell length. When a cell divides, the total length of the two new cells must be superior to 90% of the cell length before the division.

To follow protein aggregates, fluorescence images are overlapped to phase contrast images. The studied aggregates have two different shapes: foci and comet. The plugin implements a threshold based on circularity to distinguish them.

Some corrections by hand are necessary to correct the segmentation (cell and protein aggregates).

Bulk bacterial cultures

E. coli MC4100 cells, or their isogenic strains JGT20 (*dnaK756*) and JGT3 (Δ *clpB*) (Thomas & Baneyx, 1998), were inoculated in LB medium supplemented with 0.1% glucose and, if required, with ampicillin (100 $\mu\text{g.mL}^{-1}$) and chloramphenicol (30 $\mu\text{g.mL}^{-1}$). Bacteria were grown at 37 °C to exponential phase ($\text{OD}_{600\text{nm}} = 0.2$) and then were induced with 0.5 mM IPTG. 1 mL aliquots were taken each 30 min for 4 h. For dual chaperone and WH1 expression, the promoter-specific inducers (0.1 mM IPTG and 0.5% L-arabinose) were successively added, in either order, within a 30 min interval.

Protein purification

RepA-WH1(A31V) was purified as previously described (Giraldo, 2007). DnaK was purified from *E. coli* TG1 cells carrying the pMOB45-*dnaK* vector (Giraldo-Suárez *et al*, 1993). Bacterial cultures were grown in LB at 30 °C to $\text{OD}_{600\text{nm}} = 0.45$, when a temperature shift to 42 °C was performed for 1 h. They were then further incubated at 30 °C overnight. Cells were harvested, resuspended in 50 mM Tris.HCl pH 8.0, 10% sucrose and 1/8 volumes of lysis solution (0.5 M KCl, 50 mM EDTA pH 8.0, 10 mM β -MeEtOH, 1% Brij-58, 0.1 mM spermidine, 1 mM p-NH₂-benzamidine). Lysozyme (to 2 mg.mL^{-1}) and RNase-A (to 0.1 mg.mL^{-1}) were added and cell suspension left on ice for 45 min, followed by 5 min incubation at 37 °C. Ultracentrifugation (Beckman Ti45, 30,000 rpm) was carried out for 45 min at 4 °C. The supernatant was loaded in a SP-Sepharose (GE Healthcare) column equilibrated at 4 °C with buffer-A (0.025 M KCl, 20 mM HEPES pH 7.5, 1 mM β -MeEtOH, 0.1 mM EDTA, 10% glycerol). Flow-through was loaded in a Q-Sepharose (GE Healthcare) column

equilibrated in buffer-A. A linear gradient was developed between buffer-A and the same buffer supplemented with KCl to 0.75 M. Fractions containing DnaK according to SDS-PAGE were pooled and 1 mL aliquots were loaded at 0.5 mL.min⁻¹ in a Sephacryl S200-HR (GE Healthcare) 2.6x70 cm gel filtration column equilibrated in buffer-B (0.25 M KCl, 20 mM HEPES pH 7.5, 1 mM β-MeEtOH, 0.1 mM EDTA, 10% glycerol). DnaK peak fractions were concentrated under N₂ pressure through an Amicon 50 mL cell (50 kDa cut-off membrane). DnaK concentration was determined through A_{280nm} ($\epsilon^{280} = 14,650 \text{ M}^{-1} \cdot \text{cm}^{-1}$, estimated at <http://web.expasy.org/protparam>).

Fluorescence microscopy

Observation of bacterial cells from bulk cultures was performed with a Nikon Eclipse 90i microscope, equipped with a CFI PLAN APO VC 100x (NA 1.40) oil immersion objective and a Hamamatsu ORCA-R² CCD camera. The following excitation (EX) and emission (EM) filters and exposure times were used: m-Cherry (EX 543/22, EM 593/40; 200 ms), Alexa 488 (EX 482/35, EM 536/40; 200 ms), DAPI (EX 387/11, EM 447/60; 2 s), BTA-1 (EX 438/24, EM 483/32; 3 s). A DIC image of each field was also captured (100 ms). Bacterial growth in the microfluidic cell was followed with a Nikon Eclipse inverted microscope (Ti), equipped with a 100x objective, but with phase contrast optics, and similar filters and coupled to a Photometrics Cool-Snap HQ2 CCD camera. Image acquisition at this microscope set was controlled by MetaMorph software (Molecular Devices). The following exposure times were used: 100 ms of transmitted light (ND32 neutral filter, for phase contrast), 1 s (mCherry) and 2 s (YFP). Confocal laser microscopy was carried out in a Leica TCS-SP2-AOBS, mounting a HCX PL APO CS 100x (NA 1.40) objective. Images were analyzed with the Nikon NIS-Elements AR software.

Bacterial culture aliquots were fixed in formaldehyde and mounted on poly-L-lysine coated slides (Fernández-Tresguerres *et al.*, 2010). For immuno-fluorescent staining, cells were treated with lysozyme and incubated with antibodies as indicated (Fernández-Tresguerres *et al.*, 2010), using as primary antibodies either a mouse monoclonal (mAb) anti DnaK (clone 8E2/2, Enzo Life Sciences; 1:1,000 dilution) or a rabbit polyclonal anti-ClpB (1:5,000) (Tek & Zolkiewski, 2002), and as secondary antibodies Alexa 488-coupled anti-mouse or anti-rabbit (Molecular Probes; 1:100). Incubation with fluorophores specific for DNA (DAPI, 4',6-diamidino-2-phenylindole (Serva); 5 μg.mL⁻¹) or amyloid (BTA-1, 2-(4-methylaminophenyl) benzothiazole (Sigma); 0.5 mM from a stock in DMSO) was performed for 20 min during fixation, when indicated.

Transmission electron microscopy (TEM)

RepA-WH1(A31V) amyloid fibres were assembled *in vitro* as indicated (Giraldo, 2007) and then examined on carbon coated 400-mesh copper grids (Ted Pella, Inc.) after negative staining with 2% aqueous uranyl acetate. For immuno-electron microscopy (iEM), RepA-WH1(A31V) fibres were incubated with 25 μM DnaK in 0.1 M Na₂SO₄, 10 mM Mg₂SO₄, 20 mM HEPES pH 8.0, 2 mM β-MeEtOH, 10% glycerol, with or without 0.1 mM ATP, at 37 °C for 1h. When iEM was carried out

with *E. coli* cells, bacteria were processed as described (Fernández-Tresguerres *et al.*, 2010). The primary antibodies used were a rabbit polyclonal antiserum raised against purified RepA-WH1(A31V) (1:10) and the anti-DnaK (1:200) or the anti-ClpB (1:5,000) antibodies (see above). Anti-mouse or anti-rabbit antibodies conjugated with 10 nm gold particles (Sigma; 1:50) were then used as secondary probes (Fernández-Tresguerres *et al.*, 2010). When iEM was performed on fibres assembled *in vitro*, these were adsorbed on glow-discharged grids and successively floated on drops of blocking solution (0.05% Tween-20 and 2% BSA in PBS, 30min) and the anti-DnaK or anti-WH1 primary antibodies (1:500 in blocking solution, for 1 h at room temperature). After three washes (10 min each) with 0.05% Tween-20 in PBS, grids were incubated with the gold-conjugated anti-mouse or anti-rabbit antibodies (1:50, for 1 h). Specimens were then washed three times with 0.05% Tween-20 in PBS and once with bi-distilled water before air-drying and uranyl acetate staining. Specimen observation was carried out in a JEOL JEM-1230 transmission electron microscope operating at 80 kV and images were captured with a TVIPS TemCam-F416 CMOS camera.

Semi-denaturing detergent agarose gel electrophoresis (SDD-AGE)

Cells from 25 mL cultures were harvested after 2.5 h induction ($OD_{600} \approx 2.0$) and resuspended in 400 μ L of 25 mM Tris.HCl pH 6.8, 250 mM NaCl, 5 mM EDTA, 10% glycerol, supplemented with protease inhibitors (Roche). 1/3 of the volume of this suspension was added to a tube containing silica beads (1.0 mm ϕ , Lysing Matrix C, MP Biomedicals) and the mixture was shaken in a FastPrep-24 homogenizer (4 cycles at speed-IV, 30 s each, 4 °C). 3 μ L of each lysate were diluted to 40 μ L with 0.5xTAE, 5% glycerol, 2% sarkosyl, 0.1 mg.mL⁻¹ bromophenol blue, plus protease inhibitors. Samples were incubated at room temperature for 10 min and then loaded in 1.5% agarose-TAE, 1% SDS gels. Electrophoresis was run at 100 V – 400 mA, for 7.5 h at 10 °C. Gels were then electro-blotted (Transfer Blot, BioRad) to PVDF membranes in TAE – 0.1% SDS at 16 V – 400 mA, for 15 h at 10 °C. After membrane blocking, primary antibodies were incubated for 2 h: anti-WH1(A31V) (1:500), anti-DnaK (1:2,500) or anti-ClpB (1:3,000). Labeling with HRP-conjugated antibodies and chemio-luminiscent detection was carried out as indicated (Supplementary Figure S2). Washing in 8 M Gu.HCl for 1 min allowed to strip the membranes for successive antibody probing.

Acknowledgements

We are grateful to F. Baneyx for the JGT20 and JGT3 strains, V. de Lorenzo for the pSEVAs vectors and M. Zolkiewski for the anti-ClpB antibody. M.T. Seiseddos, F. González and M. Carnota are acknowledged for their technical assistance, and the animal house service at CIB-CSIC for the preparation of the anti-WH1(A31V) antiserum. FG-R thanks the JAE-CSIC program for a Predoc grant and FEMS for financing her stay at CRI with a short-term fellowship. This work has been

supported by grants [XXX; Ariel, please fill-in] to ABL and from Spanish MICINN/MEC (BIO2009-06952 and CSD2009-00088) to RG.

References

- Aguzzi A (2009) Beyond the prion principle. *Nature* **459**: 924-925
- Albanèse V, Yen-Wen Yam A, Baughman J, Parnot C, Frydman J (2006) Systems analyses reveal two chaperone networks with distinct functions in eukaryotic cells. *Cell* **124**: 75-88
- Alix JH, Guérin MF (1993) Mutant DnaK chaperones cause ribosome assembly defects in *Escherichia coli*. *Proc Natl Acad Sci USA* **90**: 9725-9729
- Allen KD, Wegrzyn RD, Chernova TA, Müller S, Newnam GP, Winslett PA, Wittich KB, Wilkinson KD, Chernoff YO (2005) Hsp70 chaperones as modulators of prion life cycle: Novel effects of Ssa and Ssb on the *Saccharomyces cerevisiae* prion [PSI⁺]. *Genetics* **169**: 1227-1242
- Bagriantsev SN, Kushnirov VV, Liebman SW (2006) Analysis of amyloid aggregates using agarose gel electrophoresis. *Methods Enzymol* **412**: 33-48
- Baldwin AJ, Knowles TP, Tartaglia GG, Fitzpatrick AW, Devlin GL, Shammass SL, Waudby CA, Mossuto MF, Meehan S, Gras SL, Christodoulou J, Anthony-Cahill SJ, Barker PD, Vendruscolo M, Dobson CM (2011) Metastability of native proteins and the phenomenon of amyloid formation. *J Amer Chem Soc* **133**: 14160-14163
- Bredèche MF, Ehrlich SD, Michel B (2001) Viability of rep recA mutants depends on their capacity to cope with spontaneous oxidative damage and on the DnaK chaperone protein. *J Bacteriol* **183**: 2165-2171
- Brown JCS, Lindquist S (2009) A heritable switch in carbon source utilization driven by an unusual yeast prion. *Genes Dev* **23**: 2320-2332
- Buchberger A, Gässler CS, Büttner M, McMacken R, Bukau B (1999) Functional defects of the DnaK756 mutant chaperone of *Escherichia coli* indicate distinct roles for amino- and carboxyl-terminal residues in substrate and co-chaperone interaction and interdomain communication. *J Biol Chem* **274**: 38017-38026
- Carrió MM, Villaverde A (2005) Localization of chaperones DnaK and GroEL in bacterial inclusion bodies. *J Bacteriol* **187**: 3599-3601
- Carrió M, González-Montalbán N, Vera A, Villaverde A, Ventura S (2005) Amyloid-like properties of bacterial inclusion bodies. *J Mol Biol* **347**: 1025-1037
- Chang L, Miyata Y, Ung PMU, Bertelsen EB, McQuade TJ, Carlson HA, Zuiderweg ERP, Gestwicki JE (2011) Chemical screens against a reconstituted multiprotein complex: Myricetin blocks DnaJ regulation of DnaK through an allosteric mechanism. *Chem Biol* **18**: 210-221
- Chiti F, Dobson CM (2006) Protein misfolding, functional amyloid, and human disease. *Annu Rev Biochem* **75**: 333-366

- Cobb NJ, Surewicz WK (2009) Prion diseases and their biochemical mechanisms. *Biochemistry* **48**: 2574-2585
- Colby DW, Prusiner SB (2011) *De novo* generation of prion strains. *Nat Rev Microbiol* **9**: 771-777
- Collins TJ (2007) ImageJ for microscopy. *BioTechniques* **43**: 25-30
- Díaz-López T, Lages-Gonzalo M, Serrano-López A, Alfonso C, Rivas G, Díaz-Orejas R, Giraldo R (2003) Structural changes in RepA, a plasmid replication initiator, upon binding to origin DNA. *J Biol Chem* **278**: 18606-18616
- Díaz-López T, Dávila-Fajardo C, Blaesing F, Lillo MP, Giraldo R (2006) Early events in the binding of the pPS10 replication protein RepA to single iteron and operator DNA sequences. *J Mol Biol* **364**: 909-920
- Doyle SM, Wickner S (2009) Hsp104 and ClpB: Protein disaggregating machines. *Trends Biochem Sci* **34**: 40-48
- Ehrnhoefer DE, Bieschke J, Boeddrich A, Herbst M, Masino L, Lurz R, Engemann S, Pastore A, Wanker EE (2008) EGCG redirects amyloidogenic polypeptides into unstructured, off-pathway oligomers. *Nat Struct Mol Biol* **15**: 558-566
- Eichner T, Radford SE (2011) A diversity of assembly mechanisms of a generic amyloid fold. *Mol Cell* **43**: 8-18
- Eisenberg D, Jucker M (2012) The amyloid state of proteins in human diseases. *Cell* **148**: 1188-1203
- Fernández-Tresguerres ME, Moreno-Díaz de la Espina S, Gasset-Rosa F, Giraldo R (2010) A DNA-promoted amyloid proteinopathy in *Escherichia coli*. *Mol Microbiol* **77**: 1456-1469
- Garrity SJ, Sivanathan V, Dong J, Lindquist S, Hochschild A (2010) Conversion of a yeast prion to an infectious form in bacteria. *Proc Natl Acad Sci USA* **107**: 10596-10601
- Gasset-Rosa F, Maté MJ, Dávila-Fajardo C, Bravo J, Giraldo R (2008) Binding of sulphonated indigo derivatives to RepA-WH1 inhibits DNA-induced protein amyloidogenesis. *Nucl Acids Res* **36**: 2249-2256
- Gatti-Lafranconi P, Natalello A, Ami D, Doglia SM, Lotti M (2011) Concepts and tools to exploit the potential of bacterial inclusion bodies in protein science and biotechnology. *FEBS J* **278**: 2408-2418
- Giraldo R, Andreu JM, Díaz-Orejas R (1998) Protein domains and conformational changes in the activation of RepA, a DNA replication initiator. *EMBO J* **17**: 4511-4526
- Giraldo R, Fernández-Tornero C, Evans PR, Díaz-Orejas R, Romero A (2003) A conformational switch between transcriptional repression and replication initiation in the RepA dimerization domain. *Nat Struct Biol* **10**: 565-571
- Giraldo R, Fernández-Tresguerres ME (2004) Twenty years of the pPS10 replicon: insights on the molecular mechanism for the activation of DNA replication in iteron-containing bacterial plasmids. *Plasmid* **52**: 69-83

- Giraldo R (2007) Defined DNA sequences promote the assembly of a bacterial protein into distinct amyloid nanostructures. *Proc Natl Acad Sci USA* **104**: 17388-17393
- Giraldo R (2010) Amyloid assemblies: Protein *Legos* at a crossroads in bottom-up synthetic biology. *ChemBioChem* **11**: 2247-2357
- Giraldo R, Moreno-Díaz de la Espina S, Fernández-Tresguerres ME, Gasset-Rosa F (2011) RepA prionoid: A synthetic amyloid proteinopathy in a minimalist host. *Prion* **5**: 60-64
- Giraldo-Suárez R, Fernández-Tresguerres ME, Díaz-Orejas R, Malki A, Kohiyama M (1993) The heat-shock DnaK protein is required for plasmid R1 replication and it is dispensable for plasmid ColE1 replication. *Nucleic Acids Res* **21**: 5495-5499
- Glover JR, Lindquist S (1998) Hsp104, Hsp70, and Hsp40: A novel chaperone system that rescues previously aggregated proteins. *Cell* **94**: 73-82
- Goloubinoff P, Mogk A, Peres-Ben Zvi A, Tomoyasu T, Bukau B (1999) Sequential mechanism of solubilization and refolding of stable protein aggregates by a bichaperone network. *Proc Nat Acad Sci USA* **96**: 13732-13737
- Halfmann R, Jarosz DF, Jones SK, Chang A, Lancaster AK, Lindquist S (2012) *Nature* **482**: 363-368
- Haslberger T, Bukau B, Mogk A (2010) Towards a unifying mechanism for ClpB/Hsp104-mediated protein disaggregation and prion propagation. *Biochem Cell Biol* **88**: 63-75
- Helmstetter CE, (1967) Rate of DNA synthesis during the division cycle of *Escherichia coli* B/r. *J Mol Biol* **24**: 417-427
- Hoshino T, Murao N, Namba T, Takehara M, Adachi H, Katsuno M, Sobue G, Matsushima T, Suzuki T, Mizushima T (2011) Suppression of Alzheimer's disease-related phenotypes by expression of heat shock protein 70 in mice. *J Neurosci* **6**: 5225-5234
- Inoue Y, Kawai-Noma S, Koike-Takeshita A, Taguchi H, Yoshida M (2011) Yeast prion protein New1 can break Sup35 amyloid fibrils into fragments in an ATP-dependent manner. *Genes Cells* **16**: 545-556
- Kaganovich D, Kopito R, Frydman J (2008) Misfolded proteins partition between two distinct quality control compartments. *Nature* **454**: 1088-1095
- Kawai-Noma S, Pack CG, Kojidani T, Asakawa H, Hiraoka Y, Kinjo M, Haraguchi T, Taguchi H, Hirata A (2010) *In vivo* evidence for the fibrillar structures of Sup35 prions in yeast cells. *J Cell Biol* **190**: 223-231
- Kim SY, Sharma S, Hoskins JR, Wickner S (2002) Interaction of the DnaK and DnaJ chaperone system with a native substrate, P1 RepA. *J Biol Chem* **277**: 44778-44783
- Knowles TPJ, White DA, Abate AR, Agresti JL, Cohen SIA, Sperling RA, De Genst EJ, Dobson CM, Weitz DA (2011) Observation of spatial propagation of amyloid assembly from single nuclei. *Proc Natl Acad Sci USA* **108**: 14746-14751
- Kumar M, Sourjik V (2012) Physical map and dynamics of the chaperone network in *Escherichia coli*. *Mol Microbiol* **84**: 736-747

- Ladiwala ARA, Dordick JS, Tessier PM (2011) Aromatic small molecules remodel toxic soluble oligomers of amyloid β through three independent pathways. *J Biol Chem* **286**: 3209-3218
- Lindner A, Madden R, Demarez A, Stewart EJ, Taddei F (2008) Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation. *Proc Natl Acad Sci USA* **105**: 3076-3081
- Luk KC, Mills IP, Trojanowski JQ, Lee VMY (2008) Interactions between Hsp70 and the hydrophobic core of α -synuclein inhibit fibril assembly. *Biochemistry* **47**: 12614-12625
- Mayer MP (2010) Gymnastics of molecular chaperones. *Mol Cell* **39**: 321-331
- Muchowski PJ, Schaffar G, Sittler A, Wanker EE, Hayer-Hartl MK, Hartl FU (2000) Hsp70 and Hsp40 chaperones can inhibit self-assembly of polyglutamine proteins into amyloid-like fibrils. *Proc Natl Acad Sci USA* **97**: 7841-7846
- Newnam GP, Wegrzyn RD, Lindquist SL, Chernoff YO (1999) Antagonistic interactions between yeast chaperones Hsp104 and Hsp70 in prion curing. *Mol Cell Biol* **19**: 1325-1333
- Newnam GP, Birchmore JL, Chernoff YO (2011) Destabilization and recovery of a yeast prion after mild heat shock. *J Mol Biol* **408**: 432-448
- Patterson KR, Ward SM, Combs B, Voss K, Kanaan NM, Morfini G, Brady ST, Gamblin TC, Binder LI (2011) Heat shock protein 70 prevents both Tau aggregation and the inhibitory effects of preexisting Tau aggregates on fast axonal transport. *Biochemistry* **50**: 10300-10310
- Prusiner SB (1998) Prions. *Proc Natl Acad Sci USA* **95**: 13363-13383
- Ratajczak E, Zietkiewicz S, Liberek K (2009) Distinct activities of *Escherichia coli* small heat shock proteins IbpA and IbpB promote efficient protein disaggregation. *J Mol Biol* **386**: 178-89
- Roberts BE, Duennwald ML, Wang H, Chung C, Lopreiato NP, Sweeny EA, Knight MN, Shorter J (2009) A synergistic small-molecule combination directly eradicates diverse prion strain structures. *Nat Chem Biol* **5**: 936-946
- Rudin L. *et al.*, *Nonlinear total variation based noise removal algorithms*, *Physica D*. 1992. 60: p. 259-268.
- Shorter J, Lindquist S (2004) Hsp104 catalyzes formation and elimination of self-replicating Sup35 prion conformers. *Science* **304**: 1793-1797
- Shorter J, Lindquist S (2006) Destruction or potentiation of different prions catalyzed by similar Hsp104 remodeling activities. *Mol Cell* **23**: 425-438
- Shorter J, Lindquist S (2008) Hsp104, Hsp70 and Hsp40 interplay regulates formation, growth and elimination of Sup35 prions. *EMBO J* **27**: 2712-2724
- Shorter J (2011) The mammalian disaggregase machinery: Hsp110 synergizes with Hsp70 and Hsp40 to catalyze protein disaggregation and reactivation in a cell-free system. *PLoS ONE* **6**: e26319
- Silva JL, Lima LM, Foguel D, Cordeiro Y (2008) Intriguing nucleic-acid-binding features of mammalian prion protein. *Trends Biochem Sci* **33**: 132-140

- Silva JL, Vieira TC, Gomes MP, Bom AP, Lima LM, Freitas MS, Ishimaru D, Cordeiro Y, Foguel D (2010) Ligand binding and hydration in protein misfolding: Insights from studies of prion and p53 tumor suppressor proteins. *Acc Chem Res* **43**: 271-279
- Soto C (2011) Prion hypothesis: The end of the controversy? *Trends Biochem Sci* **36**: 151-158
- Stewart EJ, Madden R, Paul G, Taddei F (2005) Aging and death in an organism that reproduces by morphologically symmetric division. *PLoS Biol* **3**: e45
- Tek V, Zolkiewski M (2002) Stability and interactions of the amino-terminal domain of ClpB from *Escherichia coli*. *Prot Sci* **11**: 1192-1198
- Thomas JG, Baneyx F (1998) Roles of the *Escherichia coli* small heat shock proteins IbpA and IbpB in thermal stress management: Comparison with ClpA, ClpB, and HtpG in vivo. *J Bacteriol* **180**: 5165-5172
- Toyama BH, Weissman JS (2011) Amyloid structure: Conformational diversity and consequences. *Annu Rev Biochem* **80**: 557-585
- Tuite MF, Serio TR (2010) The prion hypothesis: From biological anomaly to basic regulatory mechanism. *Nat Rev Mol Cell Biol* **11**: 823-833
- Tyedmers J, Mogk A, Bukau B (2010a) Cellular strategies for controlling protein aggregation. *Nat Rev Mol Cell Biol* **11**: 777-788
- Tyedmers J, Treusch S, Dong J, McCaffery JM, Bevis B, Lindquist, S (2010b) Prion induction involves an ancient system for the sequestration of aggregated proteins and heritable changes in prion fragmentation. *Proc Natl Acad Sci USA* **107**: 8633-8638
- Vyawahare S, Griffiths AD, Merten CA (2010) Miniaturization and parallelization of biological and chemical assays in microfluidic devices. *Chem Biol* **17**: 1052-1065
- Wang L, Maji SK, Sawaya MR, Eisenberg D, Riek R (2008) Bacterial inclusion bodies contain amyloid-like structure. *PLoS Biol* **6**: e195
- Wang P, Robert L, Pelletier J, Dang WL, Taddei F, Wright A, Jun S (2010) Robust growth of *Escherichia coli*. *Curr Biol* **20**: 1099-1103
- Weissmann C, Li J, Mahal SP, Browning S (2011) Prions on the move. *EMBO Rep* **12**: 1109-1117
- Wickner RB, Edskes HK, Shewmaker F, Nakayashiki T (2007) Prions of fungi: Inherited structures and biological roles. *Nat Rev Microbiol* **5**: 611-618
- Wickner S, Skowrya D, Hoskins J, McKenney K (1992) DnaJ, DnaK, and GrpE heat shock proteins are required in oriP1 DNA replication solely at the RepA monomerization step. *Proc Natl Acad Sci USA* **89**: 10345-10349
- Winkler J, Seybert A, König L, Prugnaller S, Haselmann U, Sourjik V, Weiss M, Frangakis AS, Mogk A, Bukau B (2010) Quantitative and spatio-temporal features of protein aggregation in *Escherichia coli* and consequences on protein quality control and cellular ageing. *EMBO J* **29**: 910-923

- Woods LA, Platt GW, Hellewell AL, Hewitt EW, Homans SW, Ashcroft AE, Radford SE (2011) Ligand binding to distinct states diverts aggregation of an amyloid-forming protein, *Nat Chem Biol* **7**: 730-739
- Wu C, Wang Z, Lei H, Duan Y, Bowers MT, Shea JE (2008) The binding of thioflavin T and its neutral analog BTA-1 to protofibrils of the Alzheimer's disease A β (16-22) peptide probed by molecular dynamics simulations. *J Mol Biol* **384**: 718-729
- Yin H, Marshall D (2012) Microfluidics for single cell analysis. *Curr Op Biotechnol* **23**: 110-119
- Zietkiewicz S, Lewandowska A, Stocki P, Liberek K (2006) Hsp70 chaperone machine remodels protein aggregates at the initial step of Hsp70-Hsp100-dependent disaggregation. *J Biol Chem* **281**: 7022-7029

Figure Legends

Figure 1. Selected time-elapsed frames showing *E. coli* cells in five of the channels (vertically aligned) in the microfluidic device. Mother cells occupy the bottom of each microchannel, whereas fresh medium flows horizontally at the top. **(A)** Fluorescence emission of the RepA-WH1-mCherry (mRFP) prionoid particles indicates the presence of two types of protein aggregates: either globular (labelled with red stars below the mother cells) or worm-shaped (yellow stars). For any given channel, globular and wormlike aggregates can switch (double coloured stars) with low frequencies. **(B)** A similar composition of microchannels with bacteria carrying the IB control RepA-WH1(Δ N37)-mCherry. Cells in all channels carry a single subpolar globular protein aggregate (red stars). **(C)** IbpA-GFP, a marker of IBs, colocalizes with some of the subpolar aggregates, but not with the globular particles when located in other positions or with the elongated aggregates. For a full account of these experiments, see Supplementary Movies 1 (A), 2 (B) and 3 (C).

Figure 2. The amyloid-specific fluorophore BTA-1 targets the RepA-WH1 intracellular aggregates. **(A)** DIC (*left panels*), and red (*centre*) and cyan (*right*) fluorescence images of single cells bearing the globular (*top*) and the wormlike (*middle*; arrows) RepA-WH1(A31V) aggregates, or the RepA-WH1(Δ N37) IBs (*bottom*). **(B)** Box plot of the ratios between the fluorescence intensities of the cyan (BTA-1) and red (mCherry) channels for each type of RepA-WH1 aggregate. Particles were picked-up from 25 images in these two channels (16 bit greyscale mode) acquired under constant optical settings (see Materials and methods). Calculations were performed using ImageJ (v. 1.43u), using as a probe an elliptical sector (comprising 166 pixels) that was centred on single, neat particles (N= 75) of each type. In the plot, dots mark the mean values, bars span the minimal and maximum intervals of dispersion, boxes comprise 25-75% of the data points and the thick horizontal lines indicate the median values. One-way ANOVA statistical analysis (F= 98.31, p < 0.0001) followed by Tukey's

pairwise difference test indicate that each kind of aggregate belongs to a distinct class in terms of BTA-1 staining.

Figure 3. Microfluidic channels, in a display analogue to that in Figure 1, showing bacteria expressing RepA-WH1-mRFP in which it was tested the effect of either the DnaK inhibitor myricetin (**A**), the overexpression of the chaperones DnaK (Hsp70) (**B**) or ClpB (Hsp104) (**C**). While DnaK seems to have an effect on the interconversion between the globular (red stars) and the wormlike (yellow stars) aggregated species of RepA-WH1, because the former accumulate upon DnaK inhibition (A) and the latter under DnaK overexpression (B), ClpB has no apparent role, since both species still coexists along the whole time course of the experiment. For a full account of these assays, see Supplementary Movies 4 (A), 5 (B) and 6 (C).

Figure 4. Expression of the amyloidogenic RepA-WH1-mRFP in the *E. coli* WT strain MC4100 or in isogenic mutants for the two chaperones under study: *dnaK756^{ts}* (JGT20) or $\Delta clpB$ (JGT3). Representative views from culture aliquots taken after IPTG induction are displayed superposed to the DIC and red fluorescence images. Right column shows (red fluorescence emission) a 6-fold magnification of the sectors boxed in yellow. For further details, see Supplementary Figure S2A.

Figure 5. Expression of the amyloidogenic RepA-WH1-mRFP in the chaperone-deficient strains (Figure 4) complemented by the expression of either DnaK or ClpB from a compatible plasmid (Supplementary Figure S2B). IPTG and Ara inducers were added, in either order (+ind. 1 \rightarrow +ind. 2), within a 30 min interval, which result in prior overexpression of the prionoid (Amyl \rightarrow DnaK/ClpB) or a chaperone (DnaK/ClpB \rightarrow Amyl), respectively. The Hsp70 chaperone partially disorganizes the globular prionoid particles into more diffused, elongated aggregates (arrows).

Figure 6. Intracellular distribution of RepA-WH1-mRFP and the DnaK/ClpB chaperones explored by epifluorescence and confocal microscopies. (**A**) Complementation of *dnaK756* and $\Delta clpB$ strains was performed with the DnaK and ClpB chaperones, respectively, by consecutive Ara \rightarrow IPTG induction (Figure 5). Endogenous mCherry (red) and extrinsic, immuno-labelled (green) fluorescence localize the prionoid and chaperone molecules, respectively. Nucleoid DNA staining with DAPI was also performed for confocal microscopy. (**B**) Color (RGB) channels, plotted along longitudinal sections (yellow dashed lines in A) through chaperone-expressing cells, highlight the mutually exclusive distribution of aggregates (red) and nucleoid (blue), and the differential location of DnaK and ClpB chaperones.

Figure 7. Intracellular distribution of DnaK and ClpB chaperones and RepA-WH1-mRFP revealed by iEM. Thin sections of *E. coli* cells bearing the RepA-WH1 aggregates and complemented with

DnaK (**A**) or ClpB (**B**) were incubated with specific antibodies against the chaperones and then localized with Au-conjugated antibodies (dots). N: nucleoid. The electron-dense areas are RepA-WH1 aggregates (Fernández-Tresguerres *et al*, 2010). The boundaries of the elongated wormlike RepA-WH1 aggregates, which concentrate most of the anti-DnaK labelling, are encircled in red whereas smaller aggregated particles, also bound by the chaperone, are highlighted in green. The large globular RepA-WH1 aggregates do not exhibit significant DnaK binding. ClpB binds evenly across the whole cytoplasm.

Figure 8. RepA-WH1 amyloid fibres assembled *in vitro* (Giraldo, 2007) were incubated with purified DnaK and ATP, resulting in partial disorganization of the fibres by the action of the Hsp70 chaperone. Single Au-conjugated anti-WH1 or anti-DnaK labels localize on small, spherical particles detached from the fibre scaffold. Due to the dimension of the RepA-WH1 molecules (Giraldo *et al*, 2003) each of such complexes must be an oligomer of this protein.

Figure 9. Characterization by SDD-AGE of the presence of RepA-WH1-mRFP oligomers in *E. coli* and their dependence on DnaK/ClpB chaperones. Whole cell extracts of the WT and DnaK/ClpB-deficient bacterial strains, complemented or not by the expression of these chaperones (Figures 4-6), were analyzed by electrophoresis in non-denaturing, detergent-containing agarose gels. The distribution of RepA-WH1-mRFP (**A**), DnaK (**B**) and ClpB (**C**) was revealed by means of Western blotting with specific antibodies. The presence of detergent-resistant RepA-WH1 oligomers was strictly dependent on the Hsp70 chaperone, but not on the Hsp104 disaggregase.

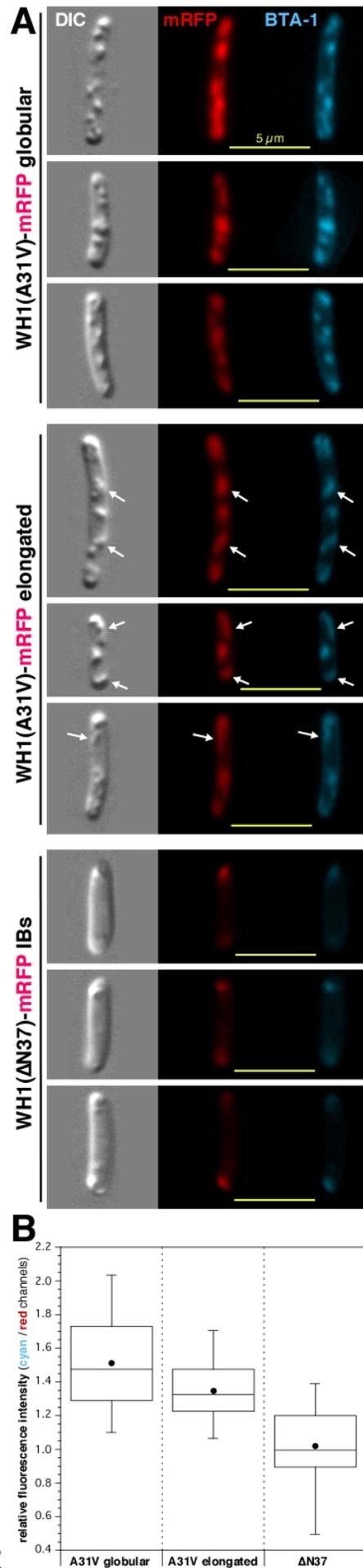
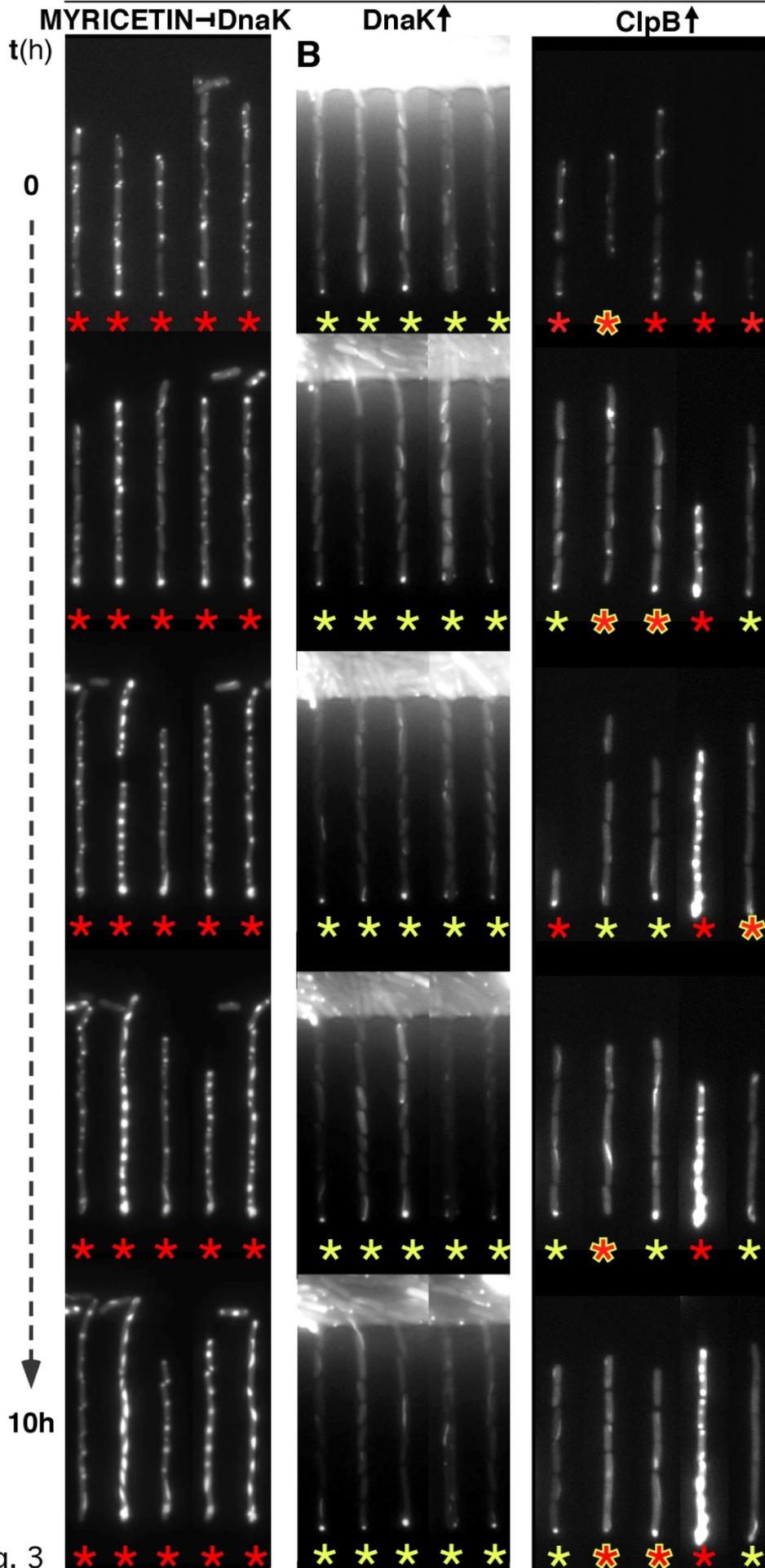


Fig. 2

WH1(A31V)-mRFP



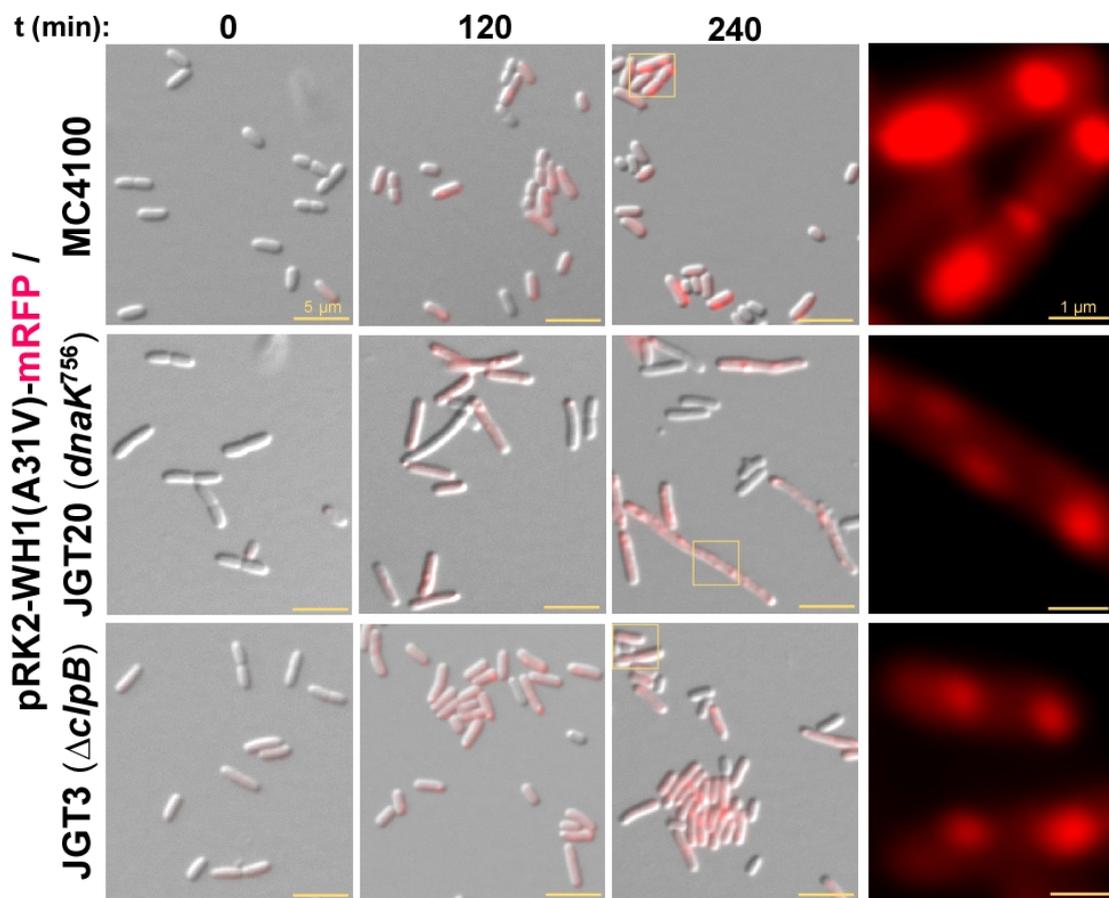


Fig. 4

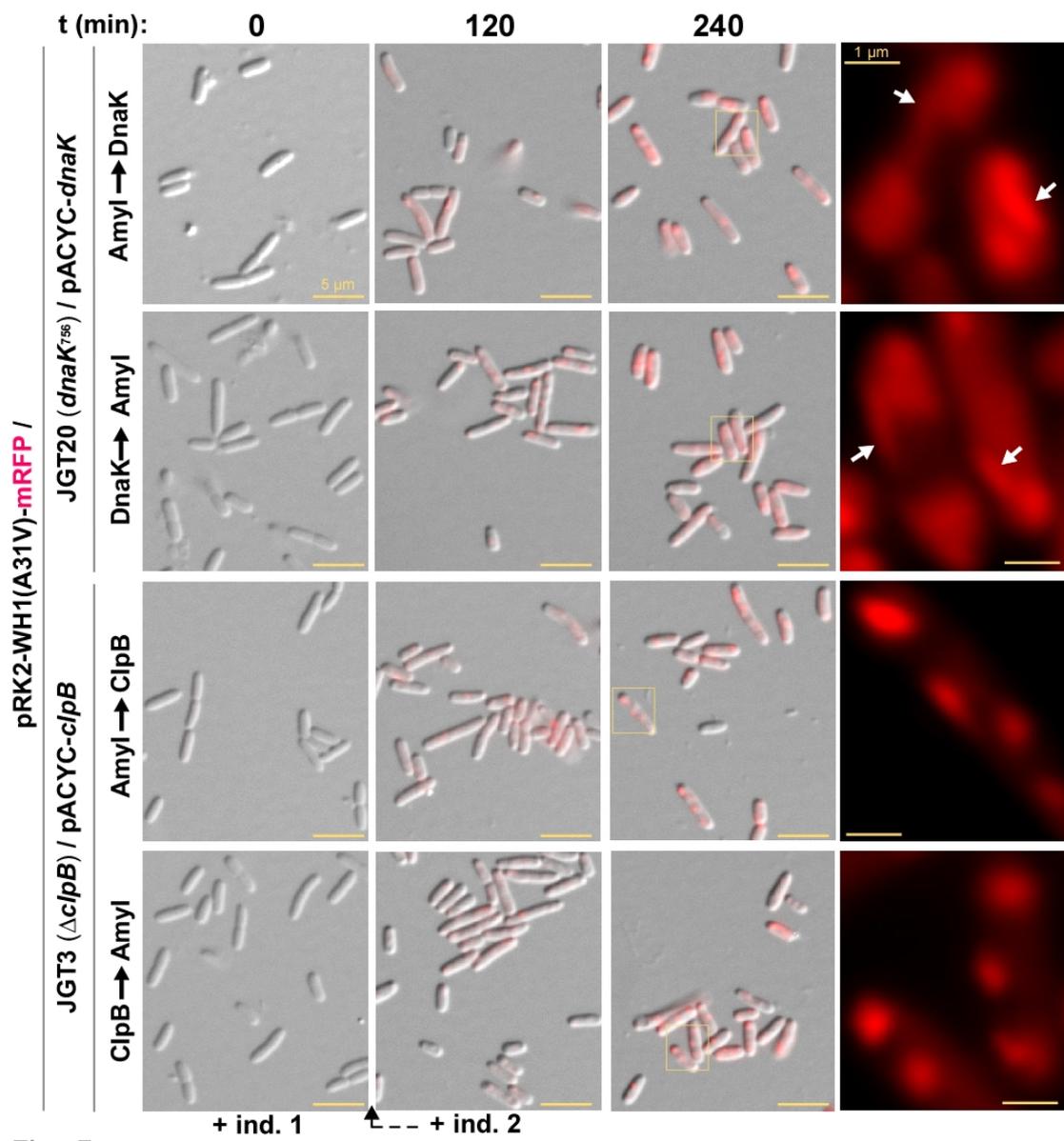


Fig. 5

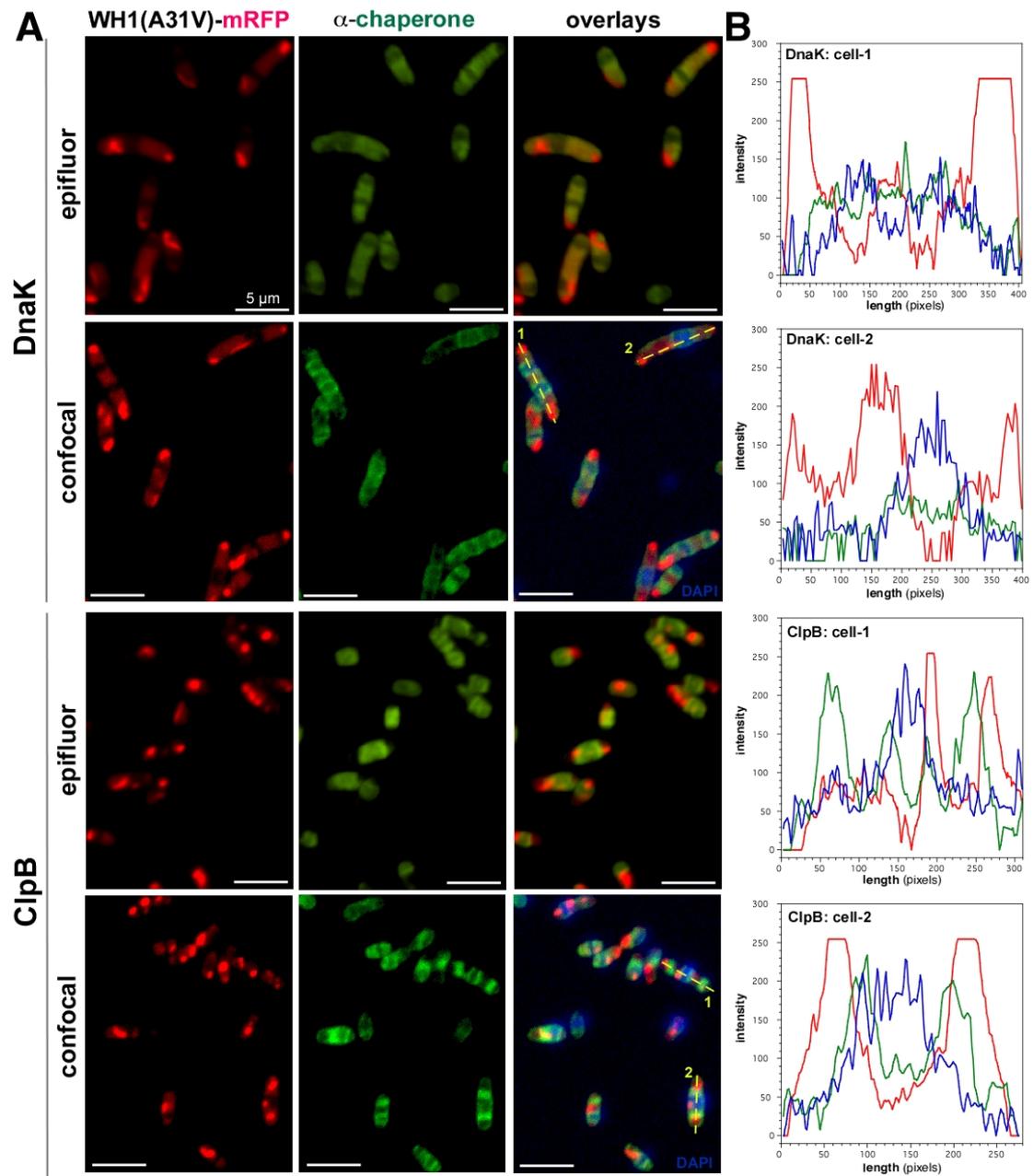


Fig. 6

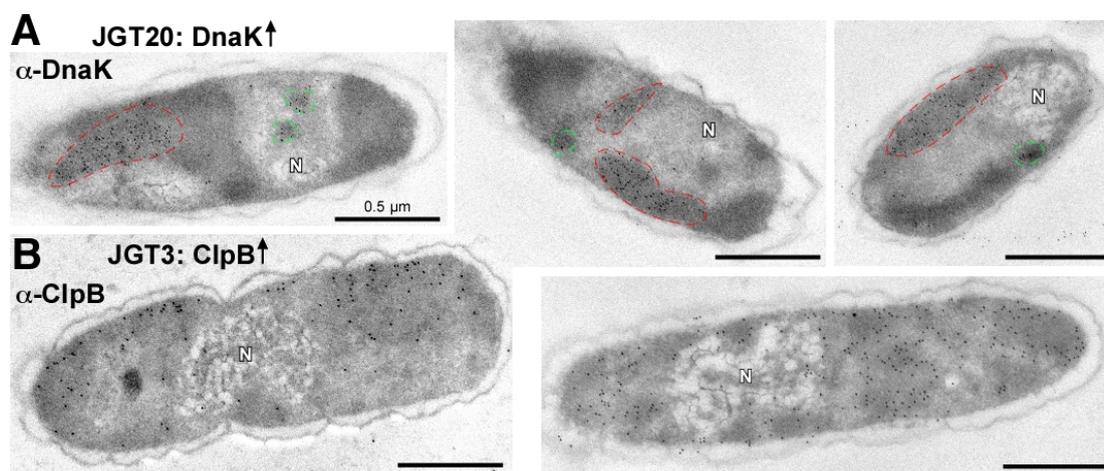


Fig. 7

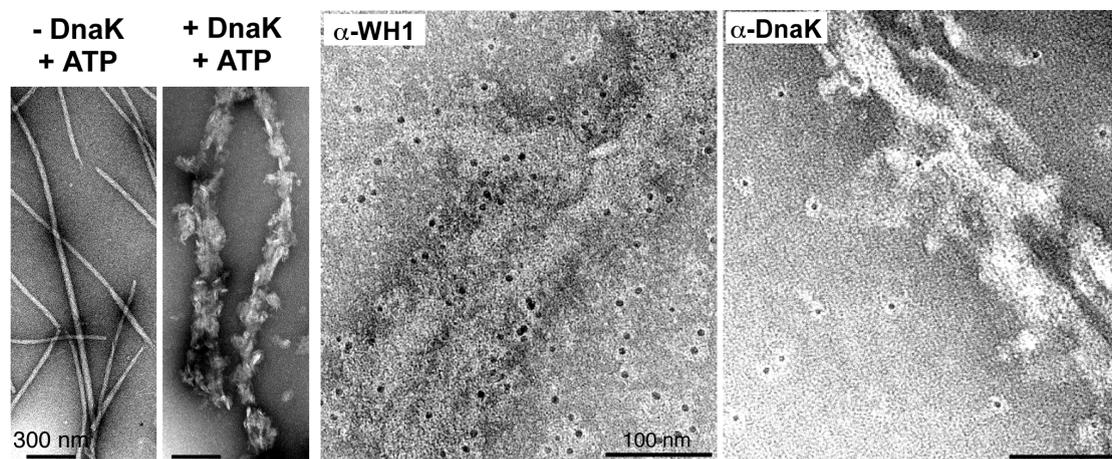


Fig. 8

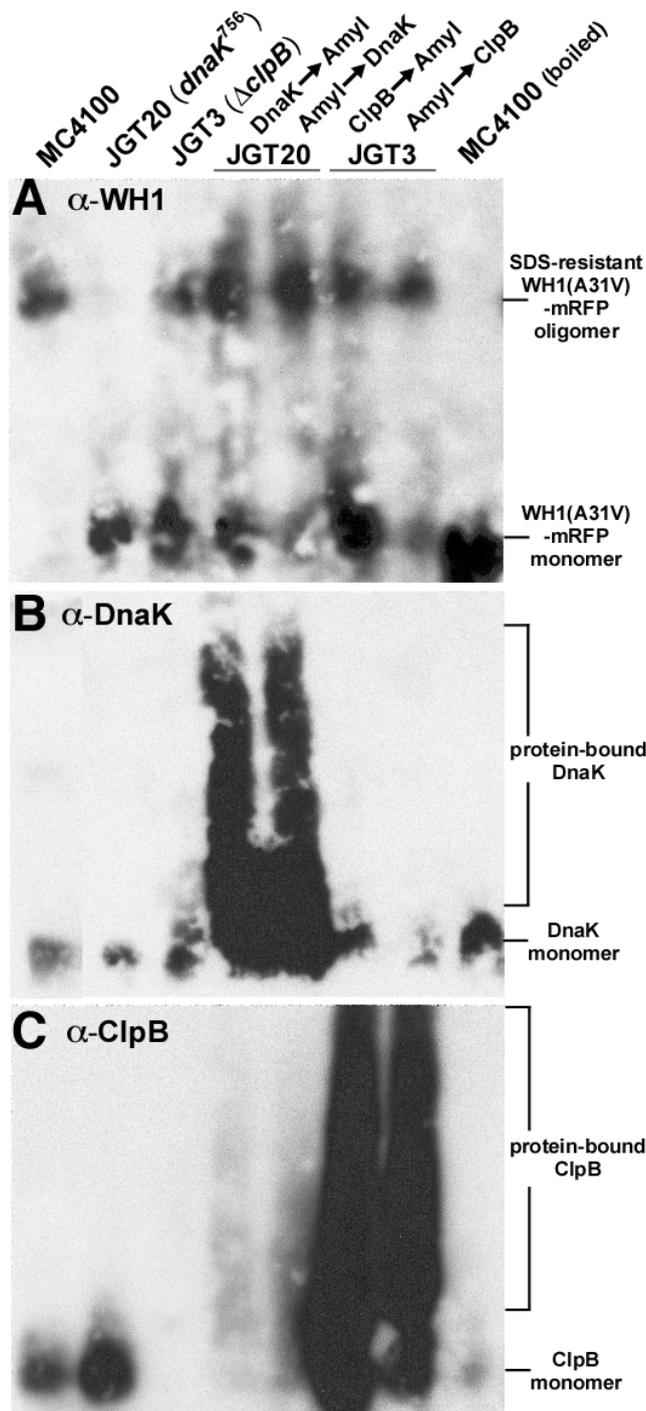


Fig. 9

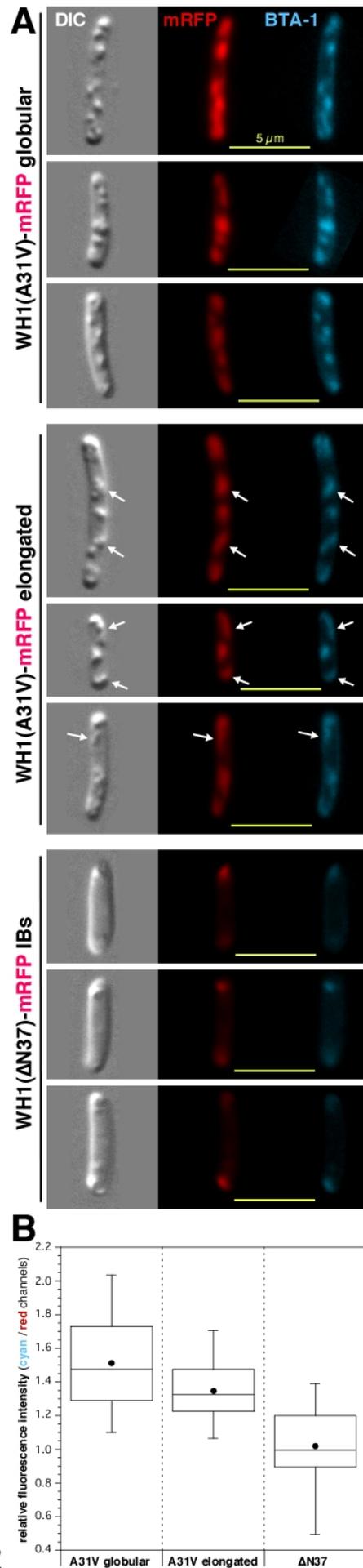
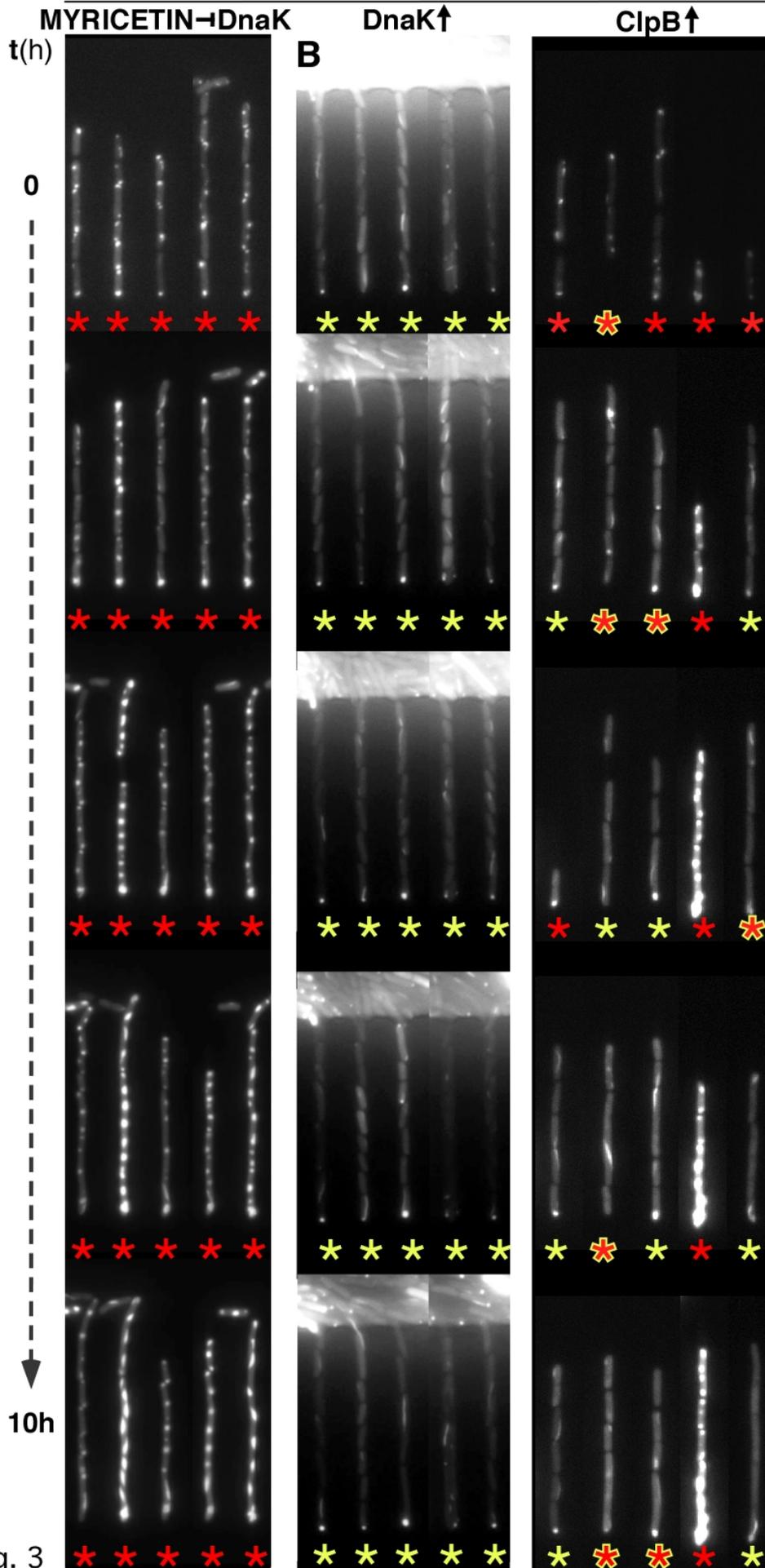


Fig. 2

WH1(A31V)-mRFP



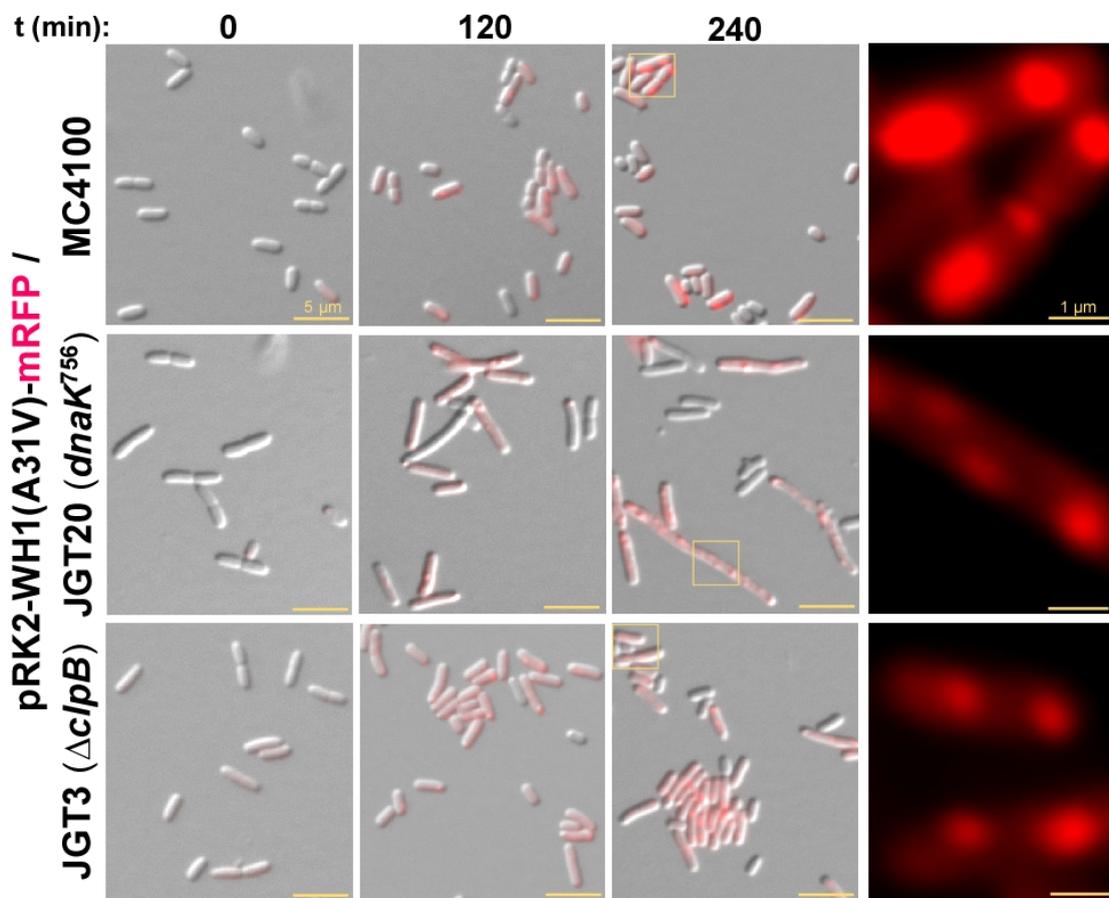


Fig. 4

Discussion

Dans ce chapitre, je discuterai des résultats obtenus sur la dynamique des agrégats protéiques de type corps d'inclusion en conditions non stressantes (article 1) et de type prionoïde (article 2) chez la bactérie *E. coli*. La première étude a combiné une double approche théorique de modélisation et expérimentale basée sur la microscopie time-lapse. La seconde étude s'appuie sur la création d'un système synthétique permettant d'étudier une protéinopathie amyloïde chez la bactérie *E. coli* au moyen de la microscopie microfluidique. J'aborderai dans cette discussion, la dynamique spatiale des agrégats protéiques ainsi que leur toxicité et lien avec le vieillissement. L'agrégation protéique est un candidat potentiel comme facteur de vieillissement. En effet, les agrégats protéiques sont néfastes pour la cellule, s'accumulent avec l'âge et sont distribués asymétriquement ce qui permet une distinction entre les parents et les descendants, caractéristique définissant le vieillissement. Plusieurs études ont tenté d'identifier les paramètres tels que le nombre, la taille ou encore la localisation cellulaire des agrégats protéiques afin de déterminer les causes de leur toxicité et le mécanisme entraînant la mort cellulaire. Pourquoi les cellules ne sont-elles pas capables de faire appel automatiquement aux protéines chaperonnes afin de prévenir l'agrégation protéique avec le vieillissement ? La cellule dispose d'un système contrôle qualité des protéines qui permet de renverser l'agrégation protéique et on pourrait s'attendre à ce que ce système soit activé constitutivement en présence d'agrégats. Certaines modifications oxydatives sont réversibles mais la réaction de carbonylation par exemple, est irréversible et peut entraîner la formation d'agrégats insolubles qui sont très difficile à dégrader [Nystrom, 2005]. La cellule doit ainsi faire face à ces agrégats et leur accumulation avec l'âge.

Sommaire

I	Dynamique des agrégats protéiques	144
I.1	Résumé des résultats obtenus	145
I.2	Localisation et mécanisme passif	145
I.3	Ségrégation des agrégats protéiques	149
II	Toxicité des agrégats protéiques	152
II.1	Résumé des résultats obtenus	152
II.2	Influence de la forme des agrégats protéiques sur la toxicité	153
II.3	Influence de la taille des agrégats protéiques sur la toxicité	154
II.4	<i>Escherichia coli</i> , modèle d'étude des prions	154
III	Conclusion	155

I Dynamique des agrégats protéiques

La cellule bactérienne n'est pas aussi simple qu'on aurait pu le penser. La preuve en est par la mise en évidence d'un cytosquelette permettant de contrôler la polarité [Gitai et al., 2004] ou la ségrégation chromosomique [Kruse et al., 2006] ou encore la synthèse de la membrane [Jones et al., 2001].

Les techniques récentes de suivi de molécules ont permis de montrer que de nombreuses molécules ont des localisations spécifiques. Nous noterons que la récente publication de Landgraf [Landgraf et al., 2012] mettant en cause l'impact des marqueurs fluorescents sur la localisation et agrégation de protéines *in vivo*, n'impacte en rien les résultats que nous obtenons. En effet, en 2008, Lindner et ses collaborateurs ont montré que le marquage de la protéine chaperonne IbpA avec la protéine fluorescente jaune n'avait pas de conséquences sur la localisation et agrégation via la protéine chaperonne IbpA. Ils observent une localisation identique des agrégats observés en fluorescence avec les corps d'inclusion visibles en contraste de phase. D'autre part, les résultats obtenus par Landgraf avec la protéine chaperonne IbpA ne sont pas si différents de ceux que nous obtenons quand à la localisation et agrégation. Leurs résultats révèlent également des agrégats localisés au pôle et au centre de la cellule. Cependant, pour les expériences à venir, il sera important d'effectuer d'autres contrôles préliminaires et de tester le marqueur fluorescent GFP3-mutant, suggéré par Landgraf comme le moins néfaste, pour les nouveaux marquages, qui semble minimiser les impacts sur la localisation et l'agrégation.

De nombreux mécanismes contribuent au transport intracellulaire et à la localisation des macromolécules chez la bactérie et suscitent des questions telles que : existe-t-il des mécanismes de transports actifs ? Est ce que la diffusion à elle seule peut expliquer la localisation des molécules ? Est ce que les filaments du cytosquelette sont polarisés ?

Le cytosquelette jouent un rôle clé dans la polarité chez les eucaryotes [Etienne-Manneville and Hall, 2002] [Drubin and Nelson, 1996]. Chez la bactérie *C. crescentus*, plusieurs protéines du cytosquelette ont été identifiées telles que FtsZ, homologue de la tubuline [Quardokus et al., 1996] [Lowe and Amos, 1998], MreB, protéine homologue de l'actine [Figge et al., 2004] [van den Ent et al., 2001] ou encore CreS, homologue de filament intermédiaire [Ausmees et al., 2003]. La protéine MreB est impliquée également dans la forme des organismes tels que *E. coli*, *B. subtilis* ou *C. crescentus* [Figge et al., 2004] [Jones et al., 2001] [Wachi et al., 1989].

La ségrégation chromosomique chez les organismes unicellulaires semble nécessiter des mécanismes actifs et l'implication du cytosquelette. Plusieurs molécules seraient impliquées telles que MreB [Gitai et al., 2004], qui contient les informations concernant la polarité afin de déterminer la localisation subcellulaire des protéines chez la bactérie *C. crescentus* ou le système ParA/ParB dans lequel ParB se fixe à une région spécifique du chromosome près de l'origine de réplication et ensuite ParA filamente tout en étant ancré à un des pôles. L'origine de réplication est ainsi tirée vers le pôle opposé [Ebersbach et al., 2006].

La géométrie de la cellule peut aussi contribuer à la localisation de protéines dans des régions particulières telles que les pôles. La cardiolipine est connue pour former naturellement des microdomaines au pôle chez *E. coli* et colocalise avec la protéine ProP qui est un transporteur sensible à l'osmolalité [Mileykovskaya, 2007].

Malgré la présence de transports actifs et un cytoplasme défini comme environnement très encombré chez la bactérie *E. coli*, les premières mesures de diffusion effectuées par Elowitz

[Elowitz et al., 1999] démontrent que la diffusion est possible malgré l'encombrement moléculaire afin que les molécules atteignent leur destination finale.

I.1 Résumé des résultats obtenus

Dans la première publication, nous avons décrit la dynamique des agrégats protéiques chez la bactérie *E. coli* en conditions non-stressantes.

Au moyen de la microscopie time-lapse et en marquant la protéine chaperonne IbpA avec la protéine fluorescente jaune (YFP), nous avons suivi le mouvement des agrégats protéiques dans des cellules uniques. Nous avons démontré que le mouvement des agrégats protéiques localisés au pôle est cohérent avec un mouvement Brownien, le coefficient de diffusion étant dépendant de la taille de la molécule en accord avec l'équation de Stokes-Einstein. Le MSD (Mean Square Displacement) indique un confinement localisant les agrégats protéiques dans les régions dépourvues d'ADN.

Afin de confirmer cet effet de volume exclu et la distribution spatiale des agrégats, nous avons développé un modèle 3d basé individus de diffusion-agrégation. Ce modèle reproduit la localisation des agrégats protéiques dans le cytoplasme et confirme que la diffusion est un mécanisme suffisant pour expliquer le mouvement des agrégats protéiques.

Associant les données expérimentales à des simulations avec estimations de paramètres (fitting simulations), nous avons également pu établir une corrélation entre la fluorescence initiale de l'agrégat, sa taille et son coefficient de diffusion. Le coefficient de diffusion décroît ainsi avec l'inverse du rayon.

Nos résultats démontrent ainsi par une étude à la fois robuste, qualitative et quantitative que chez la bactérie *E. coli*, la diffusion gouverne la localisation polaire des agrégats protéiques en conditions non stressantes et que leur coefficient de diffusion est dépendant de leur taille.

I.2 Localisation et mécanisme passif

La mobilité des agrégats protéiques est très controversée dans la littérature chez les organismes unicellulaires [Rokney et al., 2009] [Winkler et al., 2010] [Zhou et al., 2011] [Liu et al., 2011] mais aucune étude quantitative n'a été réalisée *in vivo*.

Deux études qualitatives ont été réalisées chez la bactérie *E. coli* en conditions de choc thermique [Rokney et al., 2009] [Winkler et al., 2010] et donc stressantes pour la cellule ce qui entraîne la formation de nombreux agrégats protéiques [Rokney et al., 2009]. En effet, le système contrôle qualité des protéines est saturé et n'est pas capable de réparer toutes les protéines mal repliées à cause du choc thermique, ce qui entraîne l'agrégation.

En 2009, Rokney et ses collaborateurs [Rokney et al., 2009] ont montré que la localisation polaire des agrégats protéiques et leur formation seraient ATP dépendantes et nécessiteraient ainsi un mécanisme actif. Ils montrent ainsi qu'en conditions de stress thermique, de multiples foci sont observés dans un premier temps et sont très mobiles à travers le cytoplasme (Figure 4.1, page 146). Dans un second temps, les petits foci se regroupent pour former des agrégats protéiques de tailles plus importantes, nommés LPA (Large Polar Aggregate). En revanche, une fois au pôle, les agrégats protéiques deviendraient stationnaires et aucune énergie ne serait nécessaire à leur maintien au pôle. Ces résultats sont plutôt contradictoires avec notre étude quantitative même si les conditions expérimentales sont différentes (conditions stressantes *versus* conditions

non-stressantes). Les agrégats localisés au pôle ne sont pas stationnaires malgré un coefficient de diffusion très faible des agrégats de grande taille en accord avec la théorie de la diffusion. En revanche, nous nous accordons sur l'absence de mécanisme actif pour le mouvement des agrégats au pôle.

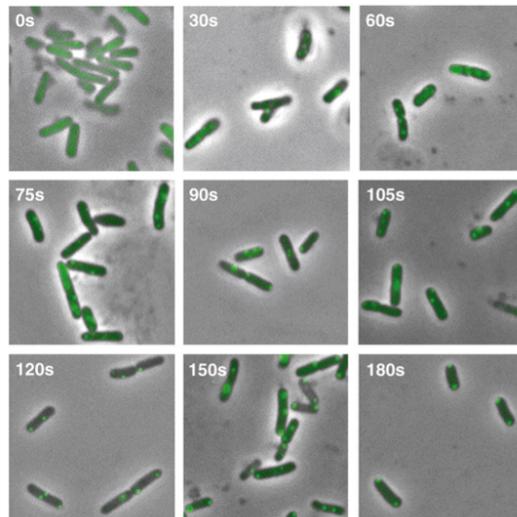


FIGURE 4.1 – **Différentes étapes de formation des agrégats protéiques et leur localisation.** On notera que les images sont indépendantes les unes des autres et ne sont pas issues d'une même séquence et même bactéries. Les images 0, 30 et 60s sont représentatives de la 1ère étape avec l'apparition de nombreux foci dans toute la cellule. Dans les images 75 à 150s, les petits foci se regroupent pour former des agrégats plus larges. Enfin, à 180s, la cellule contient 1 ou 2 agrégats protéiques localisés au pôle. (Figure extraite de [Rokney et al., 2009]).

En 2010, également dans des conditions expérimentales de choc thermique, Winkler et ses collaborateurs [Winkler et al., 2010] montrent chez la bactérie *E. coli* que le mouvement des agrégats serait un mouvement diffusif jusqu'au pôle, régions dépourvues d'ADN bactérien.

Les résultats obtenus par les équipes de Rokney et Winkler s'opposent malgré des conditions similaires de choc thermique. Notre étude en conditions non stressantes rejoint les résultats obtenus par Winkler et l'absence de mécanisme actif pour le déplacement des agrégats. L'étude de Rokney entraîne une question clé : est-il possible qu'en conditions stressantes, la cellule utilise des mécanismes actifs afin de ségréger les dommages protéiques et ainsi faciliter le travail du système contrôle qualité ?

Notre étude et celle de Winkler s'accordent sur le fait que l'occlusion due à la présence des nucléoides serait ainsi la principale force permettant de diriger les agrégats vers les pôles.

Afin de comprendre si l'effet de volume exclu peut être une force suffisante pour expliquer la localisation des agrégats protéiques, Saberi *et al.* [Saberi and Emberly, 2010] ont développé un modèle théorique en utilisant comme modèle la protéine PopZ localisée au pôle chez la bactérie *C. crescentus*. Ils confirment ainsi que la présence de régions dépourvues d'ADN, associée à la présence des nucléoi (ADN bactérien) très encombrés, pourrait être le mécanisme permettant

la localisation polaire. Ils montrent que pour un volume fixe d'ADN, les protéines restent diffuses à faible concentration. En revanche, à haute concentration, les protéines peuvent former un noyau à un des pôles par la création d'une densité élevée de protéines mal repliées dans une zone dépourvue d'ADN. En modifiant le volume occupé par l'ADN de 10 à 15%, les patterns de localisation observés sont différents. Un volume d'ADN de 15% favorise un pattern bipolaire avec formation de noyau (agrégat protéique) à chaque pôle. D'un point de vue expérimental, le pattern bipolaire est observé au moment de la division chez la bactérie *C. crescentus* tandis que l'on observe plus fréquemment un pattern unipolaire chez la bactérie *E. coli*. ceci est en adéquation avec les volumes cellulaires des deux bactéries. En effet le volume occupé par leur ADN est très similaire. En revanche, le volume cellulaire de *C. crescentus* est inférieur à celui d'*E. coli* et cela correspondrait dans le modèle de Saberi à une augmentation du volume occupé par l'ADN avec un volume cellulaire identique. Ainsi, l'oligomérisation des protéines à un pôle dépend de la concentration en protéines mal repliées mais aussi de la géométrie de la cellule et du volume bactérien. Dans notre modèle, l'absence d'agrégation ne permet pas de reproduire les localisations spécifiques des agrégats protéiques observées *in vivo* chez *E. coli* ainsi que dans le modèle de Sabéri quelles que soient les concentrations initiales protéiques.

D'autres études ont été réalisées sur le mouvement des agrégats chez la levure. Chez la levure *S. cerevisiae*, le mécanisme permettant de retenir les agrégats dans la cellule mère est sujet à discussion. En 2008, Shcheprova [Shcheprova et al., 2008] met en évidence l'existence d'une barrière de diffusion formée par l'enveloppe nucléaire et empêchant ainsi le transfert des agrégats protéiques existants dans la cellule mère à la cellule fille. De plus, très récemment, Zhou et ses collaborateurs [Zhou et al., 2011] concluent également en l'absence d'un transport rétrograde via l'actine des agrégats de la cellule fille vers la cellule mère, en conditions de stress thermique. Leur analyse des trajectoires des agrégats démontrent un mouvement diffusif (Brownien) avec confinement (Figure 4.2, page 147). Ils proposent ainsi que l'étroitesse du passage entre la cellule mère et la cellule fille suffise à la rétention des agrégats protéiques accumulés dans la cellule mère.

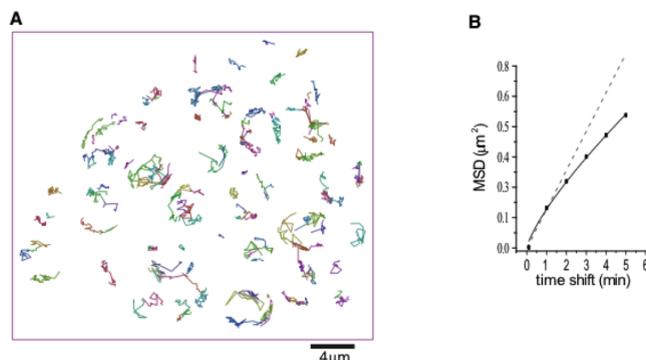


FIGURE 4.2 – **Mouvement diffusif des agrégats protéiques chez la levure *S. cerevisiae*.** **A)** Trajectoires d'agrégats protéiques suivi durant 1h en microscopie time-lapse. **B)** MSD en fonction du temps indique que le mouvement des agrégats suit une marche aléatoire (courbe en pointillés) avec un léger confinement (courbe avec marqueurs) (Figure extraite de [Zhou et al., 2011]).

Cependant, une correspondance de Liu *et al.* [Liu et al., 2011] en réponse à l'étude de Zhou *et al.* publiée dans le même volume de la revue Cell, tend à monter l'existence d'un mouvement rétrograde. Il semblerait que dans un premier temps, la cellule fille hérite des agrégats protéiques et effectue ensuite un nettoyage par un transport rétrograde vers la cellule mère. La répartition des agrégats protéiques impliquerait les filaments d'actine. Liu et ses collaborateurs suggèrent de faire une analyse plus approfondie quand à la diffusion des agrégats en conditions naturelles tout en maintenant le cytosquelette actif. Ainsi chez la levure *S. cerevisiae*, l'existence d'un mécanisme actif n'est pas exclu et nécessite des études plus approfondies.

Chez la levure *S. pombe*, Coelho et ses collaborateurs [Coelho, 2012] ont montré que le mouvement des agrégats était diffusif et dépendant de la taille des particules. Les agrégats de faible intensité bougent plus vite que les agrégats de haute intensité, ce qui indique que leur taille doit être plus petite en accord avec les équations de la théorie de la diffusion (Figure 4.3, page 148). De plus, ils évoquent également le fait que les agrégats sont absents du noyau qui doit avoir également un effet d'exclusion comme chez la bactérie *E. coli*. Le nucleus agirait ainsi comme barrière face à la division et créerait des compartiments localisant les agrégats protéiques dans les régions dépourvues d'ADN.

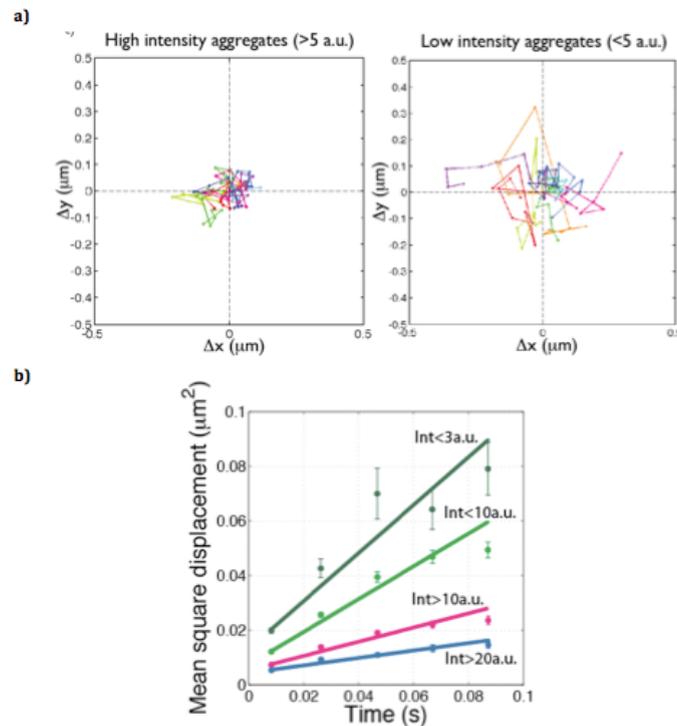


FIGURE 4.3 – Diffusion des agrégats dans le cytoplasme de *S. cerevisiae*. **a)** Trajectoires individuelles d'agrégats de haute intensité (gauche) et de faible intensité (droite). **b)** MSD en fonction du temps. Les agrégats ont été classés en fonction de leur intensité. (Figure extraite de [Coelho, 2012])

Est ce que l'exclusion due à l'encombrement moléculaire dans les nucléoi ou la géométrie cellulaire, serait un paramètre conservé chez les organismes unicellulaires et permettrait ainsi une distribution asymétrique des agrégats ? L'ensemble des études publiées dans la littérature, ajouté à nos résultats, semble indiquer que la diffusion régirait le mouvement des agrégats protéiques chez les organismes unicellulaires par un effet de volume exclu, les agrégats sont ainsi localisés dans les régions dépourvues d'ADN ou retenus dans une cellule préférentiellement par une barrière physique. Il y a une "pseudo-compartmentalisation" qui est créée permettant une distribution asymétrique des dommages.

I.3 Ségrégation des agrégats protéiques

Afin de faire face aux dommages, la cellule dispose de deux grands mécanismes : la réparation de ces dommages ou lorsque la cellule ne peut plus faire face aux dommages car leur production est supérieure aux capacités de réparation de la cellule, leur séquestration dans des agrégats protéiques. En outre, chez la bactérie *E. coli* ou encore la levure *S. cerevisiae*, on observe une distribution asymétrique [Lindner et al., 2008] des dommages entre les cellules filles. Cette ségrégation asymétrique a pour conséquence la génération d'une cellule indemne de dommages et d'une cellule vieillissante et a lieu pour les organismes dont la division est symétrique ou asymétrique.

Plusieurs paramètres semblent influencer l'agrégation protéique. En effet, l'âge [Maisonneuve et al., 2008], un choc thermique [Rokney et al., 2009] ou encore le milieu [Kwiatkowska et al., 2008], contribuent à augmenter de manière significative le nombre de protéines endommagées et la formation d'agrégats. Ainsi, en phase stationnaire (la bactérie *E. coli* entre en phase stationnaire quand tous les nutriments essentiels du milieu sont épuisés), la formation des agrégats protéiques augmente et dépend de la disponibilité en oxygène et glucose dans le milieu et donc du métabolisme [Kwiatkowska et al., 2008].

Nous avons vu que chez la bactérie *E. coli*, la ségrégation stochastique et l'encombrement moléculaire du à la présence des nucléoides étaient les éléments clés conduisant à la ségrégation asymétrique des dommages. Cependant, est ce que tous les organismes unicellulaires distribuent symétriquement leurs dommages ou existe-t-il d'autres stratégies ?

La polarisation des dommages est une stratégie permettant de restreindre la sénescence à une des deux cellules filles lors de la division. Elle implique une distribution asymétrique des dommages mais aussi une directionnalité.

I.3.1 Ségrégation asymétrique

Chez la bactérie *E. coli* [Lindner et al., 2008], ou la levure *S. cerevisiae* [Zhou et al., 2011], on assiste à une ségrégation asymétrique des dommages. Chez *E. coli*, la cellule héritant de l'agrégat protéique montre des signes de vieillissement répliatif sur les neuf premières générations, à savoir que sa croissance diminue par rapport à la cellule fille n'ayant pas hérité de l'agrégat protéique [Stewart et al., 2005]. Chez *S. cerevisiae* et *E. coli*, une des cellules filles est ainsi nettoyée des dommages. Ce mécanisme est appelé réjuvenation et permet ainsi à une cellule fille d'avoir

un potentiel répliatif non entamé et à l'autre d'hériter des dommages.

La bactérie *E. coli* montre une polarité fonctionnelle au niveau moléculaire. En effet, les agrégats protéiques s'accumulent préférentiellement dans la cellule contenant le vieux pôle. Les dommages sont ainsi hérités par les cellules vieillissantes [Lindner et al., 2008].

Comment le vieux pôle recrute-t-il les protéines endommagées et les agrégats protéiques dans la cellule? Très récemment [Lloyd-Price et al., 2012], la localisation polaire de complexes d'ARN bactériophage MS2 fusionné à la protéine fluorescente GFP (Green Fluorescent Protein) a été mise en évidence dans des cellules individuelles de la bactérie *E. coli*. Ces ARN sont des molécules non natives chez *E. coli* et ont une propension à l'agrégation. Chaque complexe a tendance à se diriger vers un pôle et à y rester. L'observation d'un biais en faveur d'un des pôles n'a pu être expliqué et fera l'objet de recherches plus approfondies par l'équipe de Lloyd.

A ce jour, la ségrégation en vers un pôle n'est pas élucidé.

Une autre définition du vieillissement peut être l'accumulation de dommages au sein d'un organisme et dans ce cas la longévité est garantie par les mécanismes de maintenance et réparation [Kirkwood and Austad, 2000]. Qu'en est-il chez la bactérie *E. coli* à long terme? Des expériences préliminaires (A.S. Coquel et M. Ni) réalisées à l'INSERM U1001, en conditions non-stressantes, associant la protéine chaperonne *ibpA* marquée avec la protéine fluorescente jaune et la microscopie time-lapse indique que la cellule mère accumulerait après 20-30 générations, plusieurs agrégats protéiques au vieux pôle (Figure 4.4, page 150). L'accumulation de ces dommages semblent inhiber la division puisque ces cellules filamentent. Il semblerait ainsi que lorsque les dommages atteignent un seuil critique ne permettant plus au système contrôle qualité de les réparer. Des études plus approfondies seront nécessaires afin de comprendre le mécanisme. Pourquoi les agrégats s'accumulent-ils au vieux pôles alors que la cellule a cessé de se diviser? Est ce que ces agrégats protéiques séquestrent des protéines fonctionnelles et essentielles à la division? Une étude récente [Mannik et al., 2012] montre que le système Min et le facteur d'occlusion des nucléoides *SlmA* que l'on pensait jouant un rôle dans la localisation du "divisome" chez les procaryotes, n'étaient pas indispensables. En effet, la délétion de *slmA* ne modifie pas la position relative des nucléoides dans des cellules de formes irrégulières. Ainsi la robustesse de la position des nucléoides permet à la cellule de maintenir la localisation du "divisome" et ainsi de se diviser en deux cellules de taille identique. Cependant, la délétion de *slmA* diminue la précision dans la localisation du divisome. Est ce que d'autres facteurs que *slmA* sont impliqués dans la division et séquestrés dans les agrégats protéiques?

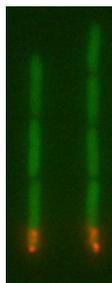


FIGURE 4.4 – Agrégation protéique à long terme chez la bactérie *E. coli* En rouge (fausse couleur) sont représentés les agrégats protéiques *Ibpa*-YFP.

I.3.2 Ségrégation stochastique

La ségrégation asymétrique n'est a priori pas la seule stratégie possible chez les organismes unicellulaires. La levure *S. pombe* semble avoir trouvé un compromis en mettant en oeuvre une agrégation passive et une ségrégation stochastique des agrégats protéiques ce qui permet à la cellule de passer d'un état "vieillissant" à un état "non vieillissant" [Coelho, 2012]. L'interrupteur activant le mode "vieillissant" est activé par l'augmentation de la quantité d'agrégats protéiques dans la cellule.

Le mécanisme d'agrégation protéique serait ainsi un mécanisme stochastique entraînant une dilution constante des dommages jusqu'à un certain seuil. Cependant, lors d'un stress environnemental par exemple thermique, le nombre d'agrégats protéiques augmente et la cellule n'est plus capable de les réparer. Elle adopte alors une stratégie de ségrégation des dommages et passe en mode "vieillissant". Ainsi, des cellules naissant avec un taux d'agrégats protéiques élevés vont avoir une probabilité d'environ 80% de mourir à la division suivante sans avoir montré précédemment des signes de vieillissement. Les hypothèses avancées seraient que les agrégats peuvent colocaliser au niveau du plan de division et peuvent alors interagir avec l'anneau acto-myosine qui permet la création de deux membranes cellulaires durant la cytokinèse. Il se pourrait alors que les agrégats interfèrent en empêchant la formation complète des deux membranes ou en capturant l'actine.

Ainsi, chez *S. pombe*, le vieillissement serait activé face à des conditions de croissance stressantes. En condition de croissance normale, les dommages seraient distribués aléatoirement. L'hypothèse de l'existence d'une boucle de régulation basée sur le niveau de dommages accumulés a été émise. Chez *E. coli*, en conditions non stressantes de laboratoire, nous n'observons pas d'état non vieillissant. Les dommages sont distribués asymétriquement, il n'y a pas de dilution des dommages ou cet état "non vieillissant" est temporaire et n'était pas observable en microscopie à notre échelle de temps. En revanche, à long terme (15-20 générations), l'accumulation d'agrégats au vieux pôle de la cellule mère semble arrêter la division. Est-ce que les agrégats protéiques piègent des protéines essentielles à la division ?

En revanche, chez la levure *S. cerevisiae*, la perte de la ségrégation asymétrique des dommages entraîne chez la cellule fille une espérance de vie répllicative plus courte. Les hypothèses avancées sont que les dommages hérités par la cellule fille seraient plus néfastes que ceux accumulés dans la cellule mère ou que la cellule mère possède un système contrôle qualité des protéines plus robuste afin de faire face aux dommages [Partridge, 2011].

De manière générale, il semblerait que pour une cellule, lorsque la quantité de dommages atteint un seuil critique, la cellule déploie une stratégie permettant de distribuer asymétriquement les dommages. Cette stratégie permet de donner naissance à une cellule fille indemne de dommages qui pourra perpétuer l'espèce et à une autre qui éliminera les dommages de la population. On peut en revanche se demander quels sont les mécanismes qui permettent de détecter lorsque le seuil critique est atteint et s'il n'existerait pas des mécanismes actifs lorsque ce seuil est dépassé afin de séquestrer et fusionner les agrégats afin qu'ils soient le moins toxiques possible pour la cellule.

II Toxicité des agrégats protéiques

La maladie du prion est une des maladies les plus effrayantes de part son caractère infectieux. D'autres agrégats protéiques ont des propriétés "prion-like" qui leur permettent de s'auto-répliquer. C'est le cas des protéines A- β et Tau impliquées dans la maladie d'Alzheimer [Nussbaum et al., 2012] ou encore de la protéine α -synucléine [Surgucheva et al., 2012] dans la maladie de Parkinson. Il semblerait que pour la plupart des maladies neurodégénératives, l'auto-réplication des agrégats soit le mécanisme principal de propagation [Polymenidou and Cleveland, 2012].

Une question d'actualité en biologie est "quels sont les facteurs impliquant un changement conformationnel des protéines solubles et fonctionnelles afin qu'elles s'assemblent en fibres amyloïdes?". En effet, les prions et les amyloïdes sont la cause de nombreuses maladies neurodégénératives. Ces maladies ont pour caractéristique commune une altération de la conformation native des protéines impliquées.

De nombreux changements moléculaires ou biophysiques comme des changements de température, pH ou encore des mutations peuvent modifier la conformation native d'une protéine. Les protéines n'ayant pas acquise leur conformation native ont une propension à s'agréger et peuvent aussi former des fibrilles amyloïdes sous certaines conditions [Stefani and Dobson, 2003]. Du fait que ces agrégats sont le résultat de l'accumulation de protéines mal repliées ou ayant acquis une conformation non native, ils exposent des régions hydrophobes, normalement enfouies au coeur des protéines, qui peuvent piéger des protéines fonctionnelles ou encore monopoliser le système contrôle qualité des protéines. Nous pouvons ainsi nous interroger sur la toxicité des agrégats protéiques pour la cellule. Est-il possible d'établir un lien entre la toxicité des agrégats protéiques et leur structure moléculaire ou encore leur taille ?

II.1 Résumé des résultats obtenus

Dans la seconde étude (article soumis courant octobre 2012), réalisée en collaboration avec l'équipe de R. Giraldo à Madrid, la construction d'un modèle synthétique basé sur la protéine RepA-WH1 a permis d'étudier une forme de protéinopathie amyloïde chez la bactérie *E. coli*. Nous avons caractérisé une transmission verticale de la cellule mère à la cellule fille, *in vivo*, sur une dizaine de générations grâce à la microscopie microfluidique. Nos résultats indiquent l'existence pour une même protéine de deux types morphologiques d'agrégats : de forme globulaire et de forme allongée (ou forme "comet"), ainsi qu'une variation dans le nombre d'agrégats par cellule ou encore de leur toxicité. En effet, les agrégats globulaires entraînent la mort de la cellule tandis que les agrégats de type "comet" semblent moins toxiques pour la cellule.

Par ailleurs, l'étude de l'influence de la protéine chaperonne DnaK et de la protéase ClpB, deux acteurs majeurs du système contrôle qualité des protéines, a permis de montrer le rôle central de DnaK afin de promouvoir la propagation du prionoïde RepA-WH1 sous forme "comet". Alors que pour des agrégats de type corps d'inclusion, la protéase ClpB désagrège les agrégats en cas de surexpression, dans notre cas, ClpB a une action très secondaire.

II.2 Influence de la forme des agrégats protéiques sur la toxicité

Les données publiées récemment indiquent que les précurseurs donnant naissance aux fibrilles matures sont les plus toxiques malgré la difficulté à les détecter *in vivo* du fait que ce sont des intermédiaires temporaires. Ils sont responsables des troubles des fonctions cellulaires observés tandis que les fibrilles matures ou plaques (maladie d'Alzheimer) sont considérées comme protectrice [Caughey and Lansbury, 2003] [Xue et al., 2010]. Ceci pourrait ainsi expliquer l'absence de corrélation directe entre la densité des plaques fibrillaires dans le cerveau de personnes atteintes de la maladie d'Alzheimer et la sévérité des symptômes [Dickson et al., 1995]. Cependant, les mécanismes de toxicité impliqués dans les maladies telles que Alzheimer ou Parkinson ne sont pas encore élucidés. Une des hypothèses énoncées est que les agrégats solubles (premiers intermédiaires ou précurseurs) sont toxiques pour la cellule en diminuant les capacités du système contrôle qualité des protéines. Ceci est supporté par une étude récente [Olzscha et al., 2011] qui a comparé la composition d'agrégats cellulaires toxiques et non toxiques dans des cellules de mammifères. Les agrégats toxiques sont enrichis en protéines cellulaires métastables. Ceci suggère que les agrégats toxiques séquestrent des composants essentiels tels que des protéines chaperonnes ou des facteurs impliqués dans l'homéostasie protéique [Olzscha et al., 2011]. Il a été montré également que la surexpression de protéines chaperonnes peut décaler dans le temps l'agrégation et empêcher la formation d'agrégats toxiques de certaines protéines associées à des maladies [Fonte et al., 2002] [Nollen et al., 2004] [van Ham et al., 2008] [Tam et al., 2009].

Un des aspects importants de la formation des amyloïdes est le mécanisme spécifique par lequel ils se forment à partir de précurseurs solubles ou non. Différentes études montrent l'existence d'étapes définies [Quintas et al., 2001] [Bucciantini et al., 2002]. Dans la plupart, des petites particules de types agrégats amorphes, de forme globulaire et "doughnut" ont été observées. Ces agrégats sont les précurseurs des proto-fibrilles (contenant de nombreux feuilletts β), qui donneront naissance ensuite aux fibrilles matures (Figure 4.5, page 153). Dans nos travaux, nous avons mis en évidence différents types d'agrégats dont la forme peut être plus ou moins toxique pour la cellule et agir ainsi sur le taux de croissance et la mortalité. Comme dans notre étude, il semblerait que les formes globulaires sont les plus toxiques.

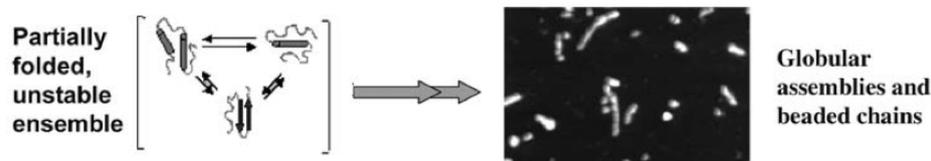


FIGURE 4.5 – **Formes des agrégats protéiques.** Certaines amyloïdes peuvent former des agrégats de forme apparente globulaire et "doughnut" avant de s'assembler en fibrilles. Figure extraite de [Stefani and Dobson, 2003].

Une étude très récente [Kachooei et al., 2012] a montré une corrélation entre la tension de surface et la neurotoxicité de fibrilles amyloïdes d'insuline dont on observe la formation chez des patients insulino-dépendants. Une diminution de la tension de surface lors de la formation des agrégats augmente les effets neurotoxiques. En effet, on a une exposition accrue des régions hydro-

phobes ce qui augmente la surface de l'agrégat et augmente la propension à piéger des protéines fonctionnelles ou les protéines chaperonnes du système contrôle qualité. En revanche, pour les proto-fibrilles et fibrilles matures, ils observent une augmentation de la tension de surface et une diminution de la toxicité par l'exposition préférentielle des parties hydrophiles et une structure plus compacte. La tension de surface pourrait ainsi être un indicateur des effets neurotoxiques.

Comment le système contrôle qualité décide-t-il si un agrégat doit être séquestré ou engagé dans le processus de repliement ou de dégradation. Une des hypothèses pourrait être que les agrégats insolubles sont reconnus en tant que tel et sont ainsi séquestrés. Quand aux intermédiaires de fibrilles amyloïdes et agrégats potentiellement toxiques, ils sont envoyés dans un "pseudo-compartiment" où ils peuvent s'agréger en fibres amyloïdes, forme moins toxique pour la cellule. Ces "pseudo-compartiments" existent-ils ? Comment la cellule les maintient-elle ? Existe-t-il des mécanismes actifs permettant d'adresser les agrégats toxiques vers ces compartiments ?

II.3 Influence de la taille des agrégats protéiques sur la toxicité

Etablir une corrélation entre la taille, la structure moléculaire et la toxicité des agrégats protéiques est un point essentiel afin de développer des stratégies contre les maladies neurodégénératives et déterminer les cibles moléculaires.

Une étude récente [Lee et al., 2011] s'est basée sur l'hypothèse que les précurseurs des fibrilles matures étaient plus toxiques. Ces précurseurs sont des fragments de taille plus petite que les fibrilles. Ils se sont intéressés ainsi à une corrélation possible entre la taille des fibrilles (fibrilles intactes ou découpées en fragments par sonication) et leur cytotoxicité dans la maladie du prion. En 2005, il avait été mis en évidence que des fragments PrP produits par sonication à partir d'un agrégat pathogène de protéines prions PrP^{sc} révélaient la plus grande toxicité observée pour des protéines prions [Silveira et al., 2005].

A partir d'une protéine prion du hamster, ils ont généré deux conformations R- et S- (chiralité) qui ont ensuite été fragmentées par sonication. On notera qu'après la sonication, les petits fragments générés conservent leur conformation R- ou S-. Ils montrent ainsi que la cellule réagit différemment à des amyloïdes de conformations différentes malgré une séquence d'acides aminés identiques. En effet, selon la conformation (R- ou S-), la diminution de taille des amyloïdes peut augmenter ou abolir l'effet cytotoxique. La relation entre la taille des amyloïdes et leur cytotoxicité ne serait donc pas unidirectionnelle et dépendrait de la structure moléculaire de l'agrégat.

II.4 *Escherichia coli*, modèle d'étude des prions

Actuellement, peu d'organismes modèle permettent d'étudier les prions. Pour beaucoup d'organismes unicellulaires, aucun prion n'a encore été identifié à l'exception de la levure. Récemment, Espargaro et ses collaborateurs [Espargaro et al., 2012] ont montré que les prions de levure Sup35 et Ure2 provoquaient la formation de corps d'inclusion avec des propriétés amyloïdes-like quand ils étaient exprimés chez la bactérie *E. coli*. Ces agrégats servent de matrice afin de promouvoir l'agrégation de molécules priogéniques solubles homologues.

Cette étude ajoutée à nos résultats renforce l'utilisation de la bactérie *E. coli* comme modèle d'étude des prions. La bactérie *E. coli* semble disposer dans son cytoplasme des éléments nécessaires à une transmission verticale et permet ainsi l'étude de la toxicité des prions. De plus,

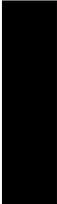
sa reproduction rapide toutes les 20 minutes à 1 heure selon les conditions ainsi que les outils moléculaires et génétiques développés en font un modèle puissant et permettent des études variées.

III Conclusion

L'agrégation protéique, ses mécanismes et conséquences sont un sujet très actuel du fait de son implication dans de nombreuses maladies neurodégénératives humaines mais aussi dans le processus normal de repliement des protéines.

Malgré la présence d'un système puissant de contrôle qualité des protéines, dont les capacités décroissent avec l'âge, la cellule doit faire face à l'accumulation de dommages qui a des conséquences sur ses fonctions. Selon les conditions environnementales, mais aussi selon leur structure et leur taille, ces agrégats peuvent être plus ou moins toxiques pour la cellule en agissant comme un aimant pour les protéines fonctionnelles ou les protéines chaperonnes.

Dans la course aux remèdes contre ces maladies, les recherches sont actuellement centrées sur les précurseurs, forme la plus toxique, des fibrilles amyloïdes afin de comprendre les mécanismes les rendant si néfastes. La compréhension de ces mécanismes permettra d'empêcher la formation de ces agrégats toxiques ou d'inhiber leur toxicité en déterminant des cibles thérapeutiques plus précises.



Bibliographie

- [Ackermann et al., 2007] Ackermann, M., Chao, L., Bergstrom, C. T., and Doebeli, M. (2007). On the evolutionary origin of aging. *Aging cell*, 6 :235–44.
- [Ackermann et al., 2003] Ackermann, M., Chao, L., Bergstrom, C. T., Doebeli, M., Ackermann, M., Stearns, S. C., and Jenal, U. (2003). Senescence in a bacterium with asymmetric division. *Aging Cell*, 300(5627) :1920.
- [Agmon, 1995] Agmon, N. (1995). The grotthuss mechanism. *Chemical Physics Letters*, 244 :456–462.
- [Aguilaniu et al., 2003] Aguilaniu, H., Gustafsson, L., Rigoulet, M., and Nystrom, T. (2003). Asymmetric inheritance of oxidatively damaged proteins during cytokinesis. *Science*, 299(5613) :1751–3.
- [Aguzzi and Rajendran, 2009] Aguzzi, A. and Rajendran, L. (2009). The transcellular spread of cytosolic amyloids, prions, and prionoids. *Neuron*, 64(6) :783–90.
- [Alberti et al., 2009] Alberti, S., Halfmann, R., King, O., Kapila, A., and Lindquist, S. (2009). A systematic survey identifies prions and illuminates sequence features of prionogenic proteins. *Cell*, 137(1) :146–58.
- [Allen et al., 1992] Allen, S. P., Polazzi, J. O., Gierse, J. K., and Easton, A. M. (1992). Two novel heat shock genes encoding proteins produced in response to heterologous protein expression in escherichia coli. *J Bacteriol*, 174(21) :6938–47.
- [Andley, 2006] Andley, U. P. (2006). Crystallins and hereditary cataracts : molecular mechanisms and potential for therapy. *Expert Rev Mol Med*, 8(25) :1–19.
- [Ando and Skolnick, 2010] Ando, T. and Skolnick, J. (2010). Crowding and hydrodynamic interactions likely dominate in vivo macromolecular motion. *Proc Natl Acad Sci U S A*, 107(43) :18457–62.
- [Anfinsen, 1973] Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. *Science*, 181(4096) :223–30.
- [Arsene et al., 2000] Arsene, F., Tomoyasu, T., and Bukau, B. (2000). The heat shock response of escherichia coli. *Int J Food Microbiol*, 55(1-3) :3–9.

- [Ausmees et al., 2003] Ausmees, N., Kuhn, J. R., and Jacobs-Wagner, C. (2003). The bacterial cytoskeleton : an intermediate filament-like function in cell shape. *Cell*, 115(6) :705–13.
- [Azriel and Gazit, 2001] Azriel, R. and Gazit, E. (2001). Analysis of the minimal amyloid-forming fragment of the islet amyloid polypeptide. an experimental support for the key role of the phenylalanine residue in amyloid formation. *J Biol Chem*, 276(36) :34156–61.
- [Bachelier, 1900] Bachelier, L. (1900). *Theorie de la speculation*. PhD thesis, Annales de l'Ecole Normale superieure.
- [Baldwin et al., 1995] Baldwin, M. A., Cohen, F. E., and Prusiner, S. B. (1995). Prion protein isoforms, a convergence of biological and structural investigations. *J Biol Chem*, 270(33) :19197–200.
- [Baldwin, 1997] Baldwin, R. L. (1997). The problem was to find the problem. *Protein Sci*, 6(9) :2031–4.
- [Baneyx and Mujacic, 2004] Baneyx, F. and Mujacic, M. (2004). Recombinant protein folding and misfolding in escherichia coli. *Nat Biotechnol*, 22(11) :1399–408.
- [Barker and Walmsley, 1999] Barker, M. G. and Walmsley, R. M. (1999). Replicative ageing in the fission yeast schizosaccharomyces pombe. *Yeast*, 15(14) :1511–8.
- [Blank et al., 1986] Blank, A., Gallant, J. A., Burgess, R. R., and Loeb, L. A. (1986). An rna polymerase mutant with reduced accuracy of chain elongation. *Biochemistry*, 25(20) :5920–8.
- [Bonabeau, 2002] Bonabeau, E. (2002). Agent-based modeling : methods and techniques for simulating human systems. *Proc Natl Acad Sci U S A*, 99 Suppl 3 :7280–7.
- [Bowman et al., 2008] Bowman, G. R., Comolli, L. R., Zhu, J., Eckart, M., Koenig, M., Downing, K. H., Moerner, W. E., Earnest, T., and Shapiro, L. (2008). A polymeric protein anchors the chromosomal origin/parb complex at a bacterial cell pole. *Cell*, 134(6) :945–55.
- [Brown, 1828] Brown, R. (1828). A brief account of microscopical observations on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *Edinburgh Journal of Science*.
- [Bucciantini et al., 2002] Bucciantini, M., Giannoni, E., Chiti, F., Baroni, F., Formigli, L., Zurdo, J., Taddei, N., Ramponi, G., Dobson, C. M., and Stefani, M. (2002). Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature*, 416(6880) :507–11.
- [Buchberger et al., 1996] Buchberger, A., Schroder, H., Hestekamp, T., Schonfeld, H. J., and Bukau, B. (1996). Substrate shuttling between the dnaK and groEL systems indicates a chaperone network promoting protein folding. *J Mol Biol*, 261(3) :328–33.
- [Bukau and Horwich, 1998] Bukau, B. and Horwich, A. L. (1998). The hsp70 and hsp60 chaperone machines. *Cell*, 92(3) :351–66.
- [Campbell et al., 2002] Campbell, R. E., Tour, O., Palmer, A. E., Steinbach, P. A., Baird, G. S., Zacharias, D. A., and Tsien, R. Y. (2002). A monomeric red fluorescent protein. *Proc Natl Acad Sci U S A*, 99(12) :7877–82.
- [Carrio et al., 2005] Carrio, M., Gonzalez-Montalban, N., Vera, A., Villaverde, A., and Ventura, S. (2005). Amyloid-like properties of bacterial inclusion bodies. *J Mol Biol*, 347(5) :1025–37.
- [Carrio et al., 1998] Carrio, M. M., Corchero, J. L., and Villaverde, A. (1998). Dynamics of in vivo protein aggregation : building inclusion bodies in recombinant bacteria. *FEMS Microbiol Lett*, 169(1) :9–15.

- [Carrio and Villaverde, 2003] Carrio, M. M. and Villaverde, A. (2003). Role of molecular chaperones in inclusion body formation. *FEBS Lett*, 537(1-3) :215–21.
- [Caughey and Lansbury, 2003] Caughey, B. and Lansbury, P. T. (2003). Protofibrils, pores, fibrils, and neurodegeneration : separating the responsible protein aggregates from the innocent bystanders. *Annu Rev Neurosci*, 26 :267–98.
- [Chan and Dill, 1990] Chan, H. S. and Dill, K. A. (1990). Origins of structure in globular proteins. *Proc Natl Acad Sci U S A*, 87(16) :6388–92.
- [Chao, 2010] Chao, L. (2010). A model for damage load and its implications for the evolution of bacterial aging. *PLoS Genet*, 6(8).
- [Chiti and Dobson, 2006] Chiti, F. and Dobson, C. M. (2006). Protein misfolding, functional amyloid, and human disease. *Annu Rev Biochem*, 75 :333–66.
- [Chiti and Dobson, 2009] Chiti, F. and Dobson, C. M. (2009). Amyloid formation by globular proteins under native conditions. *Nat Chem Biol*, 5(1) :15–22.
- [Choi et al., 2008] Choi, P. J., Cai, L., Frieda, K., and Xie, X. S. (2008). A stochastic single-molecule event triggers phenotype switching of a bacterial cell. *Science*, 322(5900) :442–6.
- [Cobb and Surewicz, 2009] Cobb, N. J. and Surewicz, W. K. (2009). Prion diseases and their biochemical mechanisms. *Biochemistry*, 48(12) :2574–85.
- [Coelho, 2012] Coelho, M. (2012). *Ageing by passive aggregation and stochastic distribution of protein aggregates*. PhD thesis, Technische Universitat Desden.
- [Coppede et al., 2006] Coppede, F., Mancuso, M., Siciliano, G., Migliore, L., and Murri, L. (2006). Genes and the environment in neurodegeneration. *Biosci Rep*, 26(5) :341–67.
- [Cox, 1965] Cox, B. S. (1965). [psi], a cytoplasmic suppressor of super-suppressor in yeast. *Heredity*, 20 :505–521.
- [Crow and Li, 2011] Crow, E. T. and Li, L. (2011). Newly identified prions in budding yeast, and their possible functions. *Semin Cell Dev Biol*, 22(5) :452–9.
- [de Groot et al., 2009] de Groot, N. S., Sabate, R., and Ventura, S. (2009). Amyloids in bacterial inclusion bodies. *Trends Biochem Sci*, 34(8) :408–16.
- [DeBurman et al., 1997] DeBurman, S. K., Raymond, G. J., Caughey, B., and Lindquist, S. (1997). Chaperone-supervised conversion of prion protein to its protease-resistant form. *Proc Natl Acad Sci U S A*, 94(25) :13938–43.
- [Deich et al., 2004] Deich, J., Judd, E. M., McAdams, H. H., and Moerner, W. E. (2004). Visualization of the movement of single histidine kinase molecules in live caulobacter cells. *Proc Natl Acad Sci U S A*, 101(45) :15921–6.
- [Demarez, 2011] Demarez, A. (2011). *Investigating proteostasis and ageing of Escherichia coli using spatio-temporal algorithms*. PhD thesis, Paris Descartes.
- [Derkatch et al., 2001] Derkatch, I. L., Bradley, M. E., Hong, J. Y., and Liebman, S. W. (2001). Prions affect the appearance of other prions : the story of [pin(+)]. *Cell*, 106(2) :171–82.
- [Dickson et al., 1995] Dickson, D. W., Crystal, H. A., Bevona, C., Honer, W., Vincent, I., and Davies, P. (1995). Correlations of synaptic and pathological markers with cognition of the elderly. *Neurobiol Aging*, 16(3) :285–98 ; discussion 298–304.

- [Dill and Chan, 1997] Dill, K. A. and Chan, H. S. (1997). From levinthal to pathways to funnels. *Nat Struct Biol*, 4(1) :10–9.
- [Dix and Verkman, 2008] Dix, J. A. and Verkman, A. S. (2008). Crowding effects on diffusion in solutions and cells. *Annu Rev Biophys*, 37 :247–63.
- [Dobson, 2001] Dobson, C. M. (2001). Protein folding and its links with human disease. *Biochem Soc Symp*, 1(68) :1–26.
- [Dobson and Ellis, 1998] Dobson, C. M. and Ellis, R. J. (1998). Protein folding and misfolding inside and outside the cell. *EMBO J*, 17(18) :5251–4.
- [Dobson and Karplus, 1999] Dobson, C. M. and Karplus, M. (1999). The fundamentals of protein folding : bringing together theory and experiment. *Curr Opin Struct Biol*, 9(1) :92–101.
- [Donachie and Begg, 1989] Donachie, W. D. and Begg, K. J. (1989). Chromosome partition in escherichia coli requires postreplication protein synthesis. *J Bacteriol*, 171(10) :5405–9.
- [Dougan et al., 2002] Dougan, D. A., Mogk, A., and Bukau, B. (2002). Protein folding and degradation in bacteria : to degrade or not to degrade ? that is the question. *Cell Mol Life Sci*, 59(10) :1607–16.
- [Drubin and Nelson, 1996] Drubin, D. G. and Nelson, W. J. (1996). Origins of cell polarity. *Cell*, 84(3) :335–44.
- [Drummond and Wilke, 2008] Drummond, D. A. and Wilke, C. O. (2008). Mistranslation-induced protein misfolding as a dominant constraint on coding-sequence evolution. *Cell*, 134(2) :341–52.
- [Ebersbach et al., 2006] Ebersbach, G., Ringgaard, S., Moller-Jensen, J., Wang, Q., Sherratt, D. J., and Gerdes, K. (2006). Regular cellular distribution of plasmids by oscillating and filament-forming para atpase of plasmid pb171. *Mol Microbiol*, 61(6) :1428–42.
- [Einstein, 1905] Einstein, A. (1905). Über einen die erzeugung und verwandlung des lichtes betreffenden heuristischen gesichtspunkt. *Annalen der Physik*, 322(6) :132–148.
- [Einstein, 1956] Einstein, A. (1956). Investigations of the theory of the brownian movement. *ed. R. FÁ $\frac{1}{4}$ rth, Dover, New York*.
- [Elf et al., 2007] Elf, J., Li, G. W., and Xie, X. S. (2007). Probing transcription factor dynamics at the single-molecule level in a living cell. *Science*, 316(5828) :1191–4.
- [Ellis, 2001a] Ellis, R. J. (2001a). Macromolecular crowding : obvious but underappreciated. *Trends Biochem Sci*, 26(10) :597–604.
- [Ellis, 2001b] Ellis, R. J. (2001b). Molecular chaperones : inside and outside the anfinen cage. *Curr Biol*, 11(24) :R1038–40.
- [Ellis and Minton, 2006] Ellis, R. J. and Minton, A. P. (2006). Protein aggregation in crowded environments. *Biol Chem*, 387(5) :485–97.
- [Ellis and van der Vies, 1991] Ellis, R. J. and van der Vies, S. M. (1991). Molecular chaperones. *Annu Rev Biochem*, 60 :321–47.
- [Elmore et al., 2005] Elmore, S., Muller, M., Vischer, N., Odijk, T., and Woldringh, C. L. (2005). Single-particle tracking of oric-gfp fluorescent spots during chromosome segregation in escherichia coli. *J Struct Biol*, 151(3) :275–87.

- [Elowitz et al., 1997] Elowitz, M. B., Surette, M. G., Wolf, P. E., Stock, J., and Leibler, S. (1997). Photoactivation turns green fluorescent protein red. *Curr Biol*, 7(10) :809–12.
- [Elowitz et al., 1999] Elowitz, M. B., Surette, M. G., Wolf, P. E., Stock, J. B., and Leibler, S. (1999). Protein mobility in the cytoplasm of escherichia coli. *J Bacteriol*, 181(1) :197–203.
- [Erjavec et al., 2008] Erjavec, N., Cvijovic, M., Klipp, E., and Nystrom, T. (2008). Selective benefits of damage partitioning in unicellular systems and its effects on aging. *Proc Natl Acad Sci U S A*, 105(48) :18764–9.
- [Espargaro et al., 2012] Espargaro, A., Villar-Pique, A., Sabate, R., and Ventura, S. (2012). Yeast prions form infectious amyloid inclusion bodies in bacteria. *Microb Cell Fact*, 11(1) :89.
- [Etienne-Manneville and Hall, 2002] Etienne-Manneville, S. and Hall, A. (2002). Rho gtpases in cell biology. *Nature*, 420(6916) :629–35.
- [Fabrizio and Longo, 2003] Fabrizio, P. and Longo, V. D. (2003). The chronological life span of saccharomyces cerevisiae. *Aging Cell*, 2(2) :73–81.
- [Fedotov et al., 2010] Fedotov, S., Al-Shamsi, H., Ivanov, A., and Zubarev, A. (2010). Anomalous transport and nonlinear reactions in spiny dendrites. *Phys Rev E Stat Nonlin Soft Matter Phys*, 82(4 Pt 1) :041103.
- [Ferber, 1995] Ferber, J. (1995). *Les systemes multi-agents, vers une intelligence collective*.
- [Fernandez-Tresguerres et al., 2010] Fernandez-Tresguerres, M. E., De La Espina, S. M., Gasset-Rosa, F., and Giraldo, R. (2010). A dna-promoted amyloid proteinopathy in escherichia coli. *Mol Microbiol*.
- [Figge et al., 2004] Figge, R. M., Divakaruni, A. V., and Gober, J. W. (2004). Mreb, the cell shape-determining bacterial actin homologue, co-ordinates cell wall morphogenesis in caulobacter crescentus. *Mol Microbiol*, 51(5) :1321–32.
- [Foley et al., 2010] Foley, P. L., Wilson, D. B., and Shuler, M. L. (2010). Macromolecular crowding can account for rnase-sensitive constraint of bacterial nucleoid structure. *Biochem Biophys Res Commun*, 395(1) :42–7.
- [Fonte et al., 2002] Fonte, V., Kapulkin, V., Taft, A., Fluet, A., Friedman, D., and Link, C. D. (2002). Interaction of intracellular beta amyloid peptide with chaperone proteins. *Proc Natl Acad Sci U S A*, 99(14) :9439–44.
- [Fumentalba et al., 2010] Fumentalba, R. A., Udan, M., Bell, S., Wegorzewska, I., Shao, J., Diamond, M. I., Weihl, C. C., and Baloh, R. H. (2010). Interaction with polyglutamine aggregates reveals a q/n-rich domain in tdp-43. *J Biol Chem*, 285(34) :26304–14.
- [Garrity et al., 2010] Garrity, S. J., Sivanathan, V., Dong, J., Lindquist, S., and Hochschild, A. (2010). Conversion of a yeast prion protein to an infectious form in bacteria. *Proc Natl Acad Sci U S A*, 107(23) :10596–601.
- [Gething and Sambrook, 1992] Gething, M. J. and Sambrook, J. (1992). Protein folding in the cell. *Nature*, 355(6355) :33–45.
- [Giraldo, 2007] Giraldo, R. (2007). Defined dna sequences promote the assembly of a bacterial protein into distinct amyloid nanostructures. *Proc Natl Acad Sci U S A*, 104(44) :17388–93.
- [Gitai et al., 2004] Gitai, Z., Dye, N., and Shapiro, L. (2004). An actin-like gene can determine cell polarity in bacteria. *Proc Natl Acad Sci U S A*, 101(23) :8643–8.

- [Golding and Cox, 2006] Golding, I. and Cox, E. C. (2006). Physical nature of bacterial cytoplasm. *Phys Rev Lett*, 96(9) :098102.
- [Gonzalez-Montalban et al., 2005] Gonzalez-Montalban, N., Carrio, M. M., Cuatrecasas, S., Aris, A., and Villaverde, A. (2005). Bacterial inclusion bodies are cytotoxic in vivo in absence of functional chaperones dnak or groel. *J Biotechnol*, 118(4) :406–12.
- [Gouy, 1888] Gouy, L. (1888). Sur la conservation de l’électricité et la thermodynamique. *Comptes Rendus des Seances de l’Academie des Sciences. Paris*, 107 :329–332.
- [Gribskov and Burgess, 1983] Gribskov, M. and Burgess, R. R. (1983). Overexpression and purification of the sigma subunit of escherichia coli rna polymerase. *Gene*, 26(2-3) :109–18.
- [Hager et al., 2009] Hager, G. L., McNally, J. G., and Misteli, T. (2009). Transcription dynamics. *Mol Cell*, 35(6) :741–53.
- [Halfmann et al., 2011] Halfmann, R., Alberti, S., Krishnan, R., Lyle, N., O’Donnell, C. W., King, O. D., Berger, B., Pappu, R. V., and Lindquist, S. (2011). Opposing effects of glutamine and asparagine govern prion formation by intrinsically disordered proteins. *Mol Cell*, 43(1) :72–84.
- [Halfmann et al., 2012] Halfmann, R., Jarosz, D. F., Jones, S. K., Chang, A., Lancaster, A. K., and Lindquist, S. (2012). Prions are a common mechanism for phenotypic inheritance in wild yeasts. *Nature*, 482(7385) :363–8.
- [Hardy and Selkoe, 2002] Hardy, J. and Selkoe, D. J. (2002). The amyloid hypothesis of alzheimer’s disease : progress and problems on the road to therapeutics. *Science*, 297(5580) :353–6.
- [Harper and Lansbury, 1997] Harper, J. D. and Lansbury, P. T., J. (1997). Models of amyloid seeding in alzheimer’s disease and scrapie : mechanistic truths and physiological consequences of the time-dependent solubility of amyloid proteins. *Annu Rev Biochem*, 66 :385–407.
- [Hartl, 1996] Hartl, F. U. (1996). Molecular chaperones in cellular protein folding. *Nature*, 381(6583) :571–9.
- [Hartl et al., 2011] Hartl, F. U., Bracher, A., and Hayer-Hartl, M. (2011). Molecular chaperones in protein folding and proteostasis. *Nature*, 475(7356) :324–32.
- [Hartl and Hayer-Hartl, 2002] Hartl, F. U. and Hayer-Hartl, M. (2002). Molecular chaperones in the cytosol : from nascent chain to folded protein. *Science*, 295(5561) :1852–8.
- [Hartl and Hayer-Hartl, 2009] Hartl, F. U. and Hayer-Hartl, M. (2009). Converging concepts of protein folding in vitro and in vivo. *Nat Struct Mol Biol*, 16(6) :574–81.
- [Hatters et al., 2002] Hatters, D. M., Minton, A. P., and Howlett, G. J. (2002). Macromolecular crowding accelerates amyloid formation by human apolipoprotein c-ii. *J Biol Chem*, 277(10) :7824–30.
- [Hemmingsen et al., 1988] Hemmingsen, S. M., Woolford, C., van der Vies, S. M., Tilly, K., Dennis, D. T., Georgopoulos, C. P., Hendrix, R. W., and Ellis, R. J. (1988). Homologous plant and bacterial proteins chaperone oligomeric protein assembly. *Nature*, 333(6171) :330–4.
- [Hornemann and Glockshuber, 1998] Hornemann, S. and Glockshuber, R. (1998). A scrapie-like unfolding intermediate of the prion protein domain prp(121-231) induced by acidic ph. *Proc Natl Acad Sci U S A*, 95(11) :6010–4.

- [Huang et al., 2001] Huang, H. C., Sherman, M. Y., Kandrór, O., and Goldberg, A. L. (2001). The molecular chaperone dnaJ is required for the degradation of a soluble abnormal protein in *Escherichia coli*. *J Biol Chem*, 276(6) :3920–8.
- [Jahn and Radford, 2005] Jahn, T. R. and Radford, S. E. (2005). The yin and yang of protein folding. *FEBS J*, 272(23) :5962–70.
- [Jamin and Baldwin, 1998] Jamin, M. and Baldwin, R. L. (1998). Two forms of the pH 4 folding intermediate of apomyoglobin. *J Mol Biol*, 276(2) :491–504.
- [Jazwinski, 1990] Jazwinski, S. M. (1990). Aging and senescence of the budding yeast *Saccharomyces cerevisiae*. *Mol Microbiol*, 4(3) :337–43.
- [Jones et al., 2001] Jones, L. J., Carballido-Lopez, R., and Errington, J. (2001). Control of cell shape in bacteria : helical, actin-like filaments in *Bacillus subtilis*. *Cell*, 104(6) :913–22.
- [Kachooei et al., 2012] Kachooei, E., Moosavi-Movahedi, A. A., Khodagholi, F., Ramshini, H., Shaerzadeh, F., and Sheibani, N. (2012). Oligomeric forms of insulin amyloid aggregation disrupt outgrowth and complexity of neuron-like pc12 cells. *PLoS One*, 7(7) :e41344.
- [Kaeberlein, 2010] Kaeberlein, M. (2010). Lessons on longevity from budding yeast. *Nature*, 464(7288) :513–9.
- [Kahn et al., 1999] Kahn, S. E., Andrikopoulos, S., and Verchere, C. B. (1999). Islet amyloid : a long-recognized but underappreciated pathological feature of type 2 diabetes. *Diabetes*, 48(2) :241–53.
- [Kennedy et al., 1994] Kennedy, B. K., Austriaco, N. R., J., and Guarente, L. (1994). Daughter cells of *Saccharomyces cerevisiae* from old mothers display a reduced life span. *J Cell Biol*, 127(6 Pt 2) :1985–93.
- [Kim and Baldwin, 1990] Kim, P. S. and Baldwin, R. L. (1990). Intermediates in the folding reactions of small proteins. *Annu Rev Biochem*, 59 :631–60.
- [Kim et al., 2006] Kim, S. Y., Gitai, Z., Kinkhabwala, A., Shapiro, L., and Moerner, W. E. (2006). Single molecules of the bacterial actin mreB undergo directed treadmilling motion in *Caulobacter crescentus*. *Proc Natl Acad Sci U S A*, 103(29) :10929–34.
- [Kirkwood and Austad, 2000] Kirkwood, T. B. and Austad, S. N. (2000). Why do we age? *Nature*, 408(6809) :233–8.
- [Konopka et al., 2009] Konopka, M. C., Sochacki, K. A., Bratton, B. P., Shkel, I. A., Record, M. T., and Weisshaar, J. C. (2009). Cytoplasmic protein mobility in osmotically stressed *Escherichia coli*. *J Bacteriol*, 191(1) :231–7.
- [Krantz et al., 2002] Krantz, B. A., Mayne, L., Rumbley, J., Englander, S. W., and Sosnick, T. R. (2002). Fast and slow intermediate accumulation and the initial barrier mechanism in protein folding. *J Mol Biol*, 324(2) :359–71.
- [Krishnan and Lindquist, 2005] Krishnan, R. and Lindquist, S. L. (2005). Structural insights into a yeast prion illuminate nucleation and strain diversity. *Nature*, 435(7043) :765–72.
- [Kruse et al., 2006] Kruse, T., Blagoev, B., Lobner-Olesen, A., Wachi, M., Sasaki, K., Iwai, N., Mann, M., and Gerdes, K. (2006). Actin homolog mreB and rna polymerase interact and are both required for chromosome segregation in *Escherichia coli*. *Genes Dev*, 20(1) :113–24.

- [Kryndushkin et al., 2003] Kryndushkin, D. S., Alexandrov, I. M., Ter-Avanesyan, M. D., and Kushnirov, V. V. (2003). Yeast [psi+] prion aggregates are formed by small sup35 polymers fragmented by hsp104. *J Biol Chem*, 278(49) :49636–43.
- [Kubelka et al., 2004] Kubelka, J., Hofrichter, J., and Eaton, W. A. (2004). The protein folding 'speed limit'. *Curr Opin Struct Biol*, 14(1) :76–88.
- [Kuczynska-Wisnik et al., 2002] Kuczynska-Wisnik, D., Kedzierska, S., Matuszewska, E., Lund, P., Taylor, A., Lipinska, B., and Laskowska, E. (2002). The escherichia coli small heat-shock proteins ibpa and ibpb prevent the aggregation of endogenous proteins denatured in vivo during extreme heat shock. *Microbiology*, 148(Pt 6) :1757–65.
- [Kumar et al., 2010] Kumar, M., Mommer, M. S., and Sourjik, V. (2010). Mobility of cytoplasmic, membrane, and dna-binding proteins in escherichia coli. *Biophys J*, 98(4) :552–9.
- [Kwiatkowska et al., 2008] Kwiatkowska, J., Matuszewska, E., Kuczynska-Wisnik, D., and Laskowska, E. (2008). Aggregation of escherichia coli proteins during stationary phase depends on glucose and oxygen availability. *Res Microbiol*, 159(9-10) :651–7.
- [Lacroute, 1971] Lacroute, F. (1971). Non-mendelian mutation allowing ureidosuccinic acid uptake in yeast. *J Bacteriol*, 106(2) :519–22.
- [Landgraf et al., 2012] Landgraf, D., Okumus, B., Chien, P., Baker, T. A., and Paulsson, J. (2012). Segregation of molecules at cell division reveals native protein localization. *Nat Methods*.
- [Langston et al., 1983] Langston, J. W., Ballard, P., Tetrud, J. W., and Irwin, I. (1983). Chronic parkinsonism in humans due to a product of meperidine-analog synthesis. *Science*, 219(4587) :979–80.
- [Laskowska et al., 1996] Laskowska, E., Wawrzynow, A., and Taylor, A. (1996). Ibpa and ibpb, the new heat-shock proteins, bind to endogenous escherichia coli proteins aggregated intracellularly by heat shock. *Biochimie*, 78(2) :117–22.
- [Lee et al., 2011] Lee, Y. J., Savtchenko, R., Ostapchenko, V. G., Makarava, N., and Baskakov, I. V. (2011). Molecular structure of amyloid fibrils controls the relationship between fibrillar size and toxicity. *PLoS One*, 6(5) :e20244.
- [Levinthal, 1968] Levinthal, C. (1968). Are these pathways for protein folding? *Journal de Chimie Physique et de Physico-Chimie biologique*, 65 :44–45.
- [Lindner and Demarez, 2009] Lindner, A. B. and Demarez, A. (2009). Protein aggregation as a paradigm of aging. *Biochim Biophys Acta*, 1790(10) :980–96.
- [Lindner et al., 2008] Lindner, A. B., Madden, R., Demarez, A., Stewart, E. J., and Taddei, F. (2008). Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation. *Proc Natl Acad Sci U S A*, 105(8) :3076–81.
- [Liu et al., 2011] Liu, B., Larsson, L., Franssens, V., Hao, X., Hill, S. M., Andersson, V., Høglund, D., Song, J., Yang, X., Oling, D., Grantham, J., Winderickx, J., and Nystrom, T. (2011). Segregation of protein aggregates involves actin and the polarity machinery. *Cell*, 147(5) :959–61.
- [Lloyd-Price et al., 2012] Lloyd-Price, J., Hakkinen, A., Kandhavelu, M., Marques, I. J., Chowdhury, S., Lihavainen, E., Yli-Harja, O., and Ribeiro, A. S. (2012). Asymmetric disposal of individual protein aggregates in escherichia coli, one aggregate at a time. *J Bacteriol*, 194(7) :1747–52.

- [Loftfield and Vanderjagt, 1972] Loftfield, R. B. and Vanderjagt, D. (1972). The frequency of errors in protein biosynthesis. *Biochem J*, 128(5) :1353–6.
- [Lomakin et al., 1997] Lomakin, A., Teplow, D. B., Kirschner, D. A., and Benedek, G. B. (1997). Kinetic theory of fibrillogenesis of amyloid beta-protein. *Proc Natl Acad Sci U S A*, 94(15) :7942–7.
- [Lowe and Amos, 1998] Lowe, J. and Amos, L. A. (1998). Crystal structure of the bacterial cell-division protein ftsz. *Nature*, 391(6663) :203–6.
- [Luby-Phelps et al., 1987] Luby-Phelps, K., Castle, P. E., Taylor, D. L., and Lanni, F. (1987). Hindered diffusion of inert tracer particles in the cytoplasm of mouse 3t3 cells. *Proc Natl Acad Sci U S A*, 84(14) :4910–3.
- [Lukacs et al., 2000] Lukacs, G. L., Haggie, P., Seksek, O., Lechardeur, D., Freedman, N., and Verkman, A. S. (2000). Size-dependent dna mobility in cytoplasm and nucleus. *J Biol Chem*, 275(3) :1625–9.
- [Lynch, 2005] Lynch, M. D. (2005). Replicative aging in e. coli. *Rejuvenation Res*, 8(2) :79–81.
- [Ma et al., 2012] Ma, Q., Fan, J. B., Zhou, Z., Zhou, B. R., Meng, S. R., Hu, J. Y., Chen, J., and Liang, Y. (2012). The contrasting effect of macromolecular crowding on amyloid fibril formation. *PLoS One*, 7(4) :e36288.
- [Maisonneuve et al., 2008] Maisonneuve, E., Ezraty, B., and Dukan, S. (2008). Protein aggregates : an aging factor involved in cell death. *J Bacteriol*, 190(18) :6070–5.
- [Maji et al., 2009] Maji, S. K., Perrin, M. H., Sawaya, M. R., Jessberger, S., Vadodaria, K., Rissman, R. A., Singru, P. S., Nilsson, K. P., Simon, R., Schubert, D., Eisenberg, D., Rivier, J., Sawchenko, P., Vale, W., and Riek, R. (2009). Functional amyloids as natural storage of peptide hormones in pituitary secretory granules. *Science*, 325(5938) :328–32.
- [Mannik et al., 2012] Mannik, J., Wu, F., Hol, F. J., Bisicchia, P., Sherratt, D. J., Keymer, J. E., and Dekker, C. (2012). Robustness and accuracy of cell division in escherichia coli in diverse cell shapes. *Proc Natl Acad Sci U S A*, 109(18) :6957–62.
- [Martin and Hartl, 1997] Martin, J. and Hartl, F. U. (1997). The effect of macromolecular crowding on chaperonin-mediated protein folding. *Proc Natl Acad Sci U S A*, 94(4) :1107–12.
- [Martinez-Alonso et al., 2009] Martinez-Alonso, M., Toledo-Rubio, V., Noad, R., Unzueta, U., Ferrer-Miralles, N., Roy, P., and Villaverde, A. (2009). Rehosting of bacterial chaperones for high-quality protein production. *Appl Environ Microbiol*, 75(24) :7850–4.
- [Maurizi, 1992] Maurizi, M. R. (1992). Proteases and protein degradation in escherichia coli. *Experientia*, 48(2) :178–201.
- [McGuffee and Elcock, 2010] McGuffee, S. R. and Elcock, A. H. (2010). Diffusion, crowding & protein stability in a dynamic molecular model of the bacterial cytoplasm. *PLoS Comput Biol*, 6(3) :e1000694.
- [Mileykovskaya, 2007] Mileykovskaya, E. (2007). Subcellular localization of escherichia coli osmosensory transporter prop : focus on cardiolipin membrane domains. *Mol Microbiol*, 64(6) :1419–22.
- [Minton, 1983] Minton, A. P. (1983). The effect of volume occupancy upon the thermodynamic activity of proteins : some biochemical consequences. *Mol Cell Biochem*, 55(2) :119–40.

- [Minton, 1992] Minton, A. P. (1992). Confinement as a determinant of macromolecular structure and reactivity. *Biophys J*, 63(4) :1090–100.
- [Minton, 2001] Minton, A. P. (2001). The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media. *J Biol Chem*, 276(14) :10577–80.
- [Mogk et al., 2003] Mogk, A., Schlieker, C., Friedrich, K. L., Schonfeld, H. J., Vierling, E., and Bukau, B. (2003). Refolding of substrates bound to small hsp80s relies on a disaggregation reaction mediated most efficiently by clpb/dnak. *J Biol Chem*, 278(33) :31033–42.
- [Monsellier and Chiti, 2007] Monsellier, E. and Chiti, F. (2007). Prevention of amyloid-like aggregation as a driving force of protein evolution. *EMBO Rep*, 8(8) :737–42.
- [Morell et al., 2008] Morell, M., Bravo, R., Espargaro, A., Sisquella, X., Aviles, F. X., Fernandez-Busquets, X., and Ventura, S. (2008). Inclusion bodies : specificity in their aggregation process and amyloid-like structure. *Biochim Biophys Acta*, 1783(10) :1815–25.
- [Mortimer and Johnston, 1959] Mortimer, R. K. and Johnston, J. R. (1959). Life span of individual yeast cells. *Nature*, 183(4677) :1751–2.
- [Muramatsu and Minton, 1988] Muramatsu, N. and Minton, A. P. (1988). Tracer diffusion of globular proteins in concentrated protein solutions. *Proc Natl Acad Sci U S A*, 85(9) :2984–8.
- [Nagai et al., 2002] Nagai, T., Ibata, K., Park, E. S., Kubota, M., Mikoshiba, K., and Miyawaki, A. (2002). A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat Biotechnol*, 20(1) :87–90.
- [Nenninger et al., 2010] Nenninger, A., Mastroianni, G., and Mullineaux, C. W. (2010). Size dependence of protein diffusion in the cytoplasm of escherichia coli. *J Bacteriol*, 192(18) :4535–40.
- [Nishihara et al., 1998] Nishihara, K., Kanemori, M., Kitagawa, M., Yanagi, H., and Yura, T. (1998). Chaperone coexpression plasmids : differential and synergistic roles of dnk-dnaj-grpe and groel-groes in assisting folding of an allergen of japanese cedar pollen, cryj2, in escherichia coli. *Appl Environ Microbiol*, 64(5) :1694–9.
- [Nollen et al., 2004] Nollen, E. A., Garcia, S. M., van Haften, G., Kim, S., Chavez, A., Morimoto, R. I., and Plasterk, R. H. (2004). Genome-wide rna interference screen identifies previously undescribed regulators of polyglutamine aggregation. *Proc Natl Acad Sci U S A*, 101(17) :6403–8.
- [Nucifora et al., 2001] Nucifora, F. C., J., Sasaki, M., Peters, M. F., Huang, H., Cooper, J. K., Yamada, M., Takahashi, H., Tsuji, S., Troncoso, J., Dawson, V. L., Dawson, T. M., and Ross, C. A. (2001). Interference by huntingtin and atrophin-1 with cbp-mediated transcription leading to cellular toxicity. *Science*, 291(5512) :2423–8.
- [Nussbaum et al., 2012] Nussbaum, J. M., Schilling, S., Cynis, H., Silva, A., Swanson, E., Wangsanut, T., Tayler, K., Wiltgen, B., Hatami, A., Ronicke, R., Reymann, K., Hutter-Paier, B., Alexandru, A., Jagla, W., Graubner, S., Glabe, C. G., Demuth, H. U., and Bloom, G. S. (2012). Prion-like behaviour and tau-dependent cytotoxicity of pyroglutamylated amyloid-beta. *Nature*, 485(7400) :651–5.
- [Nystrom, 2005] Nystrom, T. (2005). Role of oxidative carbonylation in protein quality control and senescence. *EMBO J*, 24(7) :1311–7.

- [Odiijk, 1998] Odiijk, T. (1998). Osmotic compaction of supercoiled dna into a bacterial nucleoid. *Biophys Chem*, 73(1-2) :23–9.
- [Olzscha et al., 2011] Olzscha, H., Schermann, S. M., Woerner, A. C., Pinkert, S., Hecht, M. H., Tartaglia, G. G., Vendruscolo, M., Hayer-Hartl, M., Hartl, F. U., and Vabulas, R. M. (2011). Amyloid-like aggregates sequester numerous metastable proteins with essential cellular functions. *Cell*, 144(1) :67–78.
- [Pan et al., 1993] Pan, K. M., Baldwin, M., Nguyen, J., Gasset, M., Serban, A., Groth, D., Mehlhorn, I., Huang, Z., Fletterick, R. J., Cohen, F. E., and et al. (1993). Conversion of alpha-helices into beta-sheets features in the formation of the scrapie prion proteins. *Proc Natl Acad Sci U S A*, 90(23) :10962–6.
- [Parsell et al., 1994] Parsell, D. A., Kowal, A. S., Singer, M. A., and Lindquist, S. (1994). Protein disaggregation mediated by heat-shock protein hsp104. *Nature*, 372(6505) :475–8.
- [Partikian et al., 1998] Partikian, A., Olveczky, B., Swaminathan, R., Li, Y., and Verkman, A. S. (1998). Rapid diffusion of green fluorescent protein in the mitochondrial matrix. *J Cell Biol*, 140(4) :821–9.
- [Partridge, 2011] Partridge, L. (2011). Some highlights of research on aging with invertebrates, 2010. *Aging Cell*, 10(1) :5–9.
- [Paul, 2007] Paul, G. (2007). *Causes and consequences of physiological constraints on evolutionary models. New insights on the evolution of ageing and heredity*. PhD thesis, Universite Pierre et Marie curie.
- [Pearson, 1905] Pearson, K. (1905). The problem of the random walk. *Nature*, 72(1867) :342.
- [Perrin, 1913] Perrin, J. (1913). *Les atomes*. Felix Alcan, Paris.
- [Polymenidou and Cleveland, 2012] Polymenidou, M. and Cleveland, D. W. (2012). Prion-like spread of protein aggregates in neurodegeneration. *J Exp Med*, 209(5) :889–93.
- [Prasher et al., 1992] Prasher, D. C., Eckenrode, V. K., Ward, W. W., Prendergast, F. G., and Cormier, M. J. (1992). Primary structure of the aequorea victoria green-fluorescent protein. *Gene*, 111(2) :229–33.
- [Primet, 2011] Primet, M. (2011). *Probabilistic methods for point tracking and biological image analysis*. PhD thesis, Paris Descartes.
- [Prusiner, 1982] Prusiner, S. B. (1982). Novel proteinaceous infectious particles cause scrapie. *Science*, 216(4542) :136–44.
- [Qin and Lu, 2006] Qin, H. and Lu, M. (2006). Natural variation in replicative and chronological life spans of *saccharomyces cerevisiae*. *Exp Gerontol*, 41(4) :448–56.
- [Quardokus et al., 1996] Quardokus, E., Din, N., and Brun, Y. V. (1996). Cell cycle regulation and cell type-specific localization of the *ftsZ* division initiation protein in *caulobacter*. *Proc Natl Acad Sci U S A*, 93(13) :6314–9.
- [Quintas et al., 2001] Quintas, A., Vaz, D. C., Cardoso, I., Saraiva, M. J., and Brito, R. M. (2001). Tetramer dissociation and monomer partial unfolding precedes protofibril formation in amyloidogenic transthyretin variants. *J Biol Chem*, 276(29) :27207–13.
- [Ramachandran and Sasisekharan, 1968] Ramachandran, G. N. and Sasisekharan, V. (1968). Conformation of polypeptides and proteins. *Adv Protein Chem*, 23 :283–438.

- [Ratajczak et al., 2010] Ratajczak, E., Strozecka, J., Matuszewska, M., Zietkiewicz, S., Kuczynska-Wisnik, D., Laskowska, E., and Liberek, K. (2010). IbpA the small heat shock protein from *Escherichia coli* forms fibrils in the absence of its cochaperone IbpB. *FEBS Lett*, 584(11) :2253–7.
- [Ratajczak et al., 2009] Ratajczak, E., Zietkiewicz, S., and Liberek, K. (2009). Distinct activities of *Escherichia coli* small heat shock proteins IbpA and IbpB promote efficient protein disaggregation. *J Mol Biol*, 386(1) :178–89.
- [Ritossa, 1962] Ritossa, F. (1962). A new puffing pattern induced by temperature shock and dnp in *Drosophila*. *Experientia*, 18 :571–573.
- [Rivas et al., 1999] Rivas, G., Fernandez, J. A., and Minton, A. P. (1999). Direct observation of the self-association of dilute proteins in the presence of inert macromolecules at high concentration via tracer sedimentation equilibrium : theory, experiment, and biological significance. *Biochemistry*, 38(29) :9379–88.
- [Rivas et al., 2001] Rivas, G., Fernandez, J. A., and Minton, A. P. (2001). Direct observation of the enhancement of noncooperative protein self-assembly by macromolecular crowding : indefinite linear self-association of bacterial cell division protein FtsZ. *Proc Natl Acad Sci U S A*, 98(6) :3150–5.
- [Rokney et al., 2009] Rokney, A., Shagan, M., Kessel, M., Smith, Y., Rosenshine, I., and Oppenheim, A. B. (2009). *E. coli* transports aggregated proteins to the poles by a specific and energy-dependent process. *J Mol Biol*, 392(3) :589–601.
- [Rosen et al., 2002] Rosen, R., Biran, D., Gur, E., Becher, D., Hecker, M., and Ron, E. Z. (2002). Protein aggregation in *Escherichia coli* : role of proteases. *FEMS Microbiol Lett*, 207(1) :9–12.
- [Rudin et al., 1992] Rudin, L. I., Osher, S., and Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D*, 60(1-4) :259–268.
- [Saberli and Emberly, 2010] Saberli, S. and Emberly, E. (2010). Chromosome driven spatial patterning of proteins in bacteria. *PLoS Comput Biol*, 6(11) :e1000986.
- [Sawaya et al., 2007] Sawaya, M. R., Sambashivan, S., Nelson, R., Ivanova, M. I., Sievers, S. A., Apostol, M. I., Thompson, M. J., Balbirnie, M., Wiltzius, J. J., McFarlane, H. T., Madsen, A. O., Riek, C., and Eisenberg, D. (2007). Atomic structures of amyloid cross- β spines reveal varied steric zippers. *Nature*, 447(7143) :453–7.
- [Saxton, 1994] Saxton, M. J. (1994). Anomalous diffusion due to obstacles : a monte carlo study. *Biophys J*, 66(2 Pt 1) :394–401.
- [Schmidt and Weiss, 2011] Schmidt, U. and Weiss, M. (2011). Anomalous diffusion of oligomerized transmembrane proteins. *J Chem Phys*, 134(16) :165101.
- [Schubert et al., 2000] Schubert, U., Anton, L. C., Gibbs, J., Norbury, C. C., Yewdell, J. W., and Bannik, J. R. (2000). Rapid degradation of a large fraction of newly synthesized proteins by proteasomes. *Nature*, 404(6779) :770–4.
- [Shaner et al., 2004] Shaner, N. C., Campbell, R. E., Steinbach, P. A., Giepmans, B. N., Palmer, A. E., and Tsien, R. Y. (2004). Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nat Biotechnol*, 22(12) :1567–72.
- [Shaner et al., 2008] Shaner, N. C., Lin, M. Z., McKeown, M. R., Steinbach, P. A., Hazelwood, K. L., Davidson, M. W., and Tsien, R. Y. (2008). Improving the photostability of bright monomeric orange and red fluorescent proteins. *Nat Methods*, 5(6) :545–51.

- [Shaner et al., 2005] Shaner, N. C., Steinbach, P. A., and Tsien, R. Y. (2005). A guide to choosing fluorescent proteins. *Nat Methods*, 2(12) :905–9.
- [Shcheprova et al., 2008] Shcheprova, Z., Baldi, S., Frei, S. B., Gonnet, G., and Barral, Y. (2008). A mechanism for asymmetric segregation of age during yeast budding. *Nature*, 454(7205) :728–34.
- [Silow and Oliveberg, 2003] Silow, M. and Oliveberg, M. (2003). High concentrations of viscosogens decrease the protein folding rate constant by prematurely collapsing the coil. *J Mol Biol*, 326(1) :263–71.
- [Silveira et al., 2005] Silveira, J. R., Raymond, G. J., Hughson, A. G., Race, R. E., Sim, V. L., Hayes, S. F., and Caughey, B. (2005). The most infectious prion protein particles. *Nature*, 437(7056) :257–61.
- [Slepenkov and Witt, 2002] Slepenkov, S. V. and Witt, S. N. (2002). The unfolding story of the escherichia coli hsp70 dnak : is dnak a holdase or an unfoldase? *Mol Microbiol*, 45(5) :1197–206.
- [Squire and Himmel, 1979] Squire, P. G. and Himmel, M. E. (1979). Hydrodynamics and protein hydration. *Arch Biochem Biophys*, 196(1) :165–77.
- [Stadtman and Levine, 2000] Stadtman, E. R. and Levine, R. L. (2000). Protein oxidation. *Ann N Y Acad Sci*, 899 :191–208.
- [Stefani and Dobson, 2003] Stefani, M. and Dobson, C. M. (2003). Protein aggregation and aggregate toxicity : new insights into protein folding, misfolding diseases and biological evolution. *J Mol Med (Berl)*, 81(11) :678–99.
- [Steffan et al., 2000] Steffan, J. S., Kazantsev, A., Spasic-Boskovic, O., Greenwald, M., Zhu, Y. Z., Gohler, H., Wanker, E. E., Bates, G. P., Housman, D. E., and Thompson, L. M. (2000). The huntington’s disease protein interacts with p53 and creb-binding protein and represses transcription. *Proc Natl Acad Sci U S A*, 97(12) :6763–8.
- [Steinkraus et al., 2008] Steinkraus, K. A., Kaeberlein, M., and Kennedy, B. K. (2008). Replicative aging in yeast : the means to the end. *Annu Rev Cell Dev Biol*, 24 :29–54.
- [Stewart et al., 2005] Stewart, E. J., Madden, R., Paul, G., and Taddei, F. (2005). Aging and death in an organism that reproduces by morphologically symmetric division. *PLoS Biol*, 3(2) :e45.
- [Stockel and Hartl, 2001] Stockel, J. and Hartl, F. U. (2001). Chaperonin-mediated de novo generation of prion protein aggregates. *J Mol Biol*, 313(4) :861–72.
- [Stokes, 1851] Stokes, G. (1851). On the effect of the internal friction of fluids on the motion of pendulums. *Proceedings Cambridge Philosophical Society*, 9 :8–106.
- [Surgucheva et al., 2012] Surgucheva, I., Sharov, V. S., and Surguchov, A. (2012). gamma-synuclein : Seeding of alpha-synuclein aggregation and transmission between cells. *Biochemistry*.
- [Swaminathan et al., 1997] Swaminathan, R., Hoang, C. P., and Verkman, A. S. (1997). Photo-bleaching recovery and anisotropy decay of green fluorescent protein gfp-s65t in solution and cells : cytoplasmic viscosity probed by green fluorescent protein translational and rotational diffusion. *Biophys J*, 72(4) :1900–7.
- [Swamy and Goldberg, 1981] Swamy, K. H. and Goldberg, A. L. (1981). E. coli contains eight soluble proteolytic activities, one being atp dependent. *Nature*, 292(5824) :652–4.

- [Tam et al., 2009] Tam, S., Spiess, C., Auyeung, W., Joachimiak, L., Chen, B., Poirier, M. A., and Frydman, J. (2009). The chaperonin tric blocks a huntingtin sequence element that promotes the conformational switch to aggregation. *Nat Struct Mol Biol*, 16(12) :1279–85.
- [Tartaglia et al., 2007] Tartaglia, G. G., Pechmann, S., Dobson, C. M., and Vendruscolo, M. (2007). Life on the edge : a link between gene expression levels and aggregation rates of human proteins. *Trends Biochem Sci*, 32(5) :204–6.
- [Tartaglia and Vendruscolo, 2010] Tartaglia, G. G. and Vendruscolo, M. (2010). Proteome-level interplay between folding and aggregation propensities of proteins. *J Mol Biol*, 402(5) :919–28.
- [Terry et al., 1995] Terry, B. R., Matthews, E. K., and Haseloff, J. (1995). Molecular characterisation of recombinant green fluorescent protein by fluorescence correlation microscopy. *Biochem Biophys Res Commun*, 217(1) :21–7.
- [Thomas and Baneyx, 1998] Thomas, J. G. and Baneyx, F. (1998). Roles of the escherichia coli small heat shock proteins ibpa and ibpb in thermal stress management : comparison with clpa, clpb, and htpg in vivo. *J Bacteriol*, 180(19) :5165–72.
- [Tomoyasu et al., 2001] Tomoyasu, T., Mogk, A., Langen, H., Goloubinoff, P., and Bukau, B. (2001). Genetic dissection of the roles of chaperones and proteases in protein folding and degradation in the escherichia coli cytosol. *Mol Microbiol*, 40(2) :397–413.
- [Tyedmers et al., 2010] Tyedmers, J., Mogk, A., and Bukau, B. (2010). Cellular strategies for controlling protein aggregation. *Nat Rev Mol Cell Biol*, 11(11) :777–88.
- [Uemura et al., 2008] Uemura, S., Iizuka, R., Ueno, T., Shimizu, Y., Taguchi, H., Ueda, T., Puglisi, J. D., and Funatsu, T. (2008). Single-molecule imaging of full protein synthesis by immobilized ribosomes. *Nucleic Acids Res*, 36(12) :e70.
- [Uptain and Lindquist, 2002] Uptain, S. M. and Lindquist, S. (2002). Prions as protein-based genetic elements. *Annu Rev Microbiol*, 56 :703–41.
- [Uversky et al., 2002] Uversky, V. N., Li, J., Bower, K., and Fink, A. L. (2002). Synergistic effects of pesticides and metals on the fibrillation of alpha-synuclein : implications for parkinson's disease. *Neurotoxicology*, 23(4-5) :527–36.
- [van den Berg et al., 1999] van den Berg, B., Ellis, R. J., and Dobson, C. M. (1999). Effects of macromolecular crowding on protein folding and aggregation. *EMBO J*, 18(24) :6927–33.
- [van den Ent et al., 2001] van den Ent, F., Amos, L. A., and Lowe, J. (2001). Prokaryotic origin of the actin cytoskeleton. *Nature*, 413(6851) :39–44.
- [van Ham et al., 2008] van Ham, T. J., Thijssen, K. L., Breitling, R., Hofstra, R. M., Plasterk, R. H., and Nollen, E. A. (2008). C. elegans model identifies genetic modifiers of alpha-synuclein inclusion formation during aging. *PLoS Genet*, 4(3) :e1000027.
- [Ventura, 2005] Ventura, S. (2005). Sequence determinants of protein aggregation : tools to increase protein solubility. *Microb Cell Fact*, 4(1) :11.
- [Verkman, 2002] Verkman, A. S. (2002). Solute and macromolecule diffusion in cellular aqueous compartments. *Trends Biochem Sci*, 27(1) :27–33.
- [Wachi et al., 1989] Wachi, M., Doi, M., Okada, Y., and Matsushashi, M. (1989). New mre genes mreC and mreD, responsible for formation of the rod shape of escherichia coli cells. *J Bacteriol*, 171(12) :6511–6.

- [Wang et al., 2008] Wang, L., Maji, S. K., Sawaya, M. R., Eisenberg, D., and Riek, R. (2008). Bacterial inclusion bodies contain amyloid-like structure. *PLoS Biol*, 6(8) :e195.
- [Wang et al., 2010] Wang, P., Robert, L., Pelletier, J., Dang, W. L., Taddei, F., Wright, A., and Jun, S. (2010). Robust growth of escherichia coli. *Curr Biol*.
- [Watve et al., 2006] Watve, M., Parab, S., Jogdand, P., and Keni, S. (2006). Aging may be a conditional strategic choice and not an inevitable outcome for bacteria. *Proc Natl Acad Sci U S A*, 103(40) :14831–5.
- [Wetzel, 2006] Wetzel, R. (2006). Kinetics and thermodynamics of amyloid fibril assembly. *Acc Chem Res*, 39(9) :671–9.
- [Wickner, 1994] Wickner, R. B. (1994). [ure3] as an altered ure2 protein : evidence for a prion analog in saccharomyces cerevisiae. *Science*, 264(5158) :566–9.
- [Wickner et al., 1999] Wickner, S., Maurizi, M. R., and Gottesman, S. (1999). Posttranslational quality control : folding, refolding, and degrading proteins. *Science*, 286(5446) :1888–93.
- [Williams et al., 1982] Williams, D. C., Van Frank, R. M., Muth, W. L., and Burnett, J. P. (1982). Cytoplasmic inclusion bodies in escherichia coli producing biosynthetic human insulin proteins. *Science*, 215(4533) :687–9.
- [Winkler et al., 2010] Winkler, J., Seybert, A., König, L., Pruggnaller, S., Haselmann, U., Sourjik, V., Weiss, M., Frangakis, A. S., Mogk, A., and Bukau, B. (2010). Quantitative and spatio-temporal features of protein aggregation in escherichia coli and consequences on protein quality control and cellular ageing. *EMBO J*, 29(5) :910–23.
- [Woldringh and Odijk, 1999] Woldringh, C. L. and Odijk, T. (1999). *Structure of dna within the bacterial cell : physics and physiology*. American Society of Microbiology Press.
- [Wolynes et al., 1995] Wolynes, P. G., Onuchic, J. N., and Thirumalai, D. (1995). Navigating the folding routes. *Science*, 267(5204) :1619–20.
- [Xue et al., 2010] Xue, W. F., Hellewell, A. L., Hewitt, E. W., and Radford, S. E. (2010). Fibril fragmentation in amyloid assembly and cytotoxicity : when size matters. *Prion*, 4(1) :20–5.
- [Yildiz et al., 2003] Yildiz, A., Forkey, J. N., McKinney, S. A., Ha, T., Goldman, Y. E., and Selvin, P. R. (2003). Myosin v walks hand-over-hand : single fluorophore imaging with 1.5-nm localization. *Science*, 300(5628) :2061–5.
- [Yu et al., 2006] Yu, J., Xiao, J., Ren, X., Lao, K., and Xie, X. S. (2006). Probing gene expression in live cells, one protein molecule at a time. *Science*, 311(5767) :1600–3.
- [Zhou et al., 2011] Zhou, C., Slaughter, B. D., Unruh, J. R., Eldakak, A., Rubinstein, B., and Li, R. (2011). Motility and segregation of hsp104-associated protein aggregates in budding yeast. *Cell*, 147(5) :1186–96.
- [Zimmerman, 2006] Zimmerman, S. B. (2006). Shape and compaction of escherichia coli nucleoids. *J Struct Biol*, 156(2) :255–61.
- [Zimmerman and Trach, 1991] Zimmerman, S. B. and Trach, S. O. (1991). Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of escherichia coli. *J Mol Biol*, 222(3) :599–620.

BIBLIOGRAPHIE
