



**HAL**  
open science

# Repliement des protéines et formation de fibres amyloïdes. Le cas de l'alpha-lactalbumine

Clement Blanchet

► **To cite this version:**

Clement Blanchet. Repliement des protéines et formation de fibres amyloïdes. Le cas de l'alpha-lactalbumine. Biophysique [physics.bio-ph]. Université Joseph-Fourier - Grenoble I, 2008. Français. NNT: . tel-00327686

**HAL Id: tel-00327686**

**<https://theses.hal.science/tel-00327686>**

Submitted on 9 Oct 2008

**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.

THESE DE DOCTORAT DE L'UNIVERSITE JOSEPH FOURIER

- ECOLE DOCTORALE DE PHYSIQUE -

Spécialité : Physique pour les Sciences du Vivant

Présentée par Clément BLANCHET

# **Repliement des protéines et formation de fibres amyloides.**

Le cas de l' $\alpha$ -lactalbumine

Directeurs de thèse : Dr Vincent FORGE, Dr Florent GUILLAIN

Jury :

Pr Franz BRUCKERT

Président

Pr Patrice GOUET

Rapporteur

Dr Saïd BOUHALLAB

Rapporteur

Dr Stéphanie FINET

Examineur

Dr Florent GUILLAIN

Examineur

Dr Vincent FORGE

Examineur

## Remerciements

Je tiens à remercier en tout premier lieu Vincent Forge qui a dirigé cette thèse dans la continuité de mon stage de master. Tout au long de ce doctorat, il a su partager ses connaissances en biophysique et son expérience de recherche. Il a su être présent lorsque j'avais besoin de ses conseils éclairés et je le remercie pour son soutien et les quatre agréables années passées à ses côtés.

Je remercie Florent Guillain, co-directeur de thèse et directeur du laboratoire à l'époque où j'ai commencé ce travail, de m'avoir accueilli au sein de son laboratoire. Je lui suis reconnaissant de s'être montré très disponible lorsque j'avais besoin de lui.

Je remercie les membres de mon jury.

Le Professeur Franz Bruckert, qui m'a suivi pendant une bonne partie de ma scolarité supérieure et qui m'a fait l'honneur de présider le jury de cette thèse.

Le Professeur Patrice Gouet pour ses pertinentes remarques de structuraliste. Je le remercie aussi pour la très agréable collaboration effectuée pour les expériences de diffusion quasi-élastique de lumière.

Le Docteur Saïd Bouhallab dont la connaissance de la biophysique en général et de l' $\alpha$ -lactalbumine en particulier m'ont permis de comprendre encore mieux l'objet de mon travail à travers ses judicieuses remarques.

Le Docteur Stéphanie Finet qui a acceptée de participer à mon jury de thèse. Je tiens aussi à remercier Stéphanie pour sa collaboration pour les expériences de diffusion de rayons X aux petits angles ; elle a été d'une grande aide pour les expériences mais aussi, et surtout, pour le traitement des données qui s'est avéré assez délicat.

Cette thèse n'aurait été possible sans les personnes suivantes avec qui j'ai collaboré :

Jean-Luc Garden et Hervé Guillou du CNRS Grenoble pour les expériences de calorimétrie.

Frédéric Pignon du Laboratoire de Rhéologie de Grenoble.

Paul Schanda et Bernard Brutscher de l'IBS Grenoble pour les expériences de RMN.

Je remercie toutes les personnes que j'ai pu croiser au cours de ces quatre années et en particulier les membres du groupe « modélisation interaction repliement » ; Serge pour son aide en ce qui concerne l'informatique ; Michel qui fut le premier à m'encourager à venir faire une thèse dans ce groupe et qui sait à merveille partager son énergie et son enthousiasme ; et toutes les personnes que j'ai pu croiser, qui ont supporté mes méditations métaphysiques et m'ont apporté les petits coup de main du quotidien : Nathalia, Grégory, Charlotte, Caro, Babeth, Nicolas, Thomas, Laurence, Fanny, Hélène... et tous les autres. Un merci tout particulier à Heidi qui a relu cette thèse.

Enfin, je remercie ma famille et mes amis qui m'ont toujours supporté dans ce travail et dans mes études en générale. Un grand merci en particulier ma mère, profane en science, qui a relu les 250 pages de cette thèse pour en vérifier l'orthographe.

# Tables des matières

## I. Introduction

### I.0. Protéines

#### I.0.1. Généralités

#### I.0.2. Structure des protéines

### I.1. Repliement des protéines

#### I.1.1. Genèse du repliement des protéines

#### I.1.2. Développements expérimentaux

#### I.1.3. Simulation

#### I.1.4. Point de vue actuel

### I.2. Cofacteurs

#### I.2.1. Les coenzymes

#### I.2.2. Les cofacteurs métalliques

### I.3. Fibres amyloïdes

#### I.3.1. Quelques maladies dans lesquelles sont impliquées les fibres amyloïdes

#### I.3.2. Protéines susceptibles de former des fibres amyloïdes

#### I.3.3. Structure des fibres

#### I.3.4. Formation des fibres amyloïdes

### I.4. $\alpha$ -lactalbumine

#### I.4.1. Fonctions

#### I.4.2. Structure

#### I.4.3. L'état *Molten Globule*

#### I.4.4. Repliement

## II. Repliement de l' $\alpha$ -lactalbumine en présence de métaux

## II.1. Calcul des constantes de vitesse de repliement et des constantes de dissociation

### II.1.1. Repliement suivi par spectroscopie de fluorescence

### II.1.2. Courbes de liaisons

### II.1.3. Constantes de dissociation

## II.2. Comparaison des variations d'énergie des différents états et implications pour le repliement

### II.2.1. Equation de Brønsted

### II.2.2. Analyse des valeurs $\Phi$

## II.3. Etude des paramètres thermodynamique de la réaction de repliement

### II.3.1. Variation d'enthalpie libre

### II.3.2. Variation d'enthalpie – Courbes d'Eyring

### II.3.3. Variation d'entropie

### II.3.4. Résultats – Interprétation

## II.4. Dénaturation thermique

### II.4.1. Dénaturation thermique suivie par dichroïsme circulaire

### II.4.2. Dénaturation thermique suivie par calorimétrie différentielle à balayage

### II.4.3. Discussion

## III. Repliement de l' $\alpha$ -lactalbumine en absence de métaux

### III.1. Rhéologie

#### III.1.1. Description de l'expérience

#### III.1.2. Mesures à l'équilibre

#### III.1.3. Mesures en cours de cinétique

#### III.1.4. Interprétation et discussion

### III.2. Dichroïsme circulaire

#### III.2.1. Conditions expérimentales – Résultats

- III.2.2. Interprétation – Construction d'un modèle de réaction
- III.3. Spectres de corrélation 2D
  - III.3.1. Principe
  - III.3.2. Exemples
- III.4. Résonance Magnétique Nucléaire du proton
  - III.4.1. Description de l'expérience
  - III.4.2. Spectres 1D
  - III.4.3. Spectres de corrélation 2D
  - III.4.4. Implications pour le modèle de réaction
- III.5. Spectroscopie infrarouge
  - III.5.1. Description de l'expérience
  - III.5.2. Spectres 1D
  - III.5.3. Spectres de corrélation 2D
  - III.5.4. Implications pour le modèle de réaction
- III.6. Chromatographie d'exclusion de taille
  - III.6.1. Description de l'expérience
  - III.6.2. Résultats
- III.7. Diffusion quasi élastique de lumière
  - III.7.1. Description de l'expérience
  - III.7.2. Résultats et interprétation
- III.8. Etude de la séquence de l' $\alpha$ -lactalbumine
  - III.8.1. Méthodes utilisées
  - III.8.2. Résultats
- III.9. Synthèse, discussion
- IV. Formation de fibres amyloïdes

#### IV.1. Formation de fibres amyloïdes à partir de l' $\alpha$ -lactalbumine

##### IV.1.1. Microscopie électronique

##### IV.1.2. Formation de fibres suivie par fluorescence de la thioflavine T

##### IV.1.3. Spectroscopie infrarouge

#### IV.2. Effets du sel dans les solutions d' $\alpha$ -lactalbumine dans l'état *molten globule*

##### IV.2.1. Dénaturation thermique du *molten globule*

##### IV.2.2. Interactions protéine-protéine

##### IV.2.3. Diffusion de rayons X aux petits angles

#### IV.3. Rhéologie

#### IV.4. Résumé – Discussion

#### V. Conclusion, perspectives

# **I. Introduction**

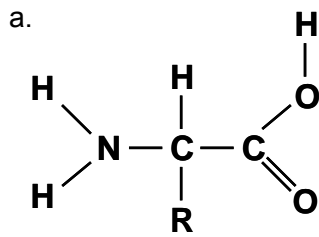


## **I.0. Protéines**

### **I.0.1. Généralités**

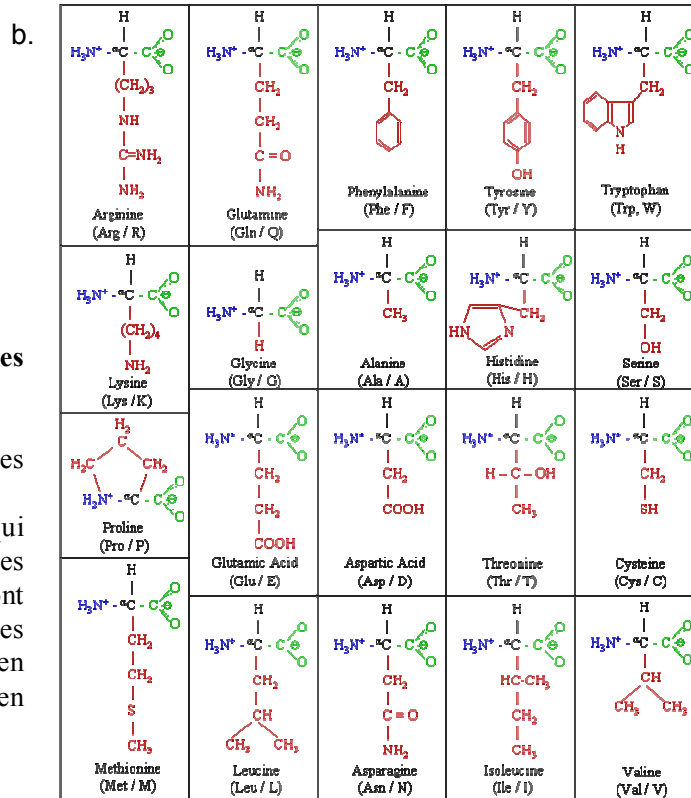
Les protéines sont découvertes à la fin des années 1830 par un chimiste hollandais, Gerhard Mulder<sup>1</sup>. Le terme « protéine » vient du grec « prôtos » qui signifie « premier, essentiel ». En effet, ces macromolécules sont indispensables à la vie et elles sont très abondantes dans les organismes vivants : elles représentent la moitié du poids sec des cellules<sup>2</sup>.

Sur le plan moléculaire, les protéines sont des polymères d'acides aminés. Ces derniers sont des molécules composées d'un squelette carboné et contenant une fonction amine et une fonction acide carboxylique. Lorsque le groupe carboxyle et le groupe amine sont portés par le même carbone, on parle d'acide aminé  $\alpha$  (fig I.1). Il existe plus d'une centaine de ces molécules dans la nature. Cependant, seulement vingt acides aminés sont utilisés pour former l'ensemble des protéines. Ils diffèrent par leurs chaînes latérales, notées R dans la figure I.1a et représentées en détail dans la figure I.1b. Les propriétés des chaînes latérales permettent de classer les acides aminés en quatre groupes : acide, basique, hydrophile, hydrophobe. Au sein des protéines, ils sont reliés entre eux par des liaisons peptidiques qui sont des liaisons covalentes entre le carbone du groupe carboxyle et l'azote du groupe amine. Toutes les protéines, aussi différentes soient-elles, sont des polymères formés par combinaison de ces vingt acides aminés. Leurs caractéristiques sont donc déterminées par le nombre et l'ordre des acides aminés dans la chaîne. Ces paramètres sont codés dans l'ADN et constituent la structure primaire (ou séquence) de la protéine.



**Fig. I.1:** Structure des acides aminés

- a. Structure générale des acides aminés  $\alpha$ .
- b. Les vingt acides aminés qui composent les protéines. Les fonctions amines sont représentées en bleu, les fonctions acide carboxylique en vert et les chaînes latérales en rouge.



© Magnus Manske

La synthèse des protéines à partir de l'ADN est réalisée en deux étapes :

- Dans un premier temps, l'ADN contenu dans le noyau de la cellule est transcrit en ARN messenger. Cette étape permet notamment d'éviter la détérioration de l'ADN qu'une traduction directe pourrait engendrer. Elle permet aussi d'avoir plusieurs copies de l'ADN à traduire, et donc de produire plus rapidement des protéines.
- Dans un second temps, l'ARN messenger est mûré et rejoint le cytoplasme dans lequel son association avec un ribosome permettra sa traduction en protéine. Le ribosome lit l'ARN messenger et assemble un à un les acides aminés pour former la chaîne polypeptidique.

Les protéines sont utilisées par les organismes vivants pour effectuer une multitude de fonctions :

- Catalyse : La grande majorité des enzymes sont des protéines. Elles permettent d'accélérer, jusqu'à plusieurs millions de fois, les réactions chimiques du métabolisme dans des conditions de température et de pression propices à la vie.
- Transport : L'hémoglobine, par exemple, permet de distribuer l'oxygène en provenance des poumons dans tout l'organisme. Les protéines sont aussi impliquées dans le transport de métabolites à travers les membranes ; on parle de canaux ioniques lorsque l'énergie du transport est fournie par le gradient électrochimique, et de pompes lorsque les protéines transporteuses ont besoin d'une source d'énergie supplémentaire comme l'hydrolyse de l'ATP.
- Communication : De nombreuses hormones comme l'insuline permettent de délivrer un message dans tout l'organisme. D'autres protéines ont une fonction de récepteur aux hormones.
- Reconnaissance : Les anticorps du système immunitaire sont des protéines qui permettent la reconnaissance du soi et du non-soi.
- Mobilité cellulaire : L'actine et la myosine sont à l'origine du mouvement cellulaire lors de la contraction du muscle. Elles sont aussi impliquées dans la mitose. D'autres protéines composent les flagelles qui permettent la mobilité des spermatozoïdes et de certaines bactéries.
- Structure: La forme des cellules ainsi que leur résistance aux contraintes physiques extérieures sont assurées par le collagène et les protéines du cytosquelette.

## I.0.2. Structure des protéines

### Introduction

Comme nous venons de le voir, les protéines constituent un ensemble de macromolécules indispensables au bon fonctionnement des organismes vivants. Pour pouvoir accomplir leurs fonctions biologiques, les protéines doivent se replier pour acquérir une structure tridimensionnelle stable. La fonction d'une protéine est intrinsèquement liée à sa structure.

Les atomes qui composent la protéine interagissent entre eux, et avec le solvant qui entoure la macromolécule. Ces interactions vont contraindre la chaîne à se replier et à adopter une structure tridimensionnelle unique et stable. Cette structure est appelée « état natif » de la protéine. C'est cette conformation qui va lui permettre de réaliser la fonction biologique pour laquelle elle est produite.

Pour décrire cette conformation, différents niveaux de structure sont utilisés:

- *La structure primaire* décrit les atomes et les liaisons covalentes qui composent la protéine. Les protéines sont des polymères linéaires, leurs structures primaires sont donc souvent assimilées à leurs séquences. Plus rigoureusement, la séquence d'une protéine, contrairement à sa structure primaire, ne renseigne pas sur les ponts disulfures qui peuvent être formés entre les soufres des cystéines et sur les autres modifications post-traductionnelles qui peuvent avoir lieu.
- *La structure secondaire* précise les différentes conformations locales de la chaîne polypeptidique. Deux principaux types de structure secondaire sont distingués : les hélices  $\alpha$  et les feuillets  $\beta$ . Ils seront décrits plus longuement par la suite.

- *La structure tertiaire* spécifie l'arrangement des différentes structures secondaires entre elles. Elle donne donc la forme globale de la protéine, les coordonnées de chacun de ses atomes dans un espace tridimensionnel.
- *La structure quaternaire* décrit, lorsque la protéine fonctionnelle est composée de plusieurs chaînes polypeptidiques, l'arrangement de ces différentes chaînes entre elles.

## Structure primaire

### *Historique*

L'idée selon laquelle les protéines contiennent des liaisons peptidiques est lancée par un chimiste français, Edouard Grimaux à la fin du 19<sup>ème</sup> siècle<sup>3</sup>. En 1902, au cours du 74<sup>ème</sup> congrès des physiciens et scientifiques allemands, deux intervenants émettent l'hypothèse que les protéines sont des chaînes d'acides aminés. Franz Hofmeister en a l'intuition en observant une réaction du biuret dans les protéines<sup>4</sup>. Une réaction du biuret correspond à la formation de complexes, à pH basique, entre des ions  $\text{Cu}^{2+}$  et des liaisons peptidiques. Quelques heures plus tard, Emil Fischer apporte une multitude de détails chimiques confirmant cette hypothèse<sup>5</sup>.

La description des protéines comme des longues chaînes d'acides aminés est alors longuement débattue, la communauté scientifique de l'époque n'admettant pas que des liaisons covalentes puissent être suffisamment solides pour former des chaînes aussi longues.

Plusieurs hypothèses alternatives existent à l'époque. Il est proposé, par exemple que les protéines sont des amas de petites molécules. Cette idée est écartée dans les années 20 grâce aux expériences d'ultracentrifugation de Theodore Svedberg et d'électrophorèse d'Arne Tiselius<sup>6</sup>. Dorothy Wrinch propose que la chaîne polypeptidique subit des changements, ce

qui lui permet de former des liaisons covalentes entre les différentes parties de la chaîne<sup>7</sup>. Ces descriptions erronées sont réfutées après le séquençage de l'insuline par Frederick Sanger<sup>8</sup> en 1952 et la résolution de la structure tridimensionnelle de la myoglobine par Max Perutz<sup>9</sup> et John Kendrew<sup>10</sup> en 1958.

### *Intérêt*

La structure primaire d'une protéine est d'un grand intérêt. En effet, elle permet de connaître tous les atomes de la protéine et toutes les liaisons covalentes entre ces atomes. On verra, dans la suite de l'introduction, que le phénomène de repliement des protéines est un processus spontané réversible. La forme native de la protéine correspond donc à la conformation de minimum d'énergie que peut adopter cette molécule. En conséquence, les structures secondaire et tertiaire de la protéine ainsi que la manière dont ces structures sont formées dépendent uniquement des interactions entre les différents atomes de la protéine et avec le solvant. Elles doivent donc pouvoir être prédites à partir de la structure primaire. Toute l'information nécessaire au repliement de la protéine, et donc à la formation des structures secondaire et tertiaire, est contenue dans la structure primaire.

### Structure secondaire

#### *Définition*

Sous le terme « structure secondaire » sont rassemblées les structures formées localement par la chaîne polypeptidique. Elles sont dues à la formation de liaisons hydrogènes entre les

atomes d'oxygène du groupe carboxyle et les hydrogènes du groupe amine de la chaîne polypeptidique.

### *Historique*

Au début des années 1930, des études de diffraction de rayons X par des cheveux et de la laine sont réalisées par William Astbury<sup>11</sup>. Les cheveux et la laine sont composés d'une protéine, la kératine. Les résultats de ces expériences conduisent Astbury à supposer l'existence d'hélices et de feuillets à l'intérieur des protéines et c'est à lui qu'on doit la nomenclature hélice  $\alpha$ , feuillet  $\beta$ . Cependant, les connaissances de l'époque ne lui permettent pas de proposer un modèle « solide ».

Linus Pauling et ses collaborateurs développent l'idée d'Astbury et publient, dans une série d'article de 1951, les structures secondaires telles qu'on les connaît aujourd'hui<sup>12,13</sup>. A partir de considérations géométriques (distances inter atomiques, angles entre les liaisons covalentes, ...) et en supposant que les oxygènes du groupe carboxyle et les hydrogènes du groupe amine de la chaîne polypeptidique sont impliqués dans des liaisons hydrogènes, Pauling et ses collaborateurs prédisent l'existence de ces structures alors qu'aucune n'a alors été observée.

## *Différents types de structure secondaire*

### - Hélices $\alpha$

Les structures secondaires les plus répandues dans les protéines sont les hélices  $\alpha$  (Fig I.2a). Dans ces structures, les liaisons hydrogènes sont créées entre l'oxygène du groupe carboxyle du résidu N et l'hydrogène du groupe amide du résidu N+4. Ces liaisons confèrent à la chaîne polypeptidique une structure en hélice avec 3,6 résidus par tour.

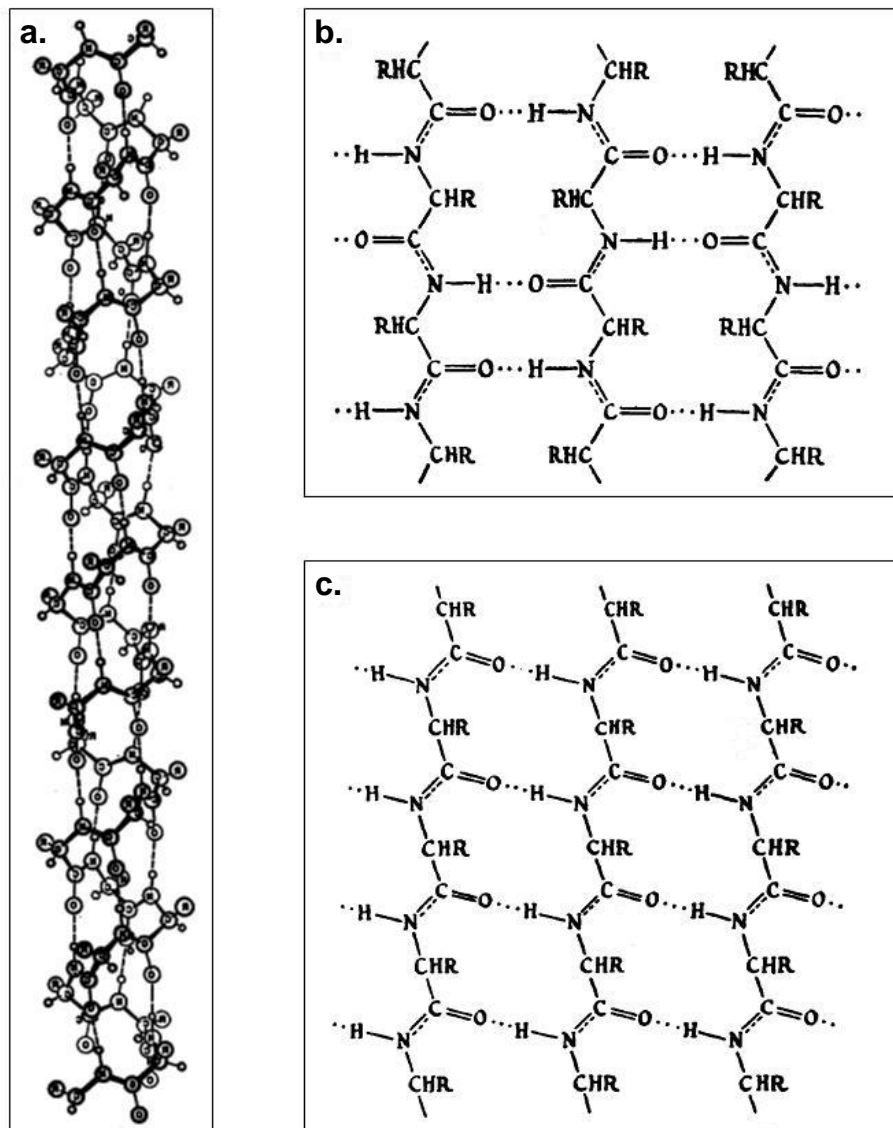
On peut expliquer l'abondance des hélices  $\alpha$  dans les protéines par le fait que les résidus impliqués dans les liaisons hydrogènes sont très proches dans la séquence de la protéine (3 résidus entre les résidus impliqués dans les liaisons hydrogènes). Ce n'est pas le cas pour les feuillets  $\beta$ .

### - Feuillet $\beta$

Les feuillets  $\beta$  (Fig I.2 b et c) sont aussi des structures secondaires très répandues. Les feuillets  $\beta$  sont composés de brins  $\beta$  qui sont des portions de la chaîne polypeptidique, généralement d'une longueur comprise entre 5 et 10 résidus, dans une configuration étendue. Ces brins  $\beta$  sont reliés entre eux et donc stabilisés par des liaisons hydrogènes. La chaîne polypeptidique est orientée : à une extrémité (N-ter) de la chaîne se trouve une fonction amine, à l'autre extrémité (C-ter) se trouve une fonction acide carboxylique. Les séquences des protéines sont par convention orientées Nter→Cter. Les feuillets  $\beta$  sont donc eux aussi orientés. On parle de feuillets  $\beta$  parallèles (Fig I.2c) lorsque les brins qui composent le feuillet ont la même orientation. Le feuillet  $\beta$  est dit antiparallèle (Fig I.2b) lorsque l'orientation des



brins est alternée. Les feuillets antiparallèles sont les plus répandus. Dans ce cas, les liaisons hydrogènes sont dans le plan du feuillet, ce qui stabilise la structure.



**Fig 1.2 :** Structures secondaires formées par la chaîne polypeptidique

- a. Hélices  $\alpha$
- b. Feuilletts  $\beta$  anti-parallèles
- c. Feuilletts  $\beta$  parallèles

Ces figures sont extraites des articles de Pauling et al. parus en 1951. Ce sont les premières représentations de structure secondaire.

## - Autres structures secondaires

D'autres structures secondaires, moins abondantes que les hélices  $\alpha$  et les feuillets  $\beta$ , peuvent être formées. Les hélices  $3_{10}$  par exemple sont des hélices stabilisées par des liaisons hydrogènes entre les résidus N et N+3 et comptent 3 résidus par tour. Les hélices  $\pi$  sont stabilisées par des liaisons entre les résidus N et N+5 et comptent 4,1 résidus par tour.

### *Prédiction*

La première génération de méthodes de prédiction de structure secondaire apparaît dans les années 60 et 70. Elle est basée sur les propriétés physico-chimiques des acides aminés considérés individuellement. La seconde génération de méthodes, utilisée jusqu'au début des années 1990, utilise des segments de 3 à 51 acides aminés. La structure secondaire de l'acide aminé au centre du segment est prédite grâce à un algorithme qui prend en considération différents facteurs tels que les propriétés physico-chimiques, des aspects statistiques, des séquences modèles, des éléments de la théorie des graphes, ... Ces méthodes de première et deuxième génération permettent d'atteindre seulement 60% d'exactitude dans la prédiction de structures secondaires<sup>14</sup>.

Les méthodes de troisième génération, basées sur l'alignement de séquence<sup>14</sup> dépassent les 70% d'exactitude. Les structures secondaires se conservent à travers l'évolution, les protéines d'une même famille ont des structures secondaires proches même si elles ne possèdent que 35% de similarité<sup>15</sup>. Ainsi, l'alignement de séquence permet d'inclure le poids, l'information de l'évolution dans la prédiction. Les méthodes actuelles ont une exactitude d'environ 77% et leur efficacité continue d'augmenter. Cependant, on estime qu'il sera difficile d'atteindre des

taux d'exactitude supérieurs à 88% car la formation de structure secondaire est aussi influencée par des interactions de longues portées et par l'environnement de la protéine<sup>14</sup>.

### Structure tertiaire

La structure tertiaire d'une protéine est sa structure tridimensionnelle définie par les coordonnées de chacun de ses atomes. Elle décrit donc l'arrangement dans l'espace des structures secondaires ainsi que l'orientation des chaînes latérales des acides aminés. La formation de la structure tertiaire est principalement liée à l'enfouissement des parties hydrophobes au cœur de la protéine. Cependant, des liaisons hydrogènes, des interactions ioniques et les ponts disulfures peuvent aussi stabiliser la structure de la protéine. Plus généralement, la structure tertiaire est stabilisée par toutes les interactions non covalentes en plus de celles considérées dans la structure secondaire.

La connaissance de la structure tertiaire d'une protéine permet de connaître la position de ses atomes dans l'espace. La fonction d'une protéine est intrinsèquement liée à sa structure. La détermination de la structure tertiaire est donc d'une importance capitale.

#### *Historique et méthode de détermination de la structure tertiaire*

La première structure tridimensionnelle d'une protéine déterminée expérimentalement est celle de la myoglobine obtenue par Max Perutz<sup>9</sup> et John Kendrew<sup>10</sup> en 1958. Cette structure a été déterminée par diffraction de rayons X sur un cristal de protéine. Cette méthode reste toujours la plus utilisée de nos jours : 90% des structures de protéine ont été résolues par diffraction. La structure tertiaire est aussi accessible par RMN (Résonance Magnétique Nucléaire). Développés initialement par Wüthrich<sup>16</sup> dans les années 80, les méthodes de

détermination de structure tertiaire par RMN présentent l'avantage de renseigner sur la dynamique des protéines en solution.

### *Prédiction*

Tout comme la structure secondaire, la structure tertiaire résulte de liaisons hydrogènes entre les différents atomes de la protéine qui sont décrits dans la structure primaire. La structure tertiaire devrait pouvoir être déduite de la structure primaire.

On distingue deux types de méthode de prédiction :

- Des méthodes « *ab initio* » qui utilisent uniquement la structure primaire de la protéine pour essayer de déterminer la structure secondaire. Cette approche nécessite beaucoup de temps de calcul et n'est en général utilisée qu'avec des superordinateurs comme Blue Gene ou Riden MDGRAPE-3 ou avec des méthodes de calcul distribué, comme c'est le cas avec les projets Folding@home et Rosetta@home<sup>17</sup>.
- Des méthodes par comparaison utilisent la structure primaire, mais aussi des structures tertiaires déjà existantes. Les structures utilisées pour la prédiction peuvent être celles de protéines de la même famille (*homology modelling*)<sup>18</sup>. Une autre méthode consiste à insérer la séquence de la protéine dans les structures déjà connues et à assigner à ces structures un score caractéristique de sa compatibilité avec la séquence à étudier (*protein threading*)<sup>19</sup>.

Les méthodes *ab initio* ne sont pas encore très efficaces et nécessitent beaucoup de temps de calcul. Les méthodes par comparaison donnent de meilleurs résultats, mais elles ne permettent pas de déterminer toutes les structures des protéines et ont besoin de structures connues pour fonctionner.

## **I.1. Repliement des protéines**

Le repliement des protéines est le processus physique par lequel les chaînes polypeptidiques acquièrent leurs structures tridimensionnelles. Les protéines sont synthétisées par le ribosome à partir de l'ARN messager. Elles doivent alors se replier dans leur conformation biologiquement active. Cette conformation est appelée état natif de la protéine.

Le repliement des protéines est un processus spontané. Cela implique que les structures tridimensionnelles que vont acquérir les protéines sont encodées dans leurs séquences. Le but des recherches effectuées dans ce domaine est de décrire comment ces macromolécules se replient à partir du code dans leurs séquences et quelles structures finales elles vont adopter.

### **I.1.1. Genèse du repliement des protéines**

#### **Origine**

Avant les années 1960, quelques protéines comme la RNase A et la chymotripsogène avaient été dénaturées puis complètement renaturées<sup>20</sup>. Mais, pour ces protéines, les ponts disulfures avaient été laissés intacts. Or, la RNase A, par exemple, possède 8 cystéines toutes impliquées dans la formation de ponts disulfures. La probabilité pour que ces ponts soient formés correctement est quasiment nulle. De plus, la majorité des protéines qui possèdent des ponts disulfures est dépliée lorsque ceux-ci sont réduits. Ils jouent donc un grand rôle dans la stabilité de la protéine. Les interrogations se portent alors sur les ponts disulfures. Comment sont-ils établis ? La protéine est-elle capable de les former correctement et spontanément ?

Anfinsen répond à ces questions en 1961. Il y parvient en étudiant la renaturation de la RNase A complètement dénaturée, avec ses ponts disulfures réduits<sup>21</sup>. La protéine parvient à se replier correctement et à former les bons ponts. Il observe même que, lorsque des ponts disulfures incorrects sont formés au sein des protéines dénaturées, ils sont détruits au cours du repliement pour former les ponts présents dans la structure native. La formation des ponts disulfures est donc dirigée par le repliement des protéines.

Ces expériences conduisent Anfinsen à formuler une hypothèse thermodynamique du repliement des protéines<sup>22</sup> : l'état natif de la protéine est l'état qui possède l'énergie libre la plus basse. Par conséquent, la conformation de l'état natif dépend uniquement de l'ordre des acides aminés dans la chaîne et donc de la structure primaire, de la séquence de la protéine. En mettant en évidence le caractère spontané du repliement des protéines (ou au moins des petites protéines), Anfinsen ouvre la voie à de nombreux expérimentateurs.

### Modèle à 2 états

Dans les premières années de l'étude du repliement-dépliement des protéines, le modèle utilisé est un modèle à deux états : un état déplié et un état replié. Dès 1931, Wu fait l'analogie entre les phénomènes de dénaturation des protéines et de déstructuration des macromolécules cristallines<sup>23</sup>. Il introduit ainsi l'idée que la dénaturation des protéines est uniquement due à un changement de conformation et que les molécules ne subissent pas d'altération chimique pendant ce processus. Kauzmann valide ce concept<sup>24</sup> à partir d'expériences de dénaturation irréversible au début des années 1950. Brandts s'attelle au milieu des années 1960 à étudier les caractéristiques thermodynamiques du dépliement du chymotrypsogène<sup>25</sup>. Il obtient de nombreuses données précises sur l'équilibre entre la forme native et la forme dénaturée en fonction du pH, de la température et de la force ionique. Ces

données lui permettent de proposer un modèle à deux états, sans intermédiaire. Tanford développe l'utilisation d'un puissant dénaturant, le chlorure de guanidium (guanidine), pour l'étude de la dénaturation et du repliement des protéines<sup>26,27</sup>. Il utilise lui aussi le modèle à deux états pour interpréter ses résultats. Dans la fin des années 60 et le début des années 70, Privalov entreprend la construction d'un microcalorimètre très sensible pour déterminer des paramètres thermodynamiques relatifs au repliement des protéines et pour tester le modèle à deux états<sup>28</sup>.

### Repliement en domaine

En 1969, Goldberg étudie la  $\beta$ -galactosidase et montre que cette protéine contient différents domaines qui se replient indépendamment<sup>29</sup>. Dans les années qui suivent, Taniuchi découvre que la RNase A est incapable de se replier quand les quatre acides aminés C-terminal sont délétés, révélant ainsi l'importance d'avoir un domaine complet pour que la protéine se replie<sup>30</sup>. En 1973, Wetlaufer corrobore ces observations en mettant en évidence différents domaines reliés par des boucles dans des structures cristallographiques de protéines<sup>31</sup>.

### Expérience de pensée de Levinthal

A partir de considérations statistiques simples, Levinthal montre en 1968 qu'une protéine ne se replie pas en explorant au hasard les configurations qui lui sont accessibles<sup>32</sup>. Le raisonnement de Levinthal est le suivant : considérons une protéine de 100 acides aminés, dans laquelle chaque acide aminé peut avoir deux configurations différentes. Une telle protéine possède  $2^{100}$  configurations possibles. Les réactions chimiques les plus rapides et les vibrations des atomes dans les molécules se déroulent sur des temps de l'ordre de  $10^{-13}$

secondes. En supposant qu'il faut  $10^{-13}$  secondes à la protéine pour passer d'une configuration à une autre, il lui faudrait  $10^{17}$  secondes, soit  $10^{10}$  années pour explorer toutes les configurations, ce qui correspond à une durée supérieure à l'âge de l'univers. Levinthal en déduit l'existence d'un chemin de repliement pour les protéines et éventuellement d'états intermédiaires. Les considérations de Levinthal encouragent d'autres scientifiques à rechercher des intermédiaires de repliement.

### Intermédiaires de repliement

En 1971, deux équipes révèlent l'existence d'intermédiaires : Ikai et Tanford enregistrent des cinétiques biphasiques grâce à des expériences de dénaturation-renaturation du cytochrome c par la guanidine à l'aide d'un appareil de mélange rapide<sup>33</sup>. Tsong et Baldwin obtiennent un résultat similaire en étudiant la dénaturation de la RNase A par sauts de température<sup>34</sup>. Des cinétiques biphasiques ne peuvent s'expliquer qu'en introduisant un troisième état dans le modèle du repliement des protéines. Ces expériences ont permis de rendre compte de la complexité de cette réaction, mais n'ont pu déterminer si les états détectés étaient sur le chemin de repliement ou en dehors de celui-ci.

### Etude structurale des intermédiaires de repliement

Les études cinétiques permirent de détecter la présence d'états intermédiaires au cours du repliement. En revanche, le caractère transitoire de ces intermédiaires empêcha leur isolement et leur étude structurale. Plusieurs méthodes ont été utilisées dans les années 70 pour étudier ces états.



En 1974, Creighton isole des intermédiaires de repliement de l'inhibiteur de trypsine pancréatique<sup>35</sup>. Pour ce faire, il stoppe la formation de ponts disulfures au cours du repliement, afin de piéger les protéines dans un état intermédiaire.

En 1979, Cook isole des intermédiaires de repliement de la RNase A en suivant le repliement de la chaîne à 0°C, et en provoquant l'isomérisation des prolines présentes dans la protéine<sup>36</sup>.

En 1979, des méthodes d'échanges hydrogène-tritium/deutérium sont développées pour caractériser les intermédiaires de repliement<sup>37</sup>. Les hydrogènes d'une protéine exposés au solvant peuvent s'échanger avec des hydrogènes de ce dernier. Si l'échange de proton est plus rapide que la réaction de repliement, les zones exposées dans les intermédiaires sont accessibles par ces méthodes.

### **I.1.2. Développements expérimentaux**

Pendant les trente dernières années, de nombreuses méthodes expérimentales ont été développées pour déclencher et suivre le repliement des protéines. Grâce à ces méthodes, les réactions peuvent être observées avec plus d'exactitude et de nouvelles informations sont accessibles.

#### Structure secondaire

Pour étudier les éléments de structure secondaire et la formation de ceux-ci, deux méthodes sont utilisées.

### *Dichroïsme circulaire*

Les expériences de dichroïsme circulaire dans l'U.V. lointain (190 – 250 nm) donnent des informations basse résolution sur la structure secondaire des protéines.

Les protéines sont des molécules optiquement actives, elles présentent un dichroïsme circulaire, c'est à dire qu'elles absorbent différemment la lumière selon que celle-ci a une polarisation gauche ou droite. Le dichroïsme circulaire dépend de la structure adoptée par la protéine. Dans l'U.V. lointain (190 – 250 nm), il provient de la structure secondaire de la protéine. La mesure de ce signal dans l'U.V. lointain est utilisée pour estimer la proportion de la chaîne polypeptidique structurée en hélice  $\alpha$ , en feuillets  $\beta$  et en pelote statistique. Pour cela, on compare le spectre de la protéine à ceux caractéristiques de ces structures secondaires. Les informations obtenues par cette méthode sont de basse résolution. Mais elles peuvent être facilement utilisées en cinétique : pour l'étude du repliement des protéines en suivant l'apparition de la structure secondaire, et pour l'étude des intermédiaires de repliement<sup>38</sup>.

### *Spectroscopie infrarouge à transformée de Fourier (FTIR)*

L'absorption infrarouge donne des informations beaucoup plus précises sur la structure secondaire des protéines.

L'absorption infrarouge permet de mesurer les énergies de vibrations des liaisons covalentes, notamment celles des doubles liaisons C=O de la chaîne polypeptidique. Les structures secondaires sont formées par des liaisons hydrogènes entre les hydrogènes portés par l'azote et les oxygènes de la chaîne polypeptidique. Ces liaisons hydrogènes modifient les énergies des doubles liaisons C=O, ces modifications dépendent du type de structure secondaire

formée. Ainsi, en étudiant l'absorption infrarouge de cette liaison, on a accès aux différents éléments de la structure secondaire.

Par rapport au dichroïsme circulaire, cette méthode présente l'avantage de détecter et de discerner les différents types de structure secondaire avec une grande précision. Leur formation peut être suivie pendant la cinétique<sup>39</sup>.

### Structure tertiaire

Différentes méthodes ont été développées pour étudier la formation de structure tertiaire. Certaines peuvent être appliquées à toutes les protéines et ne dépendent pas directement de leurs séquences. D'autres sont plus spécifiques et sont basées sur l'exploitation des propriétés physico-chimiques des acides aminés qui composent la chaîne polypeptidique.

#### *Méthodes ne dépendant pas de la structure*

La diffusion de rayons X et de neutrons permet d'avoir des informations globales sur la taille et la forme des protéines. Au cours de ces expériences, des rayons X ou des neutrons sont envoyés sur une solution de protéine. Les protéines diffusent ces particules ; la taille et la forme des molécules peuvent être déduites du profil de diffusion obtenu. Ces méthodes peuvent être utilisées pendant le repliement de la protéine pour observer l'évolution des paramètres mesurés<sup>40,41</sup>.

Les rayons X peuvent aussi être utilisés pour effectuer des expériences de diffraction. Cette technique permet de déterminer la structure tridimensionnelle des protéines<sup>9,10</sup>. En revanche, elle ne permet pas de détecter la formation de structure au cours du repliement car les protéines doivent être sous forme de cristaux pour ces expériences.

La résonance magnétique nucléaire (RMN) est un outil très puissant pour l'étude des protéines. Une multitude d'informations peuvent être déduites des nombreuses méthodes développées en RMN. On peut notamment surveiller les spins des hydrogènes de la protéine. Ces spins sont influencés par l'environnement des atomes auxquels ils sont associés. Ils permettent d'avoir une idée de l'entourage des atomes sondés et donc de la structure tertiaire<sup>16,42,43</sup>.

Les expériences d'échanges hydrogène/deutérium ou tritium sont basées sur les propriétés qu'ont les acides aminés d'échanger leurs atomes d'hydrogènes avec ceux du solvant. En utilisant une protéine hydrogénée et en la plaçant dans un milieu deutéré (ou inversement), les zones de la protéine où les protons s'échangent le plus rapidement peuvent être observées. Cette technique permet de déterminer quelles zones sont exposées au solvant et quelles sont celles qui sont enfouies. Elle est généralement couplée avec la spectrométrie de masse ou la RMN pour déterminer quels protons ont été échangés<sup>44,45</sup>.

Des méthodes ont été développées dans les années 1990 pour étudier les énergies mises en jeu dans le dépliement de protéine unique. Pour ce type d'expérience, la chaîne polypeptidique est accrochée à une surface par une de ses extrémités. A son autre extrémité est greffée une bille de silice. Celle-ci peut-être manipulée à l'aide d'une pince optique ou d'un microscope à force atomique. En tirant sur la bille, la protéine se déplie et les énergies impliquées dans ce phénomène sont mesurées<sup>46</sup>.

#### *Méthodes basées sur les propriétés physico-chimiques des acides aminés*

De nombreuses techniques sont basées sur l'utilisation des propriétés particulières que possèdent les acides aminés aromatiques.

La mesure de l'absorbance d'une solution de protéine renseigne sur l'environnement des chaînes latérales des résidus aromatiques, et les changements qui peuvent s'y produire au cours de la cinétique<sup>47</sup>. Cette méthode peut être couplée à l'utilisation de cofacteurs susceptibles de se lier à la protéine et d'en modifier le spectre d'absorption.

Dans l'U.V. proche (250 – 320 nm), un dichroïsme circulaire peut être mesuré sur une solution de protéines lorsque celles-ci possèdent une structure tertiaire. Ce dichroïsme dépend de l'absorption, de l'orientation dipolaire et de l'environnement des phénylalanines, tyrosines, tryptophanes et ponts disulfures de la protéine. Contrairement aux spectres enregistrés dans l'U.V. lointain qui permettent d'identifier des éléments de structures secondaires, les spectres dans l'U.V. proche ne peuvent être assignés à un type particulier de structure tertiaire. La mesure du signal permet tout de même de détecter la formation de structure tertiaire au cours du repliement<sup>48</sup>.

De nombreuses méthodes utilisant la spectroscopie de fluorescence sont utilisées pour suivre le repliement des protéines<sup>49,50</sup>.

La fluorescence intrinsèque des tryptophanes donne accès à l'environnement et l'orientation des chaînes latérales des tryptophanes. En regardant l'évolution de cette fluorescence au cours du repliement, des cinétiques peuvent être enregistrées<sup>51</sup>.

La fixation de l'ANS (1-anilino-8-naphthalene sulfonate) peut être mesurée par fluorescence. L'ANS est une petite molécule qui a la particularité de se fixer aux régions hydrophobes. Dans ce cas, son spectre de fluorescence change. Ainsi, l'utilisation de l'ANS permet de suivre l'enfouissement des résidus hydrophobes de la protéine au cours du repliement<sup>52</sup>.

L'utilisation de substrat de la protéine peut être couplée à la spectroscopie de fluorescence pour surveiller, au cours de la cinétique, la formation de sites susceptibles de lier ce substrat.

La technique FRET (*Förster Resonance Energy Transfer*) est aussi utilisée pour suivre le repliement de la protéine. Comme son nom l'indique, cette technique correspond à un

transfert de fluorescence. Au cours d'une expérience de fluorescence, un électron d'une protéine est excité. Lorsque l'électron retourne à son état d'énergie initiale, il émet un photon. C'est ce photon que l'on détecte au cours d'une expérience de fluorescence « traditionnelle ». Or, ce photon peut, sous certaines conditions, exciter lui-même un autre électron de la protéine si celui-ci est suffisamment proche. Cette fluorescence peut être observée en greffant des chromophores adéquats sur la protéine. La mesure de ce phénomène donne la distance entre les deux points de la protéine qui participent à ce transfert, et l'évolution de cette distance lorsque l'on effectue ce type de mesure au cours du repliement de la protéine<sup>53</sup>.

Les expériences d'anisotropie de fluorescence sont des mesures de fluorescence réalisées avec de la lumière polarisée. Cette technique permet de déterminer l'orientation relative de deux sites à l'origine d'un phénomène de fluorescence dans une protéine. L'orientation des deux sites pouvant varier lorsque la protéine change de conformation, cette méthode a été appliquée au repliement des protéines<sup>54</sup>.

Enfin, l'influence de chacun des résidus peut être estimée en utilisant la mutagenèse dirigée. En comparant le repliement de la protéine mutée avec le repliement de la protéine sauvage, des informations sur le rôle de l'acide aminé muté peuvent être déduites<sup>55</sup>.

### Techniques expérimentales utilisées pour déclencher le repliement des protéines

Le repliement des protéines est une réaction rapide. Il est donc très difficile d'avoir des informations sur les événements qui se déroulent dans les premiers instants. Le temps entre le début du repliement et l'acquisition de la première mesure est appelé temps mort de la réaction. De nombreux progrès technologiques ont été réalisés ces dernières années pour diminuer ce temps mort<sup>56</sup>.

Certaines conditions expérimentales sont favorables à un état déplié de la protéine : les hautes températures, l'utilisation d'agents dénaturants, le pH acide dans certains cas, permettent de déplier les protéines. En modifiant ces conditions expérimentales pour les rendre plus favorables à l'état natif, la protéine va se replier spontanément. Ces conditions expérimentales peuvent être modifiées par saut de pH ou de température, ou par dilution de la solution lorsque des dénaturants sont utilisés.

Pour effectuer ces modifications dans un temps très court et diminuer ainsi le temps mort de la réaction, des appareils de mélange ultra-rapide ont été développés. Les temps mort associés à ces mélanges peuvent descendre en dessous de 50 $\mu$ s.

Un laser peut aussi être utilisé pour réchauffer les molécules d'eau et provoquer ainsi le dépliement ou le repliement des protéines. Cette technique diminue le temps mort de la réaction qui peut descendre en dessous de la nanoseconde.

Lorsqu'elles sont soumises à de hautes pressions, les protéines sont partiellement dénaturées (Pour dénaturer complètement une protéine, il faudrait des pressions supérieures à 1000 bars).

Le repliement partiel des protéines peut donc être étudié en effectuant des sauts de pression.

Le temps mort de cette technique est d'environ 50 $\mu$ s.

### **I.1.3. Simulation**

Le développement de l'informatique a permis l'utilisation de cet outil pour étudier le repliement des protéines. En théorie, un ordinateur peut calculer toutes les formes accessibles pour une protéine et sélectionner la forme de minimum d'énergie. En pratique, cependant, cette approche prendrait un temps sidéral du fait de la multitude de conformations que peut adopter la chaîne polypeptidique. Pour pallier ce problème, différents modèles ont été

développés. Dans la suite de ce paragraphe, deux principaux types de modèles utilisés pour simuler le repliement des protéines seront présentés succinctement.

#### « All-atom models »

La forme repliée de la protéine est la forme de minimum d'énergie. L'énergie d'une protéine en solution provient d'une part, des interactions entre ses différents atomes et d'autre part, de ses interactions avec le solvant.

Les modèles de simulation les plus complexes utilisés pour étudier le repliement sont des modèles où tous les atomes de la protéine et du solvant sont représentés<sup>57</sup>. Il existe aussi des modèles où tous les atomes de la protéine sont représentés mais où le solvant est traité comme un continuum<sup>58</sup>.

Ces modèles sont très détaillés et très précis, mais ils nécessitent de nombreux calculs. Ils ne sont généralement pas utilisés pour simuler le repliement de la protéine depuis sa forme dépliée. Ils prennent plutôt comme point de départ la forme complètement repliée de la protéine, et étudient sa dénaturation lorsque les conditions de température, de solvant, de pression varient ou lorsqu'une force extérieure est exercée sur la protéine. La précision de ces modèles permet une comparaison directe avec l'expérience. Cette méthode peut aussi être utilisée pour étudier le repliement depuis des états partiellement dépliés obtenus par simulation ou pour des petits peptides.



## « Lattice models »

Les « *Lattice models* » ou « modèles en treillis » ont été développés pour étudier le repliement des protéines dès les années 60<sup>59</sup>. Ils sont inspirés du modèle d'Ising : les acides aminés sont considérés comme des points qui peuvent se placer à différents endroits d'un treillis.

Les modèles en treillis sont très simples et autorisent un échantillonnage efficace de l'espace des conformations. Ils permettent aussi d'utiliser des méthodes Monte-Carlo qui donnent des valeurs moyennées des paramètres calculés. Ce qui présentait un avantage lorsque les expériences effectuées sur des molécules uniques n'étaient pas encore développées et que seules les valeurs moyennées étaient accessibles.

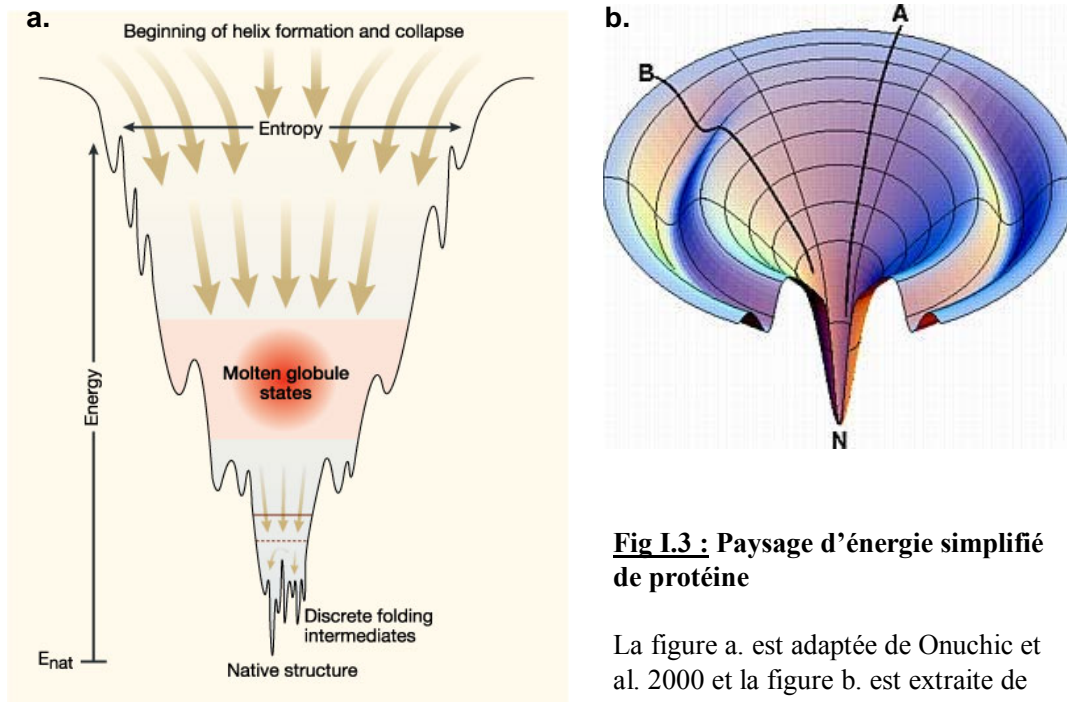
On distingue deux principaux types de modèles en treillis, chacun ayant son propre objectif. Le premier, développé par Go<sup>60</sup> a pour but d'expliquer les principes physiques de base à l'origine du repliement des protéines. Dans ce premier type de modèle, les treillis utilisés sont très simples. Il s'agit en fait de réseaux cubiques où les résidus peuvent se placer sur les sites, et où l'énergie d'interaction entre ces résidus est calculée. Ce modèle a permis à Wolynes de postuler l'existence d'un paysage d'énergie en forme d'entonnoir pour les protéines<sup>61</sup> et d'établir le principe de frustration minimal expliqué dans la suite. Ce modèle a aussi permis à Dill de souligner l'importance des interactions hydrophobes pour le repliement des protéines<sup>62</sup>. Les seconds types de modèle en treillis ont été utilisés par Skolnick<sup>63</sup>, Miyazawa et Jernigan<sup>64</sup>. Ils ont pour objectif une description plus réaliste du phénomène de repliement des protéines. Pour ce faire, les treillis utilisés sont plus compliqués et utilisent des structures de protéines déjà connues. Ces modèles restent très simplifiés, mais les résultats obtenus jusqu'à présent sont très encourageants.

### **I.1.4. Point de vue actuel**

Tous ces progrès technologiques ont permis de proposer des modèles décrivant le repliement des protéines, et l'expérimentation a pu écarter ou confirmer les modèles proposés. La description retenue aujourd'hui pour décrire le repliement des protéines est la description en termes de paysage d'énergie.

#### Paysage d'énergie

Les protéines sont des objets complexes, un très grand nombre d'interactions a lieu à l'intérieur de ces macromolécules, et entre celles-ci et le solvant. Or, elles parviennent toujours à se replier pour former une structure bien définie et très stable. Le repliement de l'état dénaturé à l'état natif se fait en passant par une multitude de conformations. Le mécanisme de repliement devrait donc impliquer un grand nombre de réactions élémentaires. Cependant, les équations cinétiques décrivant la réaction de repliement sont très simples, similaires aux équations utilisées pour décrire des réactions chimiques impliquant des petites molécules. D'où vient cette apparente simplicité? Pour répondre à cette question, on utilise la notion de paysage d'énergie de la protéine (Fig I.3).



**Fig I.3 : Paysage d'énergie simplifié de protéine**

La figure a. est adaptée de Onuchic et al. 2000 et la figure b. est extraite de Dill & Chan 1997.

La description du repliement en terme de paysage d'énergie est développée par Joseph Bryngelson<sup>67</sup> et Peter Wolynes<sup>68</sup> à la fin des années 80 et au début des années 90. Le paysage d'énergie théorique d'une protéine est construit en considérant toutes les configurations géométriques que peut adopter la chaîne polypeptidique et en mesurant leurs énergies internes. Un tel paysage d'énergie ne peut pas être représenté. En effet, pour décrire toute les configurations que peut adopter une protéine de N atomes, il faudrait un espace à 3N dimensions (3 coordonnées par atomes). Les paysages d'énergie sont simplifiés pour permettre de les représenter et d'en comprendre l'intérêt : les configurations sont décrites selon un ou deux axes, et l'énergie de chaque configuration est reportée sur un troisième axe. Un polypeptide construit à partir d'une séquence aléatoire d'acides aminés a peu de chance de se replier, de privilégier une configuration particulière. Il n'y aura pas beaucoup de différence d'énergie entre les conformations qui lui sont accessibles. Le paysage d'énergie sera donc

globalement plat, mais très rugueux, avec une multitude de minima locaux, correspondant à des petits éléments de structure stabilisant transitoirement le polypeptide.

La nature au cours de l'évolution a sélectionné les protéines capables de former rapidement des structures stables. Toutes les interactions au sein de la chaîne polypeptidique vont tendre à stabiliser l'état natif ; les interactions sont constructives. On parle de « consistance » ou du principe de « frustration minimale ». Cela se traduit par un paysage énergétique en forme d'entonnoir (figure I.3). La largeur de l'entonnoir à une énergie donnée est liée au nombre de configurations accessibles pour cette énergie et permet donc d'estimer l'entropie conformationnelle. En haut de l'entonnoir, se situent les conformations dépliées de la protéine. Ces conformations sont de hautes énergies car elles ne contiennent pas de structures susceptibles de stabiliser la protéine. Ces états sont nombreux, ils se situent dans la partie la plus évasée de l'entonnoir. La protéine dépliée possède donc une grande entropie conformationnelle. Au fond de l'entonnoir, on trouve le minimum d'énergie qui correspond à la forme native. C'est cette forme du paysage d'énergie qui confère au repliement des protéines son apparente simplicité. Au cours de son repliement, la chaîne polypeptidique passe d'un état dénaturé à la forme native. En raisonnant en terme de paysage d'énergie, elle va donc tomber du haut de l'entonnoir jusqu'au fond. Elle privilégie le chemin qui lui est le plus accessible le long de l'entonnoir, elle peut éventuellement peupler des états intermédiaires qui correspondent à des minima locaux sur le chemin de repliement.

La forme du paysage énergétique explique aussi la robustesse de la forme native par rapport aux mutations. En effet, une mutation ne modifie pas, en général, la forme globale en entonnoir, mais induit de petites modifications dans le paysage énergétique. Même si le chemin de repliement de la protéine sauvage est modifié et rendu impraticable, étant donné la forme en entonnoir du paysage d'énergie, la protéine trouve toujours un chemin annexe lui permettant de former sa structure native. Cette représentation en entonnoir d'énergie illustre

bien le principe de consistance : une mutation ne modifie que quelques interactions au sein de la protéine, mais les autres interactions tendent à replier la protéine dans sa forme native. En général, la mutation ne modifie donc pas la forme native de la protéine.

### Chemin de repliement

Le chemin de repliement est le chemin qu'emprunte la protéine dans le paysage énergétique lorsqu'elle se replie. La forme d'entonnoir du paysage énergétique n'est pas parfaite. Certains chemins seront privilégiés. En général, la protéine ne passe pas directement du bord au fond de l'entonnoir. Elle va peupler un ou plusieurs intermédiaires de repliement qui seront plus ou moins stables. En se plaçant le long du chemin de repliement, on pourra avoir accès aux différents états peuplés au cours de la cinétique.

Par exemple, sur la figure I.3b, deux chemins A et B sont représentés. Les deux chemins permettent à la protéine de se replier. Si la protéine prend le chemin A, elle se replie directement et rapidement. En revanche, si elle prend le chemin B, un minimum local est rencontré sur son passage et la protéine passe donc par un état intermédiaire.

### Théorie de l'état de transition

Cette théorie est développée au début des années 1930 pour étudier les réactions chimiques simples<sup>69,70</sup>. Dans cette théorie, seuls les états initiaux, finaux et l'état de transition sont considérés. L'état de transition est l'état le plus instable par lequel doivent passer les réactifs pour que la réaction ait lieu. L'état initial et l'état de transition sont considérés en équilibre thermodynamique. Cette hypothèse permet de calculer la concentration de réactif dans l'état

de transition. La constante de vitesse globale de la réaction peut ensuite être calculée en utilisant le taux de décomposition de l'état de transition.

La population de l'état de transition  $X^*$  est déterminée à partir de la population de l'état initial  $X$  grâce à la différence d'énergie entre les deux états  $\Delta G^*$  :

$$[X^*] = [X] \exp\left(-\frac{\Delta G^*}{RT}\right)$$

Pour des réactions chimiques simples, la décomposition de l'état intermédiaire pour former l'état final correspond à la rupture d'une liaison chimique. La fréquence à laquelle se décompose l'état de transition correspond donc à la fréquence à laquelle la liaison chimique est cassée. Elle peut être estimée en utilisant l'équivalence entre l'énergie d'un oscillateur d'un point de vue quantique ( $E=h\nu$ ) et classique ( $E=kT$ ).  $\nu=kT/h$  où  $k$  est la constante de Boltzmann et  $h$  la constante de Planck.

Le taux de décomposition de  $X$  est donc donné par la relation suivante :

$$-\frac{d[X]}{dt} = \nu[X^*]$$

$$-\frac{d[X]}{dt} = [X] \left(\frac{kT}{h}\right) \exp\left(-\frac{\Delta G^*}{RT}\right)$$

La constante de la réaction est donc :

$$k_1 = \left(\frac{kT}{h}\right) \exp\left(-\frac{\Delta G^*}{RT}\right)$$

Dans le cas du repliement des protéines, l'état de transition est beaucoup plus compliqué, et la décomposition ne correspond pas à la simple rupture d'une liaison chimique. La théorie de l'état de transition ne peut donc pas être appliquée telle quelle. Cependant, des équations analogues peuvent être utilisées. Elles permettent de calculer la variation des niveaux d'énergie entre l'état initial et l'état de transition dans différentes conditions expérimentales ( $\Delta\Delta G^*$ )<sup>71</sup>.

La théorie de l'état de transition indique que :

$$k_{repl} = \nu\kappa \exp\left(-\frac{\Delta G^*}{RT}\right)$$

Avec  $\nu$  la vibration caractéristique pour l'état de transition et  $\kappa$  un coefficient de transmission. La valeur de  $\kappa$  est inconnue et difficile à déterminer. L'équation est donc utilisée pour calculer des  $\Delta\Delta G$ , ce qui permet de s'affranchir de ce facteur : lorsque différentes conditions expérimentales sont utilisées, la différence d'énergie entre l'état natif et l'état de transition ( $\Delta G^*$ ) va être modifiée. Cette variation ( $\Delta\Delta G^*$ ) est obtenue grâce à l'équation suivante :

$$\Delta\Delta G^* = RT \ln\left(\frac{k_{repl}^*}{k_{repl}}\right)$$

La variation de la hauteur d'énergie peut être due à différentes perturbations comme des mutations par exemple. L'équation obtenue ouvre la voie aux études par analyse de  $\Phi$  ou par les valeurs de Brønsted qui seront exposées dans la suite.

## **I.2. Cofacteurs**

On distingue deux types de cofacteurs :

- Les cofacteurs constitués par une petite molécule organique appelé aussi coenzymes
- Les cofacteurs constitués par un ion métallique

### **I.2.1. Les coenzymes**

Certains enzymes utilisent des coenzymes pour fonctionner. Ces molécules sont non protéiques et de faibles masses moléculaires comparées à l'enzyme. Elles possèdent une structure simple, thermodynamiquement stable. Elles peuvent être libres ou liées à l'enzyme. Lorsqu'elles sont impliquées dans une réaction enzymatique, elles doivent, comme l'enzyme, être régénérées en fin de réaction, ce qui explique leur faible concentration dans l'organisme.

Le coenzyme s'associe à l'apoenzyme au moment de la catalyse. L'holoenzyme ainsi formé acquiert une spécificité pour son substrat. Cette spécificité peut être étroite (l'holoenzyme n'agit que sur un seul substrat) ou lâche (l'holoenzyme agit sur un groupe de substrats).

Les coenzymes sont des molécules indispensables au métabolisme, mais ils ne sont pas synthétisés par l'organisme. Ils sont introduits dans l'organisme par l'alimentation sous forme de vitamine ou de facteur de croissance. A l'exception des vitamines A et C toutes les vitamines connues sont impliquées dans la structure des coenzymes.



## I.2.2. Les cofacteurs métalliques

### Métaux et biologie

On considère traditionnellement que les êtres vivants sont composés de molécules organiques. Certaines molécules inorganiques sont cependant nécessaires à la vie. C'est le cas de certains métaux. Treize éléments métalliques ont été identifiés comme nécessaires à la vie animale et végétale<sup>72,73</sup>. Quatre d'entre eux sont présents en grande quantité dans les êtres vivants. Il s'agit du sodium, du potassium, du magnésium et du calcium<sup>74</sup>. Ils sont regroupés sous le terme de « bulk metals » et représentent 1 à 2% du poids du corps humain. Les neuf autres métaux, désignés par le terme « trace metals », appartiennent aux métaux de transition : vanadium, chrome, molybdène, manganèse, fer, cobalt nickel, cuivre et zinc. Ils sont présents en petite quantité dans les organismes (0,01% du poids d'un humain).

Un excès ou un manque de métal est dangereux, et peut même être léthal. La concentration en métaux est donc finement régulée. La plupart des métaux traces sont séquestrés dans des protéines. De cette manière, la nature a trouvé un moyen astucieux d'utiliser les propriétés des métaux tout en les encapsulant au sein de protéines, ce qui permet une meilleure régulation<sup>72,73</sup>.

Plus de 30% des protéines utilisent du métal pour fonctionner<sup>75</sup> et 40% des enzymes contiennent des métaux<sup>72,73</sup>. Les géométries des sites de liaison de chacun des métaux sont déterminées par le degré d'oxydation de l'ion métallique, bien que des distorsions puissent avoir lieu dans certaines protéines. Chaque métal possède des propriétés physico-chimiques qui lui sont propres et effectue différents types de fonctions biologiques.

Le fer et le cuivre possèdent des propriétés d'oxydo-réduction et participent au transfert d'électron<sup>76</sup>. Dans la respiration et la photosynthèse par exemple, des petites protéines

possédant des propriétés d'oxydo-réduction participent au transfert d'électrons en se fixant spécifiquement à des protéines membranaires qui possèdent souvent plusieurs sites de liaison. Le fer et le cuivre participent aussi au stockage et au transport du dioxygène par l'intermédiaire de métalloprotéines comme l'hémoglobine, la myoglobine et l'hémocyanine. Le zinc joue lui le rôle de centre super-acide au sein de différentes protéines permettant ainsi l'hydrolyse ou le clivage de liaisons chimiques. C'est le cas dans la carboxypeptidase, l'anhydrase carbonique et l'alcool déshydrogénase. De plus, il joue un rôle structural dans la superoxyde dismutase et les protéines contenant un domaine « Zinc finger ». Les autres « trace metals » ont été identifiés comme des constituants de métalloenzymes<sup>73,75</sup>. Le manganèse, par exemple, est un cofacteur de la superoxyde dismutase mitochondriale et de la phosphatase inorganique ; le nickel est utilisé dans les enzymes comme l'uréase et dans plusieurs hydrogénases. Le molybdène et le vanadium sont tout deux retrouvés associés au fer et au soufre dans les nitrogénases<sup>76</sup>.

### Métaux et repliement des protéines

La liaison du métal a-t-elle un rôle sur le repliement des protéines ? Le métal se lie-t-il uniquement lorsque la structure est formée ? Des études *in vitro* ont montré que beaucoup de métalloprotéines restaient fortement liées à leur cofacteur une fois dénaturées<sup>77,78,79</sup> (cytochrome b, azurine...). Ces résultats suggèrent que les métaux peuvent *in vivo* se fixer aux protéines avant que celles-ci ne se replient. Des structures locales et non locales peuvent être formées grâce à la coordination spécifique du cofacteur.

### **I.3. Fibres amyloïdes**

Le terme amyloïde se réfère, dans son sens classique, à des dépôts extra ou intra cellulaires principalement composés de protéines assemblées en fibres. Ces fibres présentent des caractéristiques particulières : structure en feuillets  $\beta$  croisés, biréfringence verte après coloration au rouge congo, ...

Ce terme est aujourd'hui plus largement utilisé et il inclue toutes les protéines qui adoptent une structure en feuillets  $\beta$  croisés (feuillets  $\beta$  intermoléculaires perpendiculaires à l'axe de la fibre).

#### **I.3.1. Quelques maladies dans lesquelles sont impliquées les fibres amyloïdes**

Les fibres amyloïdes sont impliquées dans une vingtaine de maladies regroupées parfois sous le nom d'amylose ou d'amyloïdose. Parmi celles-ci, on trouve : la maladie d'Alzheimer, la maladie de Creutzfeld-Jakob, l'encéphalopathie spongiforme bovine, la chorée de Huntington, la maladie de Parkinson, la sclérose latérale amyotrophique.

##### **Maladie d'Alzheimer**

La maladie d'Alzheimer est une maladie neurodégénérative qui entraîne la perte des fonctions mentales, suite à une détérioration du tissu cérébral. Cette maladie a été caractérisée par Aloïs Alzheimer, un médecin neuropsychiatre allemand du début du 20<sup>ème</sup> siècle<sup>80</sup>. Jusque dans les années 1960, la maladie d'Alzheimer était considérée comme une maladie rare. On s'aperçut

par la suite qu'elle était très répandue, et que de nombreux signes attribués à la sénescence étaient, en fait, dus à cette maladie.

En France, on compte environ 860 000 personnes atteintes par la maladie d'Alzheimer<sup>81</sup>. Ce chiffre pourrait atteindre 1,3 millions en 2020 et 2,1 millions en 2040. Selon une récente étude, près de 18% des personnes de plus de 75 ans sont atteintes par la maladie.

Les cerveaux des patients atteints présentent, entre autres, des dépôts amyloïdes. Dans le cas de la maladie d'Alzheimer, ces fibres et plaques amyloïdes se forment à partir de peptides  $\beta$ -amyloïde<sup>82</sup>. Ils sont constitués d'une quarantaine d'acides aminés et proviennent d'un mauvais clivage de la protéine APP (Amyloid Protein Precursor). Ces peptides sont inoffensifs et solubles sous forme monomérique. En revanche, lorsque ces monomères s'accumulent, ils changent de structure pour former des plaques amyloïdes. Ces dépôts compriment les neurones, ce qui entraîne un dysfonctionnement, puis la mort neuronale.

Les causes de cette maladie ne sont pas encore totalement connues et aucun traitement n'est actuellement disponible.

### Maladie de Creutzfeld-Jakob, l'encéphalopathie spongiforme bovine

La maladie de Creutzfeld-Jakob fut caractérisée au début du 20<sup>ème</sup> siècle par deux neurologistes allemands : Hans Gerhard Creutzfeld<sup>83</sup> et Alfons Maria Jakob<sup>84</sup>. Il s'agit d'une dégénérescence du système nerveux central caractérisée par l'accumulation d'un prion. La maladie touche une personne sur un million et se déclare généralement entre 55 et 65 ans. Son origine est inconnue dans la plupart des cas (forme sporadique). Cependant, il existe d'autres formes de la maladie : héréditaire, ou ayant pour origine une contamination iatrogénique (liée à un traitement médical).

La protéine prion (PrP) possède au moins deux conformations. La forme native, soluble, présente dans les cellules saines et dont la fonction biologique n'est pas encore connue. L'autre conformation, présente chez les patients atteints par la maladie, est qualifiée par le terme « scrapie » (Prp-sc). Elle est peu soluble et conduit à la formation de fibres amyloïdes. Ces fibres amyloïdes sont très résistantes. Elles entraînent la mort des neurones et la dégénérescence du système nerveux central.

L'encéphalopathie spongiforme bovine est la variante de la maladie de Creutzfeld-Jakob chez les bovins. Cette maladie a été connue du grand public à la suite de la crise de la vache folle. Dans les années 1990, le cours de la viande de boeuf s'est effondré à la suite d'une épidémie d'encéphalopathie spongiforme bovine qui sévissait dans les élevages européens depuis 1986. Depuis les années 2000, le nombre de cas d'ESB ne cesse de baisser.

### La chorée de Huntington

Cette maladie a été décrite dès le 19<sup>ème</sup> siècle et son origine génétique était déjà suspectée à cette époque. La description la plus complète de cette pathologie a été réalisée par George Huntington<sup>85</sup> en 1872. Les premiers symptômes de la maladie se traduisent par des troubles dans la coordination des mouvements, des changements d'humeur. Viennent ensuite des troubles cognitifs. Elle touche entre 5 et 8 personnes sur 100 000 et se déclare généralement entre 40 et 50 ans.

Cette maladie est due à une mutation sur le gène codant l'huntingtine. La fonction exacte de l'huntingtine n'est pas encore connue, mais elle semble jouer un rôle important dans les cellules nerveuses. Cette protéine est nécessaire pour un développement normal prénatal. Elle

est exprimée dans plusieurs tissus du corps avec un niveau d'expression particulièrement élevé dans le cerveau.

Chez les personnes normales, l'huntingtine possède de 6 à 35 acides glutamiques sur sa partie N-terminale. Ce nombre peut varier et la maladie a de forte chance de se développer lorsque cette protéine possède plus de 36 acides glutamiques sur la partie N-terminal<sup>82</sup>. Chez les personnes atteintes, on retrouve des agrégats formés par les parties N-terminales des huntingtines. Ils ont des caractéristiques très proches des fibres amyloïdes, mais leur nature exacte est toujours sujette à controverse.

### Maladie de Parkinson

La maladie de Parkinson est une maladie neurologique qui affecte le système nerveux central. Elle se traduit par des troubles moteurs : tremblements au repos, rigidité des membres, ralentissement des mouvements.

Sur le plan historique, les symptômes de la maladie de Parkinson sont connus et des tentatives de traitements à base d'herbes ont été effectuées dès l'antiquité en Chine, en Inde et en Amazonie<sup>86</sup>. La première description moderne de la maladie a été réalisée par le médecin britannique James Parkinson<sup>87</sup> en 1817.

La maladie de Parkinson touche entre 1 et 2 personnes sur 1000 dans la population générale. Ces chances d'apparition augmentent avec l'âge : la maladie est rare avant 50 ans, 6 à 8 personnes sur 1000 sont touchées entre 65 et 69, 26 à 35 personnes sur 1000 entre 85 et 98. L'âge habituel de début de la maladie se situe autour de 60 ans.

Les causes de la maladie sont multifactorielles, incluant des prédispositions génétiques et l'exposition aux métaux lourds, à divers polluants et pesticides. Les symptômes sont expliqués par un déficit de dopamine qui entraîne un déséquilibre dopamine/acétylcholine.

On observe aussi dans la maladie de Parkinson une accumulation de protéines sur les cellules endommagées<sup>82</sup>. Ces amas sont appelés corps de Lewy et sont constitués de fibres d' $\alpha$ -synucleïne, une protéine dont la fonction principale n'est pas encore connue, mais qui semble jouer un rôle de chaperonne dans la formation de complexes protéiniques. Ces fibres amyloïdes formées par l' $\alpha$ -synucleïne seraient à l'origine de la mort des cellules neuronales.

### La sclérose latérale amyotrophique, maladie de Charcot, maladie de Lou Gehrig

La sclérose latérale amyotrophique encore appelée maladie de Charcot (nom du médecin qui l'a caractérisée pour la première fois) ou maladie de Lou Gehrig (nom d'un célèbre joueur de base-ball atteint) est une maladie neurologique. Elle est due à une dégénérescence des neurones moteurs du cortex cérébral et de la corne antérieure de la moelle épinière, associée à une destruction des voies nerveuses destinées à transporter les messages nerveux entre le cerveau et la moelle épinière<sup>82</sup>.

Cette maladie se manifeste par la coexistence de troubles moteurs centraux et une amyotrophie (fonte des muscles).

Cette pathologie est développée chez 1 à 2 personnes sur 100 000 et les premiers symptômes apparaissent généralement entre 40 et 60 ans.

Cette maladie peut être due à la mutation dans le gène codant pour la superoxyde dismutase. La superoxyde dismutase ainsi mutée est retrouvée dans des agrégats de protéines de type amyloïde.

La description de quelques maladies dans lesquelles est impliquée la formation de fibres amyloïdes permet de relever plusieurs points intéressants :

- Les protéines impliquées dans la formation de fibres amyloïdes sont de nature très différentes.
- Les amyloïdoses ont tendance à se développer majoritairement chez les personnes âgées.

Ces deux points soulignent l'importance de l'étude des fibres amyloïdes : cette conformation peut être adoptée par des protéines très différentes et ce type d'agrégats est retrouvé dans de nombreuses maladies. Avec le vieillissement des populations, ces amyloïdoses sont de plus en plus fréquentes et un nombre croissant d'individus sera confronté, dans un avenir proche, à ce type de pathologie.

### **I.3.2. Protéines susceptibles de former des fibres amyloïdes**

Le repliement d'une protéine est dû aux interactions entre ses différents acides aminés. Des interactions peuvent aussi avoir lieu entre acides aminés appartenant à différentes chaînes polypeptidiques. Dans ce cas, elles peuvent mener à l'agrégation des protéines et parfois à la formation de fibres amyloïdes.

Une vingtaine de protéines sont impliquées dans des maladies à dépôts amyloïdes<sup>88</sup>. Elles ne partagent pas de propriétés structurales, tant au niveau de leur séquence que de leur structure tridimensionnelle. Cependant, malgré leurs différences, elles forment toutes des fibres amyloïdes qui possèdent la même structure tridimensionnelle.

Plus récemment, des fonctions non pathologiques ont été mises à jour pour des fibres amyloïdes dans différents organismes. Les premières fonctions biologiques de fibres amyloïdes ont été trouvées chez la levure.



Une protéine de *Saccharomyces cerevisiae*, Ure2p, réprime l'expression des protéines utilisées dans des milieux pauvres en azote<sup>89</sup>. Cette protéine peut former des fibres amyloïdes, elle est alors notée Ure3 et ne peut plus accomplir sa fonction de répression. La formation de fibres amyloïdes permet donc à la levure de se développer dans les milieux pauvres en azote.

Het-s, une protéine de *Podospora anserina*, contrôle la fusion cellulaire entre cellules de différentes colonies<sup>90</sup>. Il existe deux types de protéine Het : Het-s qui peut être monomérique ou sous forme de fibre amyloïde, et Het-S qui existe uniquement sous forme monomérique. Les cellules possédant des protéines Het (s ou S) sous forme monomérique seront capables de fusionner. En revanche, les cellules possédant Het-s sous forme de fibres amyloïdes ne pourront fusionner qu'avec celles qui possèdent Het-s (monomérique ou sous forme de fibres) et non avec les cellules possédant Het-S. De cette manière, les protéines Het-s sous forme amyloïde permettent aux cellules auxquelles elles appartiennent de reconnaître les cellules qui possèdent la protéine Het-s (monomérique ou amyloïde), et donc les cellules qui possèdent un patrimoine génétique proche. De plus, des protéines Het-s peuvent s'échanger entre les cellules lorsqu'elles ont fusionnées. L'apport de Het-s amyloïde dans une cellule pourvue de Het-s monomérique provoque la formation de fibres à partir de ces monomères.

L'aplysie, appelée plus communément lièvre de mer, est un organisme très utilisé pour l'étude de la mémoire. La protéine CPEB, présente dans cet organisme, montre des similarités avec les protéines prions : partie N-terminale riche en glutamine et grande flexibilité<sup>91</sup>. De plus, elle est retrouvée sous forme monomérique et multimérique et ce sont les protéines sous forme multimérique qui semblent le mieux favoriser la mémoire à long terme<sup>92</sup>. Des efforts sont actuellement faits pour caractériser la structure amyloïde de cette protéine.

Un rôle fonctionnel de fibres amyloïdes dans le corps humain a été découvert récemment. La sous-unité M $\alpha$  de la protéine Pmel17 est capable de former très rapidement des fibres amyloïdes. Ces fibres permettent d'accélérer la production de mélanine par l'organisme<sup>93</sup>.

Différentes fibres amyloïdes ont été retrouvées *in vivo*, elles peuvent être pathologiques ou avoir des fonctions biologiques. Les protéines dont elles sont composées ne partagent pas de similitudes dans leurs séquences ou leurs structures, mais elles forment toutes des fibres amyloïdes qui possèdent la même structure en feuillets  $\beta$  croisés. La conformation en fibre amyloïde semble donc être une configuration générique, un état plus ou moins accessible par toutes les protéines<sup>94</sup>. Le cas échéant, en utilisant des conditions expérimentales adaptées, des fibres pourraient être assemblées à partir de protéines qui ne forment pas de fibres *in vivo*.

Des travaux menés à la fin des années 1990 ont exposé que le domaine SH3 de la PI3 kinase formait un gel a pH acide. Une étude plus approfondie du gel a montré que celui-ci était constitué de fibres qui présentent toutes les caractéristiques des fibres formées par les protéines impliquées dans les amyloïdoses<sup>95</sup>. Ces résultats ont encouragé la communauté scientifique à rechercher des protéines susceptibles de fabriquer des fibres amyloïdes. Dans les années suivantes, des recherches ont montré qu'un grand nombre de protéines, sans lien avec les maladies amyloïdes, étaient susceptibles de former de telles fibres<sup>94</sup>. Ces découvertes, couplées à d'autres observations, ont permis de conclure que la propension des protéines à former des fibres amyloïdes n'est pas due à une spécificité de leur séquence : de nombreuses protéines peuvent adopter une configuration de fibres amyloïdes si elles sont placées dans des conditions appropriées.

### I.3.3. Structure des fibres

Les fibres amyloïdes formées *in vitro* se présentent sous la forme de tiges ou de rubans non branchés d'un diamètre typique de 10 nanomètres et d'une longueur pouvant atteindre plusieurs micromètres<sup>96</sup>. Elles sont exceptionnellement stables.

La première étude structurale des fibres fut réalisée par Astbury<sup>97</sup>. Il étudia la figure de diffraction obtenue lorsque l'on place du blanc d'œuf poché étiré dans un faisceau de rayons X. Cette figure suggérait que les chaînes de la protéine étaient dans une conformation de feuillets  $\beta$  perpendiculaires à l'axe de la fibre. Des figures de diffraction similaires furent observées par la suite pour des fibres amyloïdes prélevées sur des tissus malades et plus récemment pour des fibres formées par des protéines non pathologiques.

Des expériences de diffractions de rayons X et de RMN du solide, effectuées ces dernières années, ont permis de déterminer plus précisément la structure des fibres amyloïdes. Elles sont composées d'un feuillet  $\beta$  intermoléculaire dont les brins sont perpendiculaires à la direction de la fibre<sup>98,99</sup>. Les liaisons hydrogènes sont donc dans l'axe de la fibre (Fig I.4).



**Fig 1.4 :** Structure des fibres amyloïdes

La structure de fibre est représentée en vue de côté et vue de dessus. Seul le cœur de la fibre est représenté. Les feuillets  $\beta$  provenant de différentes protéines sont représentées par des flèches.

Le cœur de la fibre est en feuillets  $\beta$  croisés, c'est cette structure qui est commune à toutes les fibres.

Adapté de :  
 Blake C., Serpell L. (1996) *Synchrotron X-ray studies suggest that the core of the transthyretin amyloid fibril is a continuous  $\beta$ -sheet helix.* *Structure* 4(8), 989-998

Ces fibres ont la particularité d'augmenter la fluorescence de la Thioflavine T et l'absorbance du rouge congo. Ces deux méthodes, plus faciles à mettre en œuvre que la diffraction de rayons X et la spectroscopie RMN, sont particulièrement utiles pour détecter la présence de fibres en solution. Elles sont également utilisées en histologie pour étudier les tissus atteints par des amyloïdoses.

### I.3.4. Formation des fibres amyloïdes

#### Mécanisme de formation

De nombreux mécanismes ont été proposés pour la formation de fibres amyloïdes. La plupart d'entre eux font intervenir un phénomène de nucléation-croissance (ou nucléation-

polymérisation)<sup>100,101</sup>. On retrouve ce type de réaction dans de nombreux processus comme la cristallisation des protéines, l'assemblage de microtubules, la polymérisation de l'actine, ...

La première étape est la nucléation. Au cours de cette étape, une série de réactions d'assemblage de protéines mène à la formation d'un noyau stable. Les réactions d'assemblage de protéines sont thermodynamiquement défavorables : la diminution d'énergie apportée par l'interaction stabilisatrice entre deux protéines ne compense pas la diminution d'entropie<sup>102</sup>. La diminution d'entropie provient du fait que lorsque les protéines sont liées, elles ne peuvent plus se déplacer librement et indépendamment et donc, il y a moins de désordre dans la solution. L'assemblage grandit et de plus en plus de feuillets  $\beta$  intermoléculaires sont formés. Lorsque les interactions à l'intérieur de la structure sont suffisamment stabilisatrices pour compenser la diminution d'entropie, l'agrégat est thermodynamiquement stable. On parle alors de noyau.

La seconde étape de ce processus est la croissance. Une fois le noyau formé, les protéines peuvent directement se fixer à cette structure sous la forme de feuillets  $\beta$  intermoléculaires. Ces interactions sont fortement stabilisatrices et compensent la diminution de l'entropie due à la liaison de la protéine sur la structure. La réaction d'assemblage de protéine sur le noyau directement sous forme amyloïde est donc thermodynamiquement favorable. Ces réactions permettent à la fibre de grandir et constituent la phase de croissance.

### Cinétique

Trois phases peuvent être distinguées dans les cinétiques de formation de fibres.

La première phase correspond au phénomène de nucléation. Au cours de cette phase, appelée « lag phase », phase de latence, peu de changements sont détectés. Des petits multimères sont

formés, mais pas de fibres structurés en feuillets  $\beta$  intermoléculaires. Hortschansky et ses collaborateurs ont montré que la formation du noyau est stochastique : en étudiant la formation de fibres sur plusieurs échantillons issus de la même préparation, ils ont observé une grande disparité dans les durées de la première phase<sup>103</sup>. La stochasticité de la nucléation s'explique par le caractère thermodynamiquement défavorable des réactions qui conduisent à la formation du noyau. Il implique aussi plusieurs complications pour l'étude de la formation de fibres : le temps de la phase de latence varie, le nombre et la longueur moyenne des fibres formées seront différents d'une solution à l'autre, ...

La seconde phase détectée correspond au phénomène de croissance. Les protéines se fixent sur les noyaux. On pourra donc observer la croissance des fibres, la formation des feuillets  $\beta$  intermoléculaires, ... Cette phase est facilement détectable et la vitesse d'élongation des fibres est aisément déterminable.

La troisième phase est atteinte lorsqu'il n'y a plus de protéines monomériques en solution pour se fixer sur la fibre. Les fibres ne peuvent plus s'allonger et un équilibre est atteint.

On peut s'affranchir de la phase de latence en utilisant la technique du « seeding » (ensemencement). Cette technique consiste à introduire en solution des fibres déjà formées. Les fibres servent alors de noyau et les protéines peuvent s'y fixer directement. La phase de croissance est alors directement observable.

### Conditions favorables à la formation de fibre amyloïdes

Les cinétiques de formation de fibres dépendent bien sûr de la concentration en protéine mais aussi de nombreux autres paramètres<sup>104</sup>. Ceux-ci peuvent être inhérents à la séquence de la protéine comme son hydrophobicité, sa stabilité ou sa charge. Les conditions physico-

chimiques de l'environnement de la protéine jouent aussi un rôle important : le pH, la force ionique, la température, l'agitation, ...

Les fibres amyloïdes sont des multimères de protéine. Pour que ces structures soient formées, il faut donc que les chaînes polypeptidiques se rencontrent. Tous les facteurs susceptibles d'influencer cette rencontre sont des paramètres importants pour la formation de fibres. Parmi ces facteurs, il y a ceux qui ont un effet sur l'interaction entre les macromolécules et ceux qui modulent la diffusion des protéines dans la solution.

La concentration en monomère est un facteur crucial pour la formation de fibres : plus la concentration en protéine est importante, plus la distance moyenne entre ces molécules est faible ; les interactions peuvent donc plus facilement avoir lieu.

De nombreux types d'interactions existent entre les protéines. Deux principales forces sont retenues ici : les interactions électrostatiques et les interactions hydrophobes.

Les protéines en solution possèdent une charge électrique. Cette charge dépend de la séquence de la protéine et du pH de la solution. La charge des protéines induit une répulsion électrostatique entre les monomères. Plusieurs facteurs permettent de modifier cette interaction. La modification du pH permet de jouer sur la charge de la protéine. Lorsque le pH de la solution est proche du point isoélectrique (pI) de la protéine, la charge électrique de cette molécule est globalement neutre et les interactions répulsives sont plus faibles. En revanche, lorsque le pH est éloigné du pI, les charges sont nombreuses et la répulsion électrostatique est grande. Ces charges électriques peuvent être écrantées en introduisant des ions en solutions. Dans une solution de grande force ionique, l'écrantage de ces charges atténue la répulsion électrostatique qui existe entre les protéines.

Les interactions hydrophobes entre protéines sont des interactions attractives. Elles sont dues aux segments hydrophobes des protéines. Des protéines possédant des segments hydrophobes exposés au solvant ont tendance à s'associer pour leur empêcher l'accès de l'eau. C'est le même phénomène qui explique la formation de micelles à partir de lipides.

Différents facteurs permettent de moduler la diffusion des protéines au sein de la solution.

L'augmentation de température favorise l'agitation thermique à l'intérieur de la solution et la diffusion des protéines. Des études ont montré que l'augmentation de la température diminue le temps de latence<sup>105</sup> et augmente la vitesse d'élongation<sup>106</sup>. La température joue aussi un rôle sur la stabilité de la protéine, paramètre dont l'importance sera expliquée dans la suite.

Une alternative à l'agitation thermique apportée par la température est l'agitation mécanique. L'agitation mécanique déplace le solvant et les molécules, ce qui facilite leur rencontre. Elle diminue la durée de la phase de latence, et semble atténuer l'effet de la force ionique et de la concentration en protéine<sup>104</sup>.

En plus des facteurs extrinsèques, les facteurs intrinsèques aux protéines comme leur structure et leur stabilité jouent un grand rôle pour la formation de fibres amyloïdes. Les protéines susceptibles de former des fibres amyloïdes sont retrouvées dans toutes les classes structurales ( $\alpha$ ,  $\beta$ ,  $\alpha/\beta$ ). Puisque les fibres amyloïdes partagent la même structure en feuillets  $\beta$  intermoléculaires, les chaînes polypeptidiques doivent subir d'importants changements de conformation. Ils ne peuvent avoir lieu dans la structure native à cause de la structure tertiaire de la protéine qui la contraint et la maintient dans une forme compacte. La formation d'un état non natif, partiellement replié est donc nécessaire. Cet état permet les interactions électrostatiques et hydrophobes nécessaires à l'oligomérisation et la fibrillation. De nombreuses études expérimentales confirment cette hypothèse et montrent qu'un dépliement



partiel est une étape essentielle à la formation de fibres<sup>107,108,109</sup>. Ces états partiellement dépliés sont difficilement observables car ils sont peuplés de manière transitoire au cours de la réaction. Ils sont accessibles lorsque l'état natif est déstabilisé. C'est le cas pour certains mutants de protéines impliquées dans des maladies amyloïdes<sup>110</sup>. La déstabilisation de l'état natif peut aussi être obtenue de manière artificielle, en mutant certains acides aminés de la protéine, en changeant le pH, la température, la pression ou en utilisant des agents dénaturants.

Une caractéristique intéressante des séquences des protéines impliquées dans des amyloïdoses a été découverte en 2001 par Kallberg et ses collaborateurs<sup>111</sup>. Ces chercheurs ont utilisé 1324 protéines de la *protein data bank*. Ils ont ensuite utilisé des algorithmes de prédiction de structure secondaire sur la séquence de ces protéines et ont comparé les prédictions obtenues aux structures secondaires déterminées expérimentalement. Cette comparaison leur a permis de souligner un point intéressant : les protéines impliquées dans les maladies amyloïdes sont celles qui présentent les plus longs segments où des résidus sont prédits en feuillets  $\beta$  et où ils sont en configuration d'hélice  $\alpha$  dans la protéine.

Ce résultat est facilement compréhensible de manière intuitive. Les portions de protéines mises en évidence ont naturellement tendance à se mettre sous forme de brins et à interagir avec d'autres brins pour former des feuillets. Or ces brins sont isolés par le reste de la protéine et ils ne peuvent pas constituer de feuillets tous seuls. En absence de partenaire, ces portions de protéine se résolvent à adopter une configuration en hélice  $\alpha$ . Supposons maintenant que ces protéines puissent interagir entre elles. Les portions prédites en brins  $\beta$  dans chacune des protéines peuvent alors interagir et ne sont plus isolées. Elles se mettent alors sous forme de brins  $\beta$  pour constituer un feuillet  $\beta$  entre les deux protéines. Un raisonnement analogue effectué avec un plus grand nombre de protéines conduit à la formation d'agrégats qui

possèdent des caractéristiques structurales proches des fibres amyloïdes à savoir des feuillets  $\beta$  intermoléculaires.

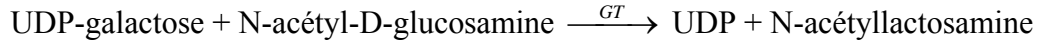
Parmi les protéines qui possèdent de telles portions de séquence et qui sont susceptibles de former des fibres amyloïdes, on trouve l' $\alpha$ -lactalbumine.

## **I.4. $\alpha$ -lactalbumine**

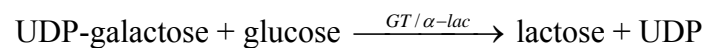
### **I.4.1. Fonctions**

Plusieurs fonctions ont été identifiées pour l' $\alpha$ -lactalbumine. La fonction principale de l' $\alpha$ -lactalbumine, et la plus étudiée, est sa capacité à former un complexe avec la galactosyltransférase<sup>112</sup>. Ce complexe, impliqué dans la synthèse de lactose, est appelé le complexe de lactose synthase

La galactosyltransférase (GT) seule catalyse le transfert de galactose de l'UDP-galactose vers des glycoprotéines et glycolipides, comme la N-acétyl-D-glucosamine :



L' $\alpha$ -lactalbumine se complexe avec la galactosyltransférase et en modifie la spécificité. Le complexe de lactose synthase ainsi formé permet de catalyser le transfert de galactose sur le glucose et de synthétiser le lactose.



L' $\alpha$ -lactalbumine contrôle ainsi la synthèse de lactose. La lactose synthase utilise du  $\text{Mn}^{2+}$  pour fonctionner. Cependant, elle peut aussi être activée par  $\text{Ca}^{2+}$ ,  $\text{Zn}^{2+}$  et  $\text{Co}^{2+}$  (ref 113).

Cette fonction, la plus connue, est celle qui a été le plus étudiée, mais l' $\alpha$ -lactalbumine joue aussi d'autres rôles découverts plus récemment.

Il a été montré qu'à pH acide, la protéine peut utiliser un acide gras (l'acide oléique, noté 18:1) comme cofacteur pour former une forme alternative d' $\alpha$ -lactalbumine appelée HAMLET (*human  $\alpha$ -lactalbumin made lethal to tumor cells*). Cette configuration possède des propriétés apoptotiques envers les cellules tumorales et immatures tout en laissant les cellules normales intactes<sup>114,115,116</sup>. Cette forme de la protéine pourrait avoir un rôle important dans le développement du nourrisson. En effet, l' $\alpha$ -lactalbumine adopte la conformation HAMLET dans des conditions semblables à celles que l'on trouve dans l'estomac des nouveaux nés: de l' $\alpha$ -lactalbumine et des triglycérides sont apportés par le lait maternel, le pH acide de l'estomac va permettre une déstabilisation de la structure de la protéine ce qui va donner lieu au relargage du calcium lié par l' $\alpha$ -lactalbumine. Ce calcium est ensuite utilisé par des lipases pour libérer les acides gras des triglycérides. Les protéines sont donc à pH acides et en présence d'acide gras, les conditions favorables à la formation de HAMLET<sup>117</sup>.

Il a aussi été démontré par Hakansson et ses collaborateurs<sup>115</sup> que la configuration HAMLET de l' $\alpha$ -lactalbumine possède des propriétés bactéricides envers les *Streptococcus pneumoniae*, que ces bactéries soient résistantes aux antibiotiques ou non.

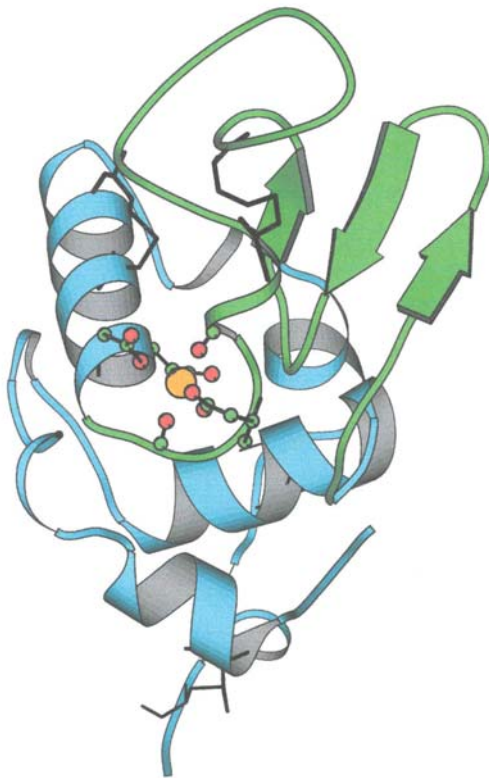
La forme HAMLET n'est pas la seule à posséder des propriétés antimicrobiennes. Trois fragments issus de la digestion protéolytique de l' $\alpha$ -lactalbumine par la trypsine et la chymotrypsine possèdent des propriétés bactéricides<sup>118</sup>.

Une étude expérimentale menée sur des rats a montré que l' $\alpha$ -lactalbumine avait aussi un effet protecteur contre les ulcères induits par l'alcool et par le stress<sup>119</sup>.

## I.4.2. Structure

L' $\alpha$ -lactalbumine possède 123 résidus et son poids moléculaire est d'environ 14,2 kDa<sup>119</sup>.

Elle est composée de deux domaines structuraux (FigI.5) : un domaine en hélice  $\alpha$  et un domaine en feuillet  $\beta$ . Le domaine en hélices  $\alpha$  contient quatre hélices (Hélices A : résidus 5 à 11, B : 23 à 34, C : 86 à 99 et D : 105 à 109) et le domaine  $\beta$  possède un feuillet  $\beta$  antiparallèle de trois brins (résidus 41 à 43, 48 à 50, 55 à 56) et une petite hélice  $3_{10}$  (résidus 77 à 81).



**Fig I.5 : Structure de l' $\alpha$ -lactalbumine.**

Cette figure présente la structure de l' $\alpha$ -lactalbumine.

Le domaine  $\beta$  est en vert et le domaine  $\alpha$  en bleu.

Les ponts disulfures sont représentés par des traits noirs.

Le calcium est indiqué sur la figure par un rond jaune et les chaînes latérales des acides aminés qui le stabilisent au sein de la structure sont aussi représentées.

L' $\alpha$ -lactalbumine est une protéine globulaire. L'hélice B est hydrophobe. Elle est située au cœur de la protéine et est isolée du solvant par le reste de la protéine. La structure est stabilisée par quatre ponts disulfures : deux dans le domaine en hélice  $\alpha$  (Cys28-Cys11 et Cys6-Cys120), un dans le domaine  $\beta$  (Cys61-Cys77) et un entre les deux domaines (Cys73-

Cys91). Dans les expériences de repliement présentées dans la suite, les ponts disulfures sont laissés intacts.

### Site de liaison du calcium

L' $\alpha$ -lactalbumine possède un site de liaison pour l'ion calcium  $\text{Ca}^{2+}$  situé entre les deux domaines. La géométrie de ce site est pentagonale bipyramidale distordue. Il est coordonné par les oxygènes du groupe carboxylique des acides aspartiques 82, 87 et 88, par les groupes carbonyles de la chaîne principale aux résidus 79 et 84 et par 2 atomes d'oxygène de molécule de l'eau<sup>120</sup>. Ce site possède la particularité d'être formé par un long segment de la chaîne (les atomes coordonnant l'ion calcium appartiennent à des résidus proches dans la séquence de la protéine).

Le calcium joue un rôle structural important pour la stabilité et le repliement de l' $\alpha$ -lactalbumine. La liaison du calcium ne modifie pas la structure secondaire de la protéine, mais entraîne des changements dans sa structure tertiaire<sup>120,121,122,123</sup>. Le spectre de fluorescence de la protéine est modifié, ainsi que celui de dichroïsme circulaire. Ces modifications vont dans le sens d'une stabilisation de la structure tertiaire.

Des structures tridimensionnelles de l' $\alpha$ -lactalbumine en présence et en absence de calcium ont été obtenues avec une résolution de 2,2 Å par Chrysina et ses collaborateurs<sup>124</sup>. Les structures globales des deux formes sont très proches. Il y a tout de même quelques changements aux alentours du site de liaison, notamment pour les atomes qui coordonnent le calcium : la chaîne latérale de l'acide aspartique 87 change d'orientation et les molécules d'eau impliquées dans la coordination du calcium disparaissent. D'autres changements induits par la liaison du calcium sont détectés dans la fente sur la face opposée au site de liaison. En présence de calcium, la Tyrosine 103, qui appartient au domaine  $\alpha$ , forme des liaisons

hydrogènes avec les résidus 54 et 56 du domaine  $\beta$ . Dans la forme apo, la chaîne latérale de la tyrosine est déplacée vers l'intérieur de la fente et les liaisons hydrogènes sont remplacées par des liaisons avec les molécules du solvant. De même, une liaison entre la tyrosine 103 et la leucine 104 est remplacée par une liaison avec une molécule de solvant. Dans la forme apo, la tyrosine forme plus de contact avec les résidus du domaine  $\alpha$ , notamment les résidus 99 et 100, et moins avec les résidus du domaine  $\beta$ . Ainsi, la fente dans cette région de la protéine est plus ouverte en absence de calcium.

Les études de RMN effectuées sur les formes apo et holo de la protéine<sup>125</sup> montrent que les pics principaux de résonance sont identiques, mais que de légers changements sont détectés pour les pics de résonance des chaînes latérales des résidus qui coordonnent le calcium. Ces données indiquent que la structure se réarrange localement selon que le calcium est présent ou non.

Les expériences de dénaturation thermique montrent que le calcium stabilise beaucoup la structure de la protéine<sup>121,126,127</sup>. En effet, la structure tertiaire de la protéine disparaît dès 25-30°C pour la forme apo, alors qu'elle est présente jusqu'à 65-70°C pour la forme holo. Cette stabilisation s'explique par la concentration de charges négatives non compensées à proximité du site de liaison du calcium. Lorsque le métal se lie à la protéine, les charges positives portées par l'ion vont compenser les charges négatives de la chaîne polypeptidique.

Les expériences d'échange hydrogène deutérium effectuées sur les deux formes de la protéine montrent que l'échange est plus rapide sur la forme apo, ce qui est cohérent avec les expériences de dénaturations thermiques : la forme holo de la protéine est plus stable.

D'autres métaux sont susceptibles de se lier au site de liaison du calcium, il s'agit des autres métaux de la famille des alcalino-terreux ( $\text{Be}^{2+}$ ,  $\text{Mg}^{2+}$ ,  $\text{Sr}^{2+}$ ,  $\text{Ba}^{2+}$  et  $\text{Ra}^{2+}$ ) et de  $\text{Mn}^{2+}$  (refs 119,

121, 122, 128, 129, 130, 131, 132, 133). Ils ont toutefois une moins bonne affinité que le calcium et stabilisent moins bien la structure de la protéine. Les lanthanides ( $Y^{3+}$ ,  $La^{3+}$ ,  $Lu^{3+}$ ,  $Pb^{2+}$ ,  $Sc^{3+}$ ) sont aussi susceptibles de se lier à ce site<sup>129,134</sup>.

Il a été proposé que l' $\alpha$ -lactalbumine possède un site de liaison pour l'ion zinc  $Zn^{2+}$ . Un site de forte affinité serait situé dans la partie N-terminale de l' $\alpha$ -lactalbumine et coordonnée par les résidus 1, 7, 11 et 37 (ref 135). Il y a aussi quelques preuves de la présence de sites secondaires pour le zinc à proximité des histidines de l' $\alpha$ -lactalbumine. Ces sites sont aussi susceptibles de lier les ions  $Co^{2+}$ ,  $Al^{2+}$  et  $Cu^{2+}$  (refs 136, 137, 138). La liaison de métaux dans les sites  $Zn^{2+}$  ne stabilise pas la structure, ils auraient plutôt tendance à déstabiliser la protéine : la stabilité de la forme holo en présence de zinc est proche de celle de la forme apo et à haute concentration en zinc ; la protéine a tendance à s'agréger.

### **I.4.3. L'état *molten globule***

L' $\alpha$ -lactalbumine peut adopter plusieurs conformations. La structure native décrite ci-dessus est obtenue à pH neutre.

Dolgikh a montré en 1981 qu'à pH2, l' $\alpha$ -lactalbumine est dans un état compact, partiellement dénaturé, qui possède la même structure secondaire que l'état natif mais pas de structure tertiaire stable<sup>139</sup>. Cet état est appelé *molten globule* (globule fondu). Cette propriété que présente l' $\alpha$ -lactalbumine en fait une protéine modèle pour l'étude structurale des états intermédiaires, et le *molten globule* de l' $\alpha$ -lactalbumine est un modèle classique d'étude des états *molten globule*<sup>140,141,142</sup>.



L'étude de l'état *molten globule* par dichroïsme circulaire dans l'UV lointain montre que cet état a une structure secondaire similaire à celle de l'état natif. Les informations apportées par ces expériences sont de basses résolutions. Pour avoir plus d'informations sur la structure, des expériences de RMN et d'échange H/D ont été réalisées : il s'agit de déterminer si les éléments de structure secondaire détectés dans l'état natif et dans l'état *molten globule* sont formés par les mêmes segments de la chaîne polypeptidique. Elles ont montré que le domaine  $\alpha$ , la région du site de liaison du calcium, possède les mêmes éléments de structure secondaire dans l'état *molten globule* et dans l'état natif<sup>143</sup>. Les expériences de RMN ont montré qu'une grande partie des interactions entre les chaînes latérales sont encore présentes dans l'état *molten globule*<sup>144</sup>. Cependant, les pics de résonance sur le spectre RMN sont plus larges dans l'état *molten globule* que dans l'état natif. Cela est dû à l'absence de structure tertiaire stable dans le *molten globule*.

La spectroscopie infrarouge permet de déterminer avec précision la quantité de chacun des types de structures secondaires dans une protéine. Les expériences réalisées mettent en évidence une petite diminution de la quantité de feuillets  $\beta$  au profit de coudes dans l'état *molten globule* par rapport à l'état natif<sup>89</sup>. De plus, les pics du spectre enregistré sur le *molten globule* sont un peu plus larges que ceux du spectre enregistré sur l'état natif. Cela indique que les structures secondaires de l'état *molten globule* sont plus flexibles que celles de l'état natif.

Des études de diffusion quasi élastique de lumière<sup>145</sup> et de diffusion de rayons X ont montré<sup>146</sup> que cet état est compact, et qu'il a une taille proche de l'état natif. La protéine a un rayon de giration égal à 15,7 Å pour l'état natif et 17,2 Å pour l'état *molten globule*, contre 30 Å pour la forme dépliée. Le rayon hydrodynamique de l'état natif est d'environ 19 Å pour l'état natif, 20,5 Å pour le *molten globule* et 24,6 Å pour la forme dépliée. Ce caractère

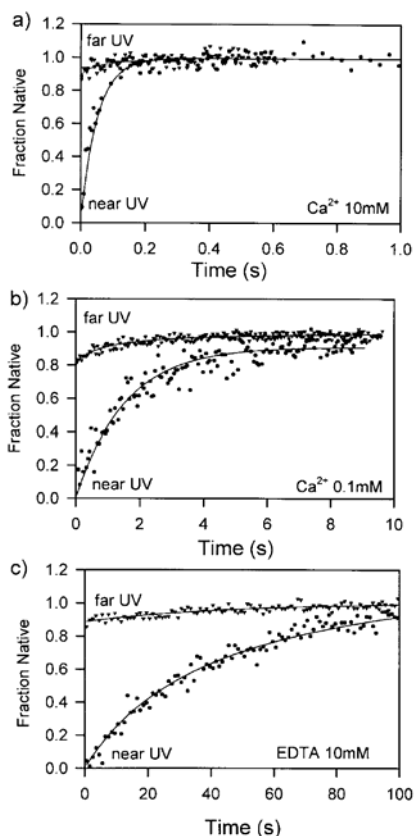
compact du *molten globule* peut être expliqué par la présence d'un noyau hydrophobe, mis en évidence par des expériences de RMN et de mutagenèse dirigée.

#### I.4.4. Repliement

Du fait de sa simplicité et de sa relative abondance, l' $\alpha$ -lactalbumine a été prise comme protéine modèle pour l'étude du repliement des protéines par plusieurs équipes<sup>39,143,147,148</sup>. Plusieurs méthodes ont été utilisées telles que l'absorbance UV, la fluorescence intrinsèque, le dichroïsme circulaire, la spectroscopie infrarouge, la RMN et la diffusion de rayons X.

Le repliement de l' $\alpha$ -lactalbumine se fait en deux étapes. La première étape est très rapide, de l'ordre de quelques millisecondes, et conduit à la formation d'un état compact contenant plus de 85% de la structure secondaire de l'état natif, mais pas de structure tertiaire stable (Figure I.6, ref 143). Cet intermédiaire de repliement est très proche de l'état *molten globule* peuplé à pH acide<sup>143,149,150</sup>. Dans une seconde phase, beaucoup plus lente, la protéine va former sa structure tertiaire pour atteindre son état natif.

Ce mode de repliement est adopté par d'autres protéines comme le lysozyme, la parvalbumine, le cytochrome c, la  $\beta$ -lactoglobuline, la ribonucléase A, l'anhydrase carbonique, la  $\beta$ -lactamase qui peuplent aussi un état *molten globule* dépourvu de structure tertiaire stable avant de trouver leur forme native.



**Figure I.6 : Repliement de l' $\alpha$ -lactalbumine suivi par dichroïsme circulaire en présence de différentes concentrations de calcium**

Le signal enregistré dans l'UV proche indique la proportion de structure tertiaire formée, le signal enregistré dans l'UV lointain donne la proportion de structure secondaire formée.

La concentration en calcium est de 10mM pour la figure a) et 0.1 mM pour la figure b). Pour la figure c), 10mM d'EDTA est utilisé. L'EDTA est un chélateur de calcium, il permet de séquestrer les éventuelles traces de calcium en solution et de suivre la réaction de repliement en absence de calcium.

Figure extraite de Forge et al. 1999

La formation de l'état *molten globule* est très rapide et se fait en général pendant le temps mort de la réaction (inférieur à une milliseconde). Le repliement de l' $\alpha$ -lactalbumine peut donc être étudié en repliant la protéine directement depuis son état *molten globule* sans perdre d'informations. Cette méthode présente en outre de grands avantages expérimentaux : Le *molten globule* est un état qu'adopte spontanément la protéine à pH acide. Le repliement de l' $\alpha$ -lactalbumine peut être déclenché par saut de pH. On s'affranchit ainsi de l'utilisation d'agent dénaturant comme la guanidine ou l'urée. Ces dénaturants peuvent même, à petite concentration, modifier la liaison des métaux à la protéine et perturber les cinétiques.

La constante de vitesse de repliement de la protéine dépend fortement de la concentration en calcium. Les mesures effectuées par dichroïsme circulaire et représentées sur la figure I.6 montrent que lorsque la concentration en calcium augmente, le repliement de l' $\alpha$ -lactalbumine est accéléré (les échelles de temps de la figure I.6 sont différentes pour les parties a, b et c). Le phénomène a été étudié plus en détail par spectroscopie infrarouge<sup>39</sup>. Le

repliement en présence de calcium est coopératif : les différents éléments de structure secondaire sont tous formés avec la même cinétique. Le repliement de la forme apo est beaucoup plus complexe, avec la formation de structures non-natives.

# **II. Repliement de l' $\alpha$ - lactalbumine en présence de métaux**

L' $\alpha$ -lactalbumine possède un site de liaison pour le calcium. Lorsque l'ion se fixe à la protéine, de nouvelles interactions se forment. Elles augmentent la stabilité de la structure de la protéine : la forme holo (en présence de métal) est moins facilement dénaturée que la forme apo sous l'action de la chaleur et d'agents dénaturants. La présence de calcium modifie aussi la constante de vitesse de repliement de l' $\alpha$ -lactalbumine.

D'autres métaux sont susceptibles de se fixer sur le même site de liaison et d'influencer la réaction de repliement<sup>130,119</sup>. Il s'agit des éléments de la famille des alcalino-terreux ( $Mg^{2+}$ ,  $Ca^{2+}$ ,  $Sr^{2+}$ ,  $Ba^{2+}$ ,  $Be^{2+}$ ,  $Ra^{2+}$ ) et de  $Mn^{2+}$ . Chacun de ces ions possède deux charges positives. Cependant, seuls les cations divalents cités ci-dessus peuvent se fixer au site du calcium ; les autres peuvent éventuellement se fixer sur d'autres sites beaucoup moins affins et spécifiques, mais n'accélèrent pas le repliement et ont plutôt tendance à induire l'agrégation de la protéine<sup>130</sup>.

L'objet de cette première partie est l'étude de l'influence des métaux sur le repliement et la dénaturation de l' $\alpha$ -lactalbumine. Les informations déduites de ces expériences permettent d'estimer l'effet de la liaison des métaux sur les paysages d'énergie.

Dans un premier temps, les protéines sont repliées en présence de  $Mn^{2+}$  et des cations de la famille des alcalino-terreux (exception faite de  $Be^{2+}$  et  $Ra^{2+}$ , écartés à cause de leur toxicité). La réaction est suivie par spectroscopie de fluorescence.

La dénaturation de la protéine est ensuite suivie par dichroïsme circulaire et calorimétrie différentielle à balayage.

Les résultats de ces expériences permettent d'obtenir des informations sur les états intermédiaires et sur le paysage d'énergie de l' $\alpha$ -lactalbumine.

## **II.1. Calcul des constantes de vitesse de repliement et des constantes de dissociation**

### **II.1.1 Repliement suivi par spectroscopie de fluorescence**

#### Description de l'expérience

##### *Principe*

La fluorescence intrinsèque de la protéine est utilisée pour suivre son repliement.

Au cours d'une expérience de fluorescence, de la lumière est envoyée sur la protéine. Cette lumière apporte de l'énergie sous forme de photons. La quantité d'énergie apportée dépend de la longueur d'onde. Pour une longueur d'onde de 280 nm, les électrons portés par les cycles des acides aminés aromatiques peuvent absorber les photons de la lumière excitatrice. Les électrons sont alors dans un état électroniquement excité. Ils subissent ensuite une relaxation vibrationnelle, c'est-à-dire qu'ils perdent un peu d'énergie par vibration. Puis ils retournent à leurs états initiaux en émettant des photons d'une énergie légèrement inférieure aux photons excitateurs, donc avec une longueur d'onde légèrement supérieure. Ces photons émis peuvent alors être comptés et donnent une mesure de la fluorescence.

L'intensité de fluorescence décroît lorsque la chaîne polypeptidique se replie. C'est la mesure de ce déclin de fluorescence qui permet de suivre le repliement des protéines et d'enregistrer des cinétiques.

### *Conditions expérimentales*

Les expériences sont effectuées sur un spectrofluorimètre couplé à un appareil de mélange rapide (SFM-400, Biologic) dans une cellule thermostatée. Les protéines sont excitées à 280 nm avec une largeur de bande d'émission de 4 nm. Pour enregistrer la lumière émise, on utilise un filtre laissant passer les longueurs d'ondes supérieures à 320 nm.

La réaction est déclenchée par mélange rapide d'1 volume d'une solution de 140  $\mu\text{M}$  (2 mg/mL) de protéines à pH 2 (protéine dans l'état *molten globule*) avec 10 volumes d'une solution tampon de 100 mM de Tris et diverses concentrations d'ions métalliques ou d'EDTA. Le pH de la solution après mélange est de 7,3, pH favorable à la forme native de la protéine. Le mélange effectué par un appareil de mélange rapide permet de réduire le temps mort à environ 3ms.

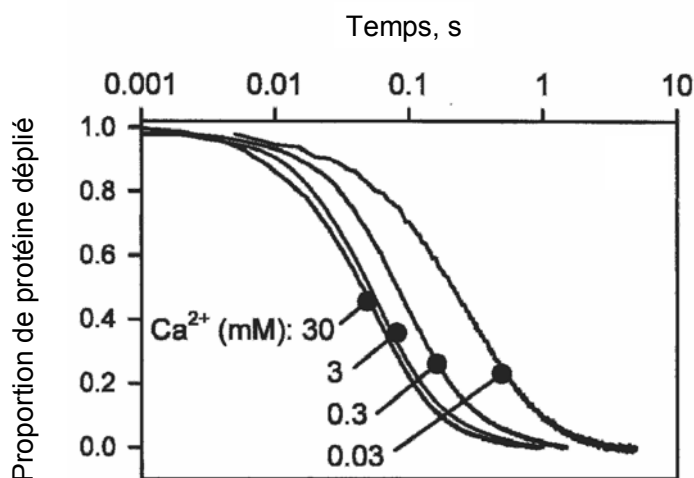
Pour les réactions lentes ( $>50$  s), le mélange est effectué à la main d'où un temps mort d'environ 3 secondes et la mesure est effectuée sur un spectrofluorimètre FP-6500 (Jasco).

#### En présence de différentes concentrations de calcium

La figure II.1 montre les cinétiques de repliement en présence de différentes concentrations de calcium. Le passage de l'état complètement déplié à l'état *molten globule* est trop rapide pour être détecté, le repliement est donc suivi depuis l'état *molten globule* à pH acide. Le repliement est déclenché depuis une solution de protéine à pH 2 en effectuant un saut de pH. Cette méthode présente l'avantage de s'affranchir d'agents dénaturants qui pourraient influencer la liaison des métaux à la protéine. Le repliement observé ici correspond plus exactement à la formation de structure tertiaire.



**Fig II.1 :** Repliement de l' $\alpha$ -lactalbumine en présence de calcium.



Pour obtenir cette figure, la fluorescence de l' $\alpha$ -lactalbumine est enregistrée au cours de son repliement : le rendement quantique des acides aminés diminue lorsque les chaînes latérales sont enfouies au cœur de la protéine. Un déclin de fluorescence peut donc être enregistré au cours du repliement. Ce signal est ensuite normalisé pour faciliter la lecture des courbes et s'affranchir d'éventuelles erreurs de concentrations. Les valeurs que prend la courbe correspondent alors à la proportion de protéine dans l'état déplié (ou plus exactement dans l'état *molten globule*).

Les courbes obtenues sont bi-exponentielles, deux phases sont distinguables. La phase rapide correspond au repliement de la protéine. La phase lente correspond à une réorganisation avec isomérisation de proline suite à la fixation du métal<sup>147</sup>. Par la suite, seule la phase rapide qui représente environ 90% du changement de fluorescence sera considérée.

Les expériences présentées dans la figure II.1 ont été réalisées en présence de 0,03, 0,3, 3 et 30 mM de calcium. La concentration en calcium influence beaucoup la cinétique. Le repliement à 0,03 mM est lent ; la protéine se replie complètement en plusieurs secondes. Lorsque la concentration en calcium est décuplée, la réaction est accélérée et la protéine se

replie en une seconde. La réaction est encore accélérée lorsque la concentration atteint 3 mM ; en revanche, la cinétique enregistrée en présence de 30 mM de calcium est très peu accélérée par rapport à celle enregistrée avec 3 mM. Ce phénomène suggère qu'une limite à l'accélération de la réaction est atteinte lorsque la concentration en calcium dépasse les quelques millimolaires.

### En présence d'autres métaux

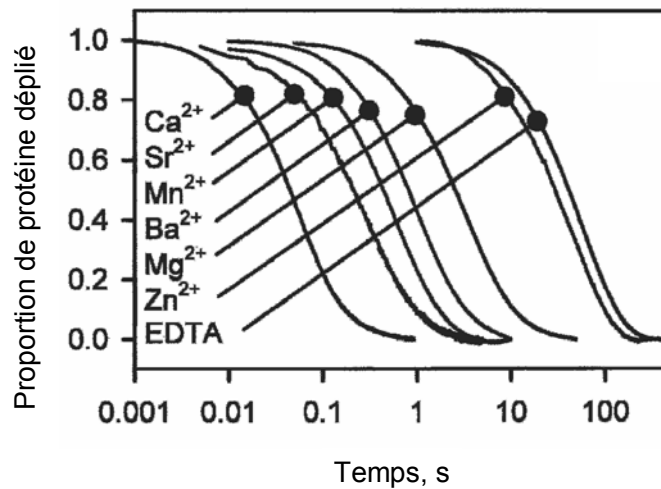
La figure II.2 représente les cinétiques de repliement enregistrées en présence de différents métaux et d'EDTA. Ce dernier est un chélateur qui se lie de façon très affine avec les ions métalliques divalents et trivalents. Il est utilisé pour suivre le repliement de la protéine sous sa forme apo. La concentration en EDTA est de 1mM lorsque le repliement de la forme apo est étudié.

La concentration en métal est de 30 mM pour les métaux susceptibles de se fixer au site de liaison du calcium ( $Mg^{2+}$ ,  $Ca^{2+}$ ,  $Sr^{2+}$ ,  $Ba^{2+}$  et  $Mn^{2+}$ ).

Le zinc est utilisé comme exemple de métal qui ne se fixe pas sur le site de liaison du calcium. La concentration en  $Zn^{2+}$  utilisée est de 10 mM car la protéine s'agrège au delà. Comme pour la réaction en présence de calcium, les cinétiques sont bi-exponentielles en présence des autres métaux susceptibles de se fixer au site de haute affinité. Elles sont mono-exponentielles en présence de Zinc et d'EDTA.

Comme dans la figure II.1, les courbes sont normalisées et montrent la proportion de protéine dans l'état *molten globule*.

**Fig II.2 : Repliement de l' $\alpha$ -lactalbumine en présence de différents métaux**



Plusieurs informations peuvent être directement déduites de ces résultats. La réaction de repliement est fortement accélérée en présence de métaux capables de se lier au site de liaison. Dans ce cas, la réaction est terminée au bout de quelques secondes alors qu'en présence d'EDTA et de zinc, il faut à la protéine une centaine de secondes pour se replier.

Tous les métaux qui se fixent au site de liaison n'accélèrent pas le repliement de la même façon. Le calcium est le métal le plus efficace. En sa présence, la protéine se replie en une demie seconde ; viennent ensuite successivement le strontium, le manganèse et le baryum avec un repliement de l'ordre de quelques secondes et enfin le magnésium en présence duquel la réaction dure quelques dizaines de secondes.

Lorsque le zinc est présent en solution, le repliement semble légèrement accéléré par rapport au repliement de la forme apo. Ces effets peuvent cependant être attribués à la présence des traces de calcium. En effet, l'EDTA ne peut être utilisé dans ce cas car il pourrait chélater le zinc. Le zinc n'accélère pas significativement le repliement de la protéine et, en sa présence, les cinétiques sont mono-exponentielles : ce fait corrobore l'idée qu'il ne se fixe pas au site de liaison du calcium.

## Vitesses de repliement

Comme pour le calcium, les cinétiques en présence des autres métaux ont été enregistrées en présence de différentes concentrations en cation métallique.

Les cinétiques obtenues sont bi-exponentielles, de la forme :

$$I = I_0 \cdot \exp(-k_{repl} \cdot t) + I_l \cdot \exp(-k_l \cdot t)$$

Avec  $I_0$  : intensité de la phase rapide

$I_l$  : intensité de la phase lente

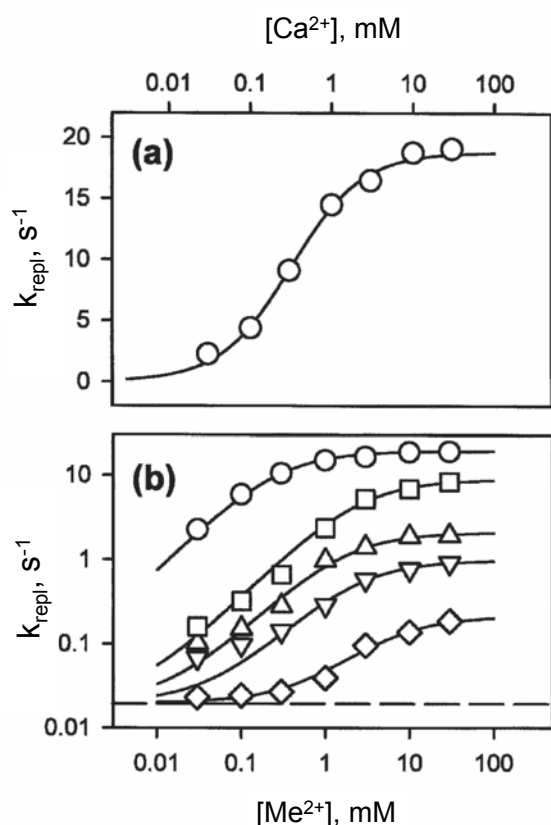
$k_{repl}$  : constante de vitesse de la phase rapide

$k_l$  : constante de vitesse de la phase lente

Le paramétrage de ces courbes permet de déduire les constantes vitesses de repliement.

Comme expliqué précédemment, seul la phase rapide qui correspond à 90 % du déclin de fluorescence sera considérée.

Les constantes de vitesse de repliement  $k_{repl}$ , déterminées en paramétrant les cinétiques, sont reportées dans la figure II.3 en fonction de la concentration en ion métallique dans la solution.



**Fig II.3 : Constantes de vitesse de repliement de l' $\alpha$ -lactalbumine en présence de différents métaux**

a) Constantes de vitesse de repliement de l' $\alpha$ -lactalbumine en présence de différentes concentrations d'ions calcium.

b) Constantes de vitesse de repliement de l' $\alpha$ -lactalbumine en présence de différentes concentrations d'ions métalliques. ( $\circ$ )  $\text{Ca}^{2+}$ , ( $\square$ )  $\text{Sr}^{2+}$ , ( $\triangle$ )  $\text{Mn}^{2+}$ , ( $\nabla$ )  $\text{Ba}^{2+}$ , ( $\diamond$ )  $\text{Mg}^{2+}$ . La ligne en pointillés indique la vitesse de repliement de la forme apo

Les lignes continues correspondent aux paramétrages des courbes permettant de calculer les constantes de dissociation.

La figure II.3a présente les constantes de vitesse de repliement de la protéine en fonction de la concentration en calcium. La gamme de concentration en métal étant très large, l'échelle utilisée pour l'axe des abscisses est logarithmique. Les constantes de vitesse obtenues traduisent bien les premières observations réalisées sur les cinétiques. A faible concentration en calcium (10  $\mu\text{M}$ ), la constante de vitesse est inférieure à une  $\text{s}^{-1}$ . Lorsque cette concentration augmente, la réaction est accélérée. Pour des concentrations supérieures à quelques millimolaires, la constante de vitesse ne varie plus et est égale à une vingtaine de  $\text{s}^{-1}$ . Les constantes de vitesse de repliement mesurées pour les autres cations métalliques sont représentées sur la figure II.3b.

Selon l'ion présent lors de la réaction, les constantes de vitesse de repliement varient beaucoup : l'axe des ordonnées de la figure est donc en échelle logarithmique pour permettre une meilleure lecture des résultats.

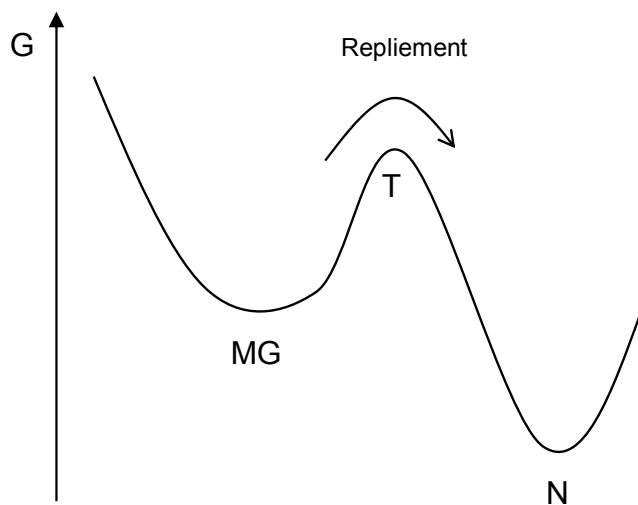
Sur cette figure, les constantes de vitesse de repliement sont représentées par des symboles propres à chaque métal (voir légende de la figure). La constante de vitesse de repliement de la forme apo est représentée par une ligne discontinue.

Bien que les constantes de vitesse soient très différentes pour chacun des métaux, l'allure générale de l'évolution des constantes de vitesse est identique. Elles augmentent progressivement avec la concentration en ion métallique pour atteindre un plateau à haute concentration : les courbes tracées pour chacun des métaux forment toutes des sigmoïdes. A basse concentration, lorsque la concentration en métal tend vers 0, la constante de vitesse de repliement s'approche logiquement de la valeur obtenue pour la forme apo. A hautes concentrations, les constantes de vitesse atteignent un plateau. La constante de vitesse de repliement maximum, valeur atteinte au plateau, est propre à chaque métal.

En ce qui concerne les valeurs absolues des constantes de vitesse de repliement, les différences entre les métaux sont clairement visibles et correspondent globalement à celles observées sur la figure II .2. Le calcium est l'ion en présence duquel la réaction est la plus rapide, suivi par le strontium, le manganèse, le baryum et le magnésium.

L'évolution sigmoïdale des constantes de vitesse de repliement peut être paramétrée, ce qui permet de déduire les constantes de dissociation entre les métaux et les états *molten globule* et de transition de l' $\alpha$ -lactalbumine.

La figure II.4 présente un schéma du chemin de repliement de la protéine.



**Fig II.4 : Chemin de repliement de l' $\alpha$ -lactalbumine.**

Cette figure présente les différents états par lesquels la protéine doit passer pour se replier. Initialement, les protéines sont dans l'état molten globule (MG). Elles doivent passer par l'état de transition (T) pour pouvoir se replier dans sa forme native (N).

Les états sont représentés en fonction de leur énergie libre (G) ce qui permet de visualiser la barrière d'énergie que la protéine doit franchir pour se replier.

Pour se replier, la protéine doit franchir la barrière énergétique entre l'état *molten globule* et l'état natif en passant par un état de transition situé au sommet de cette barrière. La constante de vitesse de repliement dépend de la hauteur de la barrière d'énergie, et donc de la différence d'énergie entre l'état *molten globule* et l'état de transition. La variation de la constante de vitesse de repliement reflète une variation de la hauteur de cette barrière d'énergie. Cette hauteur change lorsque l'énergie de l'état *molten globule* et/ou de l'état de transition est modifiée. Cette modification peut être imputée à la fixation de l'ion métallique sur ces états<sup>147</sup>.

Le paramétrage de l'évolution de la constante de vitesse de repliement avec la concentration en métal, permet de déduire la manière avec laquelle la hauteur de cette barrière varie. Ce paramétrage détermine donc les constantes de dissociation entre les états *molten globule* et de transition et les ions métalliques.

L'équation utilisée pour paramétrer les données est déterminée à l'aide de la théorie de l'état de transition<sup>71</sup>. D'après cette théorie, la constante de vitesse de repliement est égale à :

$$k_{repl} = \nu\kappa \exp\left(-\frac{\Delta G^*}{RT}\right)$$

Avec  $\nu$  la vibration caractéristique pour l'état de transition,  $\kappa$  un coefficient de transmission,  $T$  la température,  $R$  la constante des gaz parfaits, et  $\Delta G^*$  la différence d'énergie entre l'état *molten globule* et l'état de transition (hauteur de la barrière d'énergie). Cette différence d'énergie est comparée à la différence d'énergie dans le cas de la forme apo :

$$\begin{aligned}\Delta G^* &= G_T - G_{MG} \\ \Delta G^* &= (G_T^\circ - G_{MG}^\circ) + (G_T - G_T^\circ) - (G_{MG} - G_{MG}^\circ) \\ \Delta G^* &= \Delta G^{*\circ} + (G_T - G_T^\circ) - (G_{MG} - G_{MG}^\circ)\end{aligned}$$

Dans l'équation précédente, les termes de la forme  $G_X - G_X^\circ$  correspondent à la différence d'énergie entre la forme liée et la forme libre de la protéine dans l'état X, et  $\Delta G^{*\circ}$  correspond à la hauteur de la barrière d'énergie en absence de métal. Ces termes peuvent être exprimés en fonction de la constante de dissociation d'après l'équation suivante<sup>151</sup> :

$$(G_X - G_X^\circ) = -RT \ln(1 + [Me^{2+}]/K_X)$$

D'où

$$\Delta G^* = \Delta G^{*\circ} - RT \ln(1 + [Me^{2+}]/K_T) + RT \ln(1 + [Me^{2+}]/K_{MG})$$



Ainsi, les constantes de vitesse de repliement en présence de métal sont reliées à la constante de vitesse de repliement de la forme apo par la relation suivante :

$$k_{repl} = \nu\kappa \exp\left(-\frac{\Delta G^{*\circ} - RT \ln(1 + [Me^{2+}]/K_T) + RT \ln(1 + [Me^{2+}]/K_{MG})}{RT}\right)$$

$$k_{repl} = \nu\kappa \exp\left(-\frac{\Delta G^{*\circ}}{RT}\right) \cdot \exp(\ln(1 + [Me^{2+}]/K_T)) \cdot \exp(-(\ln(1 + [Me^{2+}]/K_{MG})))$$

$$k_{repl} = k^\circ \cdot \left[ \frac{1 + [Me^{2+}]/K_T}{1 + [Me^{2+}]/K_{MG}} \right]$$

Avec  $k^\circ$  la constante de vitesse de repliement de la forme apo :  $k^\circ = \nu\kappa \exp\left(-\frac{\Delta G^{*\circ}}{RT}\right)$ .

C'est cette dernière équation qui est utilisée pour paramétrer les évolutions des constantes de vitesse de repliement présentées dans la figure II.3. Ces paramétrages sont représentés par des lignes continues. Les constantes de dissociation entre le métal et les états *molten globule* et de transition de la protéine déterminées par cette méthode, sont reportées dans la table II.1. Les valeurs de ces constantes sont étudiées plus en détail par la suite.

**Table II.1 : Constantes de dissociation des ions métalliques avec les différents états de l'  $\alpha$ -lactalbumine impliqués dans la réaction de repliement**

Metal	$K_N$ , M	$K_T$ , M	$K_{MG}$ , M	$k_{max}$ , <sup>a</sup> sec <sup>-1</sup>	$\Delta H^\ddagger$ , <sup>b</sup> kJ/mol
Mg <sup>2+</sup>	$1.1 \times 10^{-3}$	$5.6 \times 10^{-4}$	$5.9 \times 10^{-3}$	0.19	$50.5 \pm 2.0$
Mn <sup>2+</sup>	$5.3 \times 10^{-6}$	$1.3 \times 10^{-5}$	$1.3 \times 10^{-3}$	1.9	$51.5 \pm 1.0$
Ca <sup>2+</sup>	$1.0 \times 10^{-8}$	$2.5 \times 10^{-7}$	$2.4 \times 10^{-4}$	19.0	$57.5 \pm 1.0$
Sr <sup>2+</sup>	$1.4 \times 10^{-6}$	$5.6 \times 10^{-6}$	$2.5 \times 10^{-3}$	8.3	$60.5 \pm 2.0$
Ba <sup>2+</sup>	$6.4 \times 10^{-5}$	$4.6 \times 10^{-5}$	$2.3 \times 10^{-3}$	0.77	$57.5 \pm 1.5$
EDTA	—	—	—	0.020	$58.5 \pm 5.0$

<sup>a</sup> Constante de vitesse de repliement enregistrée en présence d'une concentration saturante d'ions métalliques (et d'1mM d'EDTA pour la forme apo). Les erreurs sur les constantes de vitesses sont d'environ 10%.

<sup>b</sup>  $\Delta H^\ddagger$  Déterminé d'après les pentes du graphe d'Eyring (figure II.10).

Pour avoir une description plus complète de la réaction, et plus particulièrement de la fixation du cation métallique sur l' $\alpha$ -lactalbumine, il faut aussi calculer la constante de dissociation entre le métal et l'état natif de la protéine.

### II.1.2 Courbes de liaison

Les constantes de dissociation entre les métaux et l'état natif sont calculées à partir de courbes de liaisons. Ces courbes représentent la proportion de protéine sous la forme holo (qui ont lié un métal) en fonction de la concentration en ions métalliques libres.

## Description de l'expérience

### *Principe*

La proportion de protéine qui lie un ion métallique est mesurée par spectroscopie de fluorescence. La fixation du métal induit un léger déplacement du spectre de fluorescence de la protéine vers les petites longueurs d'onde. Ce déplacement est quantifié en effectuant le rapport des intensités d'émission à 320 nm et 365 nm. Ce paramètre est appelé A :

$$A = \left( \frac{I_{320}}{I_{365}} \right)_{280} . \text{ Ce n'est pas un paramètre extensif, mais son comportement en est proche}^{152}$$

pour les expériences effectuées ici (le déplacement du  $\lambda_{\text{max}}$  de fluorescence est petit). De plus, il présente l'avantage d'être très peu dépendant des conditions expérimentales : la spectroscopie de fluorescence est très sensible à la concentration en protéine, à l'usure de la lampe, ... Un facteur multiplicatif doit être utilisé pour corriger les erreurs induites. Pour les cinétiques par exemple, les courbes sont normalisées : les effets non désirables sont ainsi évités. Pour les expériences réalisées ici, les éventuelles erreurs ont le même effet sur l'intensité à 320 nm et celle à 365 nm, les facteurs correctifs sont identiques. Ainsi, ces erreurs ne seront pas visibles dans le rapport entre les deux intensités.

L'utilisation de ce paramètre réduit donc la sensibilité aux conditions expérimentales et permet une meilleure précision dans la détermination des constantes de dissociation.

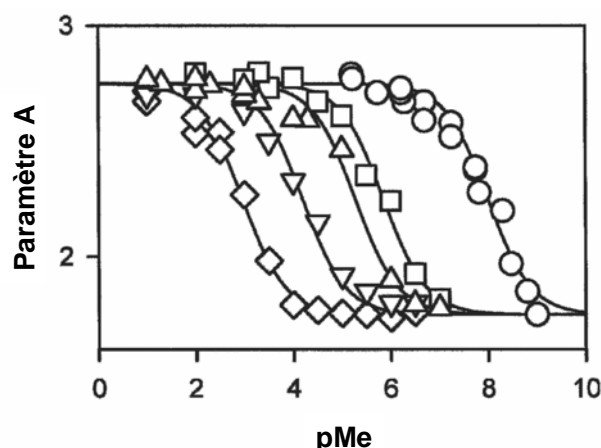
### *Conditions expérimentales*

Les expériences sont réalisées sur un spectrofluorimètre FP-6500 (JASCO) dans une cellule de 1 cm de chemin optique, thermostatée à 20°C. La fluorescence intrinsèque est mesurée à

une longueur d'onde d'excitation égale à 280 nm avec une largeur de bande de 3 nm. Les spectres d'émission sont enregistrés entre 300 et 400 nm avec une largeur de bande de 5 nm et une vitesse de balayage spectrale de 100 nm/min. La concentration en protéine est de 1,4  $\mu$ m (0,002 mg/mL). La concentration en métal est contrôlée en utilisant des tampons EDTA-métaux et EGTA-métaux (L'utilisation de l'EDTA et de l'EGTA permet de contrôler finement la concentration en ion métallique libre en solution en considérant les constantes de dissociation entre ces chélateurs et les métaux).

#### Constantes de dissociation entre les métaux et l'état natif

Les paramètres A calculés à partir des spectres de fluorescence enregistrés en présence de différentes concentrations de métaux sont représentés sur la figure II.5. L'abscisse du graphe de la figure présente le pMe, ce paramètre est égal à  $-\log[\text{Me}^{2+}]$ , avec  $[\text{Me}^{2+}]$  la concentration en métal libre ; il facilite la lecture des résultats.



**Fig II.5 :** Courbe de liaison des différents ions métalliques à l'  $\alpha$ -lactalbumine.

Le paramètre A permettant de mesurer la liaison des ions métalliques sur l'  $\alpha$ -lactalbumine dans sa forme native est représenté en fonction des  $-\log[Me^{2+}]$ . ( $\circ$ )  $Ca^{2+}$ , ( $\square$ )  $Sr^{2+}$ , ( $\triangle$ )  $Mn^{2+}$ , ( $\nabla$ )  $Ba^{2+}$ , ( $\diamond$ )  $Mg^{2+}$ .

Les lignes continues correspondent au paramétrage des courbes permettant de déterminer les constantes de dissociations entre les ions métalliques et l' $\alpha$ -lactalbumine sous sa forme native.

L'allure des courbes est la même quelque soit le métal utilisé : le paramètre A décrit une sigmoïde. Pour les grandes valeurs de pMe, lorsque la concentration en métal est faible, A est égale à 1,75. Cette valeur est celle obtenue lorsque la protéine ne fixe pas de métal. Lorsque la concentration en ions métalliques augmente (pMe diminue), A va prendre des valeurs de plus en plus élevées jusqu'à atteindre un plateau avec un paramètre A égal à 2,75. Cette augmentation traduit un décalage du spectre d'émission de fluorescence induit par la liaison du métal. Le spectre d'émission de la forme apo possède une intensité de fluorescence maximum pour  $\lambda_{max} = 335$  nm, alors que ce maximum est situé pour  $\lambda_{max} = 330$  nm pour la forme holo. Ainsi, le paramètre A dépend de la fixation du métal sur la protéine et permet d'estimer la proportion de protéine sous la forme holo dans la solution.

La fixation de l'ion sur la protéine dépend de la nature du métal utilisé ; ceci explique le décalage observé entre les courbes mesurées en présence de différents métaux. Plus l'affinité de l'ion métallique pour l'état natif est grande, plus la concentration nécessaire pour la liaison

sera faible (donc plus le pMe sera grand). Les ions métalliques qui sont les plus affins pour l'état natif sont dans l'ordre : le calcium, le strontium, le manganèse, le baryum et enfin le magnésium.

Pour le calcium, la liaison a lieu pour des concentrations en métal libre comprises entre 1 et 100 nM alors que pour le magnésium, elle n'est possible qu'entre 100 μM et 10 mM.

Les sigmoïdes obtenues peuvent être paramétrées pour déterminer la constante de dissociation entre l'état natif et le métal. Les solutions sur lesquelles sont réalisées les mesures contiennent des protéines sous la forme apo et d'autres sous la forme holo. Ces deux formes ont un paramètre A distinct. Ce paramètre a un comportement proche d'un paramètre extensif. Le paramètre A, mesuré sur la solution, correspond donc à une combinaison linéaire des paramètres de chacune des espèces : le paramètre A des formes apo et holo est pondéré par la proportion de chacune des deux formes dans la solution.

Le paramètre A de la forme apo de la protéine est celui obtenu pour des faibles concentrations en protéine :  $A_0$ . Le paramètre A de la forme holo est celui obtenu pour des fortes concentrations en métal :  $A_0 + A_a$

Ainsi, le paramètre A de chacune des solutions en fonction de la concentration en métal est égal à :

$$A = \frac{[P]}{[P.Me^{2+}] + [P]} A_0 + \frac{[P.Me^{2+}]}{[P.Me^{2+}] + [P]} (A_a + A_0)$$

$$A = A_0 + A_a \frac{[P.Me^{2+}]}{[P.Me^{2+}] + [P]}$$

Avec	$A_0$	: Valeur initiale de A
	$A_a$	: Amplitude de la variation de A
	$K_N$	: Constante de dissociation entre l'état natif et le cation métallique
	$[Me^{2+}]$	: Concentration en ion métallique

Donc

$$A = A_0 + A_a \frac{1}{1 + [P]/[P.Me^{2+}]}$$

Or

$$K_N = \frac{[P] \cdot [Me^{2+}]}{[P.Me^{2+}]}$$

D'où

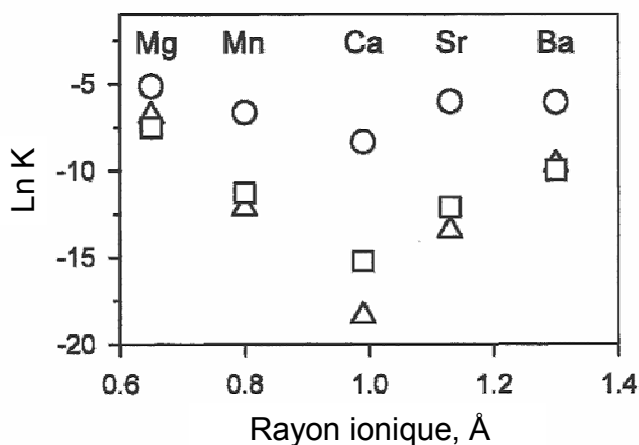
$$A = A_0 + \frac{A_a}{\left(1 + \left(K_N/[Me^{2+}]\right)\right)}$$

Cette équation permet de paramétrer les valeurs de A et d'en déduire les constantes de liaison entre les métaux et l'état natif. Sur la figure II.5, les paramétrages sont représentés par des lignes continues. Les constantes de dissociation calculées par cette méthode sont représentées dans la table II.1.

### II.1.3 Constantes de dissociation

Les constantes de dissociation entre les métaux et les différents états de la protéine (*Molten Globule*, état de transition et état natif) qui ont été calculées grâce aux expériences de

spectroscopie de fluorescence sont représentées sur la figure II.6 en fonction du rayon ionique des cations métalliques. Les rayons ioniques des ions métalliques sont les suivants :  $Mg^{2+}$  :  $0,65\text{\AA}$  ;  $Mn^{2+}$  :  $0,8\text{\AA}$  ;  $Ca^{2+}$  :  $0,99\text{\AA}$  ;  $Sr^{2+}$  :  $1,13\text{\AA}$  ;  $Ba^{2+}$  :  $1,30\text{\AA}$ .



**Fig II.6 :** Constantes de dissociation entre les ions métalliques et les différents états de l' $\alpha$ -lactalbumine représentées en fonction du rayon des ions considérés.

- : Constante de dissociation entre le métal et la protéine dans l'état molten globule.
- : ..... de transition.
- △ : ..... natif.

Quel que soit l'état considéré, toutes les constantes de dissociation présentent un profil en forme de chevron centré sur le calcium. Cela confirme que le calcium est le cation le mieux adapté pour le site de liaison de l' $\alpha$ -lactalbumine, et cela quel que soit l'état dans lequel se trouve la protéine.

En général, la forme native est celle qui fixe le mieux le métal avec des affinités de l'ordre de la dizaine de nanomolaires pour le calcium, du micromolaire pour le manganèse et le strontium, de la dizaine de micromolaires pour le baryum et du millimolaire pour le magnésium. L'état de transition est moins affiné que l'état natif excepté (i) pour le baryum dont les affinités pour l'état natif et l'état de transition sont du même ordre de grandeur, et (ii)



pour le magnésium qui semble se fixer plus facilement à l'état de transition qu'à l'état natif. L'état *molten globule* est celui qui lie le moins bien les métaux, quelle que soit leur nature. Cependant, le calcium est encore le métal qui se lie le mieux même lorsque la protéine est dans l'état *molten globule*.

## **II.2. Comparaison des variations d'énergie des différents états et implications pour le repliement**

Dans la partie suivante, la stabilisation de la protéine par la fixation des métaux est étudiée. La comparaison des variations induites sur les différents états de la protéine permet de déduire des informations sur la structure du site de liaison des métaux et sur la réaction de repliement de l' $\alpha$ -lactalbumine. Pour effectuer ces comparaisons, deux approches complémentaires sont utilisées : l'équation de Brønsted et l'analyse par valeur de  $\Phi$ .

### **II.2.1 Equation de Brønsted**

L'équation de Brønsted a originellement été développée dans les années 1920 pour comparer la force d'un acide à son activité catalytique. Elle a par la suite été utilisée pour d'autres systèmes, et est considérée ici pour étudier la fixation du métal à l' $\alpha$ -lactalbumine.

Cette équation permet de vérifier que le repliement de la protéine est bien une réaction à un seul état de transition.

Il s'agit de mettre en parallèle la stabilisation de l'état natif, à la constante de vitesse de repliement de la protéine. En d'autres termes, on compare la variation d'énergie entre l'état de transition et l'état *molten globule* en présence de différents métaux, à la variation d'énergie de l'état natif dû à la fixation de l'ion métallique. Lorsque ces deux différences d'énergie suivent une relation linéaire, le système obéit à l'équation de Brønsted. La réaction est à simple état de transition.

Sur la figure II.7a, les  $\ln(k_{\max})$  en présence des différents métaux sont représentés en fonction des  $\ln(K_N)$  en présence de ces mêmes métaux.

$k_{\max}$  correspond à la constante de vitesse de repliement maximum. C'est la constante de vitesse calculée pour les hautes concentrations en métaux, lorsque le métal est fixé aux états *molten globule* et de transition. D'après la théorie de l'état de transition,

$$k_{repl} = \nu \kappa \exp\left(-\frac{\Delta G^*}{RT}\right)$$

Avec  $\nu$  la vibration caractéristique pour l'état de transition,  $\kappa$  un coefficient de transmission,  $T$  la température,  $R$  la constante des gaz parfait et,  $\Delta G^*$  la différence d'énergie entre l'état *molten globule* et l'état de transition (hauteur de la barrière d'énergie).

$\ln(k_{\max})$  dépend donc linéairement de la hauteur de la barrière d'énergie, c'est une quantité proportionnelle à la différence d'énergie entre l'état *molten globule* et l'état de transition.

$\ln(K_N)$  représente la stabilisation de l'état natif par le métal. En effet, d'après Schellman<sup>151</sup>,

$$\Delta G = -RT \ln\left(1 + [Me^{2+}]/K_N\right)$$

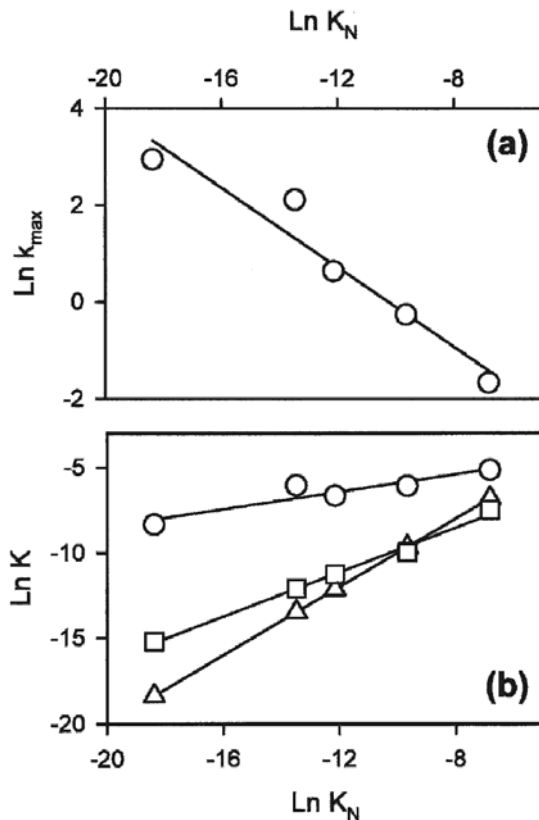
Or, la concentration en métal est égale à 30 mM, donc  $[Me^{2+}] \gg K_N$ ,  $[Me^{2+}]/K_N \gg 1$ ,

$$\ln\left(1 + [Me^{2+}]/K_N\right) \approx \ln\left([Me^{2+}]/K_N\right) = \ln([Me^{2+}]) - \ln(K_N)$$

Ainsi,

$$\Delta G \approx RT \ln(K_N) - RT \ln([Me^{2+}])$$

Pour les différentes expériences, la concentration en métal est égale à 30 mM,  $RT \ln([Me^{2+}])$  ne varie pas. Ce terme n'est donc pas pris en compte pour tracer la dépendance car il ne modifie pas la pente de la régression linéaire qui correspond à la valeur de Brønsted.



**Fig II.7 : Graphe de Brønsted**

- $\ln(k_{\max})$  représenté en fonction de  $\ln(K_N)$  pour les différents métaux. La droite est tracée par régression linéaire sur les données de la figure. Sa pente est égale à  $0,41 \pm 0,06$  et correspond à la valeur de Brønsted.
- $\ln(K)$  représenté en fonction de  $\ln(K_N)$  pour les différents métaux. Les  $K$  correspondent aux constantes de dissociation entre le métal et l'état molten globule ( $\circ$ ), l'état de transition ( $\square$ ) et l'état natif ( $\triangle$ )

Les points obtenus sur la figure II.7a suivent une relation linéaire : ceci indique que le système obéit à l'équation de Brønsted et donc, que le repliement de l'état *molten globule* à l'état natif est bien décrit par une réaction à un seul état de transition. La valeur de Brønsted correspond à la pente de la droite formée par les points, et est égale à  $0,41 \pm 0,06$ .

A titre de comparaison, les constantes de dissociation de l'état *molten globule* et de l'état de transition sont représentées en fonction de la constante de dissociation de l'état natif sur la figure II.7b. Les points obtenus forment des droites de pentes  $0,25 \pm 0,05$  pour le *molten globule* et  $0,65 \pm 0,05$  pour l'état de transition.

Plusieurs conclusions peuvent être tirées de cette analyse :

- La réaction est bien une réaction à un seul état de transition. En effet, Le système obéit à l'équation de Brønsted ; la hauteur de la barrière d'énergie à franchir varie linéairement avec la stabilisation de l'état natif lorsque différents métaux sont utilisés.

- Le site de liaison des métaux est le même dans chaque état peuplé par les protéines. En effet, la liaison d'un métal stabilise la protéine quelque soit son état ; les stabilisations des états *molten globule* et de transition varient linéairement avec la stabilisation de l'état natif. En d'autres termes, lorsqu'un métal stabilise beaucoup l'état natif, il stabilise également beaucoup les états *molten globule* et de transition ; à l'inverse, lorsqu'un métal stabilise peu l'état natif, il stabilise peu les états *molten globule* et de transition.
- Les pentes des droites de régression linéaire illustrent le degré de structuration du site de liaison pour les différents états de la protéine. Plus la pente associée à un état intermédiaire est proche de celle associée à l'état natif, plus la liaison d'un métal sur cet état a le même effet que la liaison d'un métal sur l'état natif, et donc plus le site de liaison de cet état aura la même structure que l'état natif.

## II.2.2 Analyse des valeurs $\Phi$

### Principe

L'analyse des valeurs  $\Phi$  pour l'étude du repliement des protéines a été développée par Fersht et ses collaborateurs à la fin des années 80 (ref 71).

L'état de transition est un état par lequel passe la protéine lorsqu'elle se replie, mais il n'est pas accumulé. L'étude expérimentale de cet état se fait donc par des expériences de cinétiques. L'idée de Fersht est d'introduire des changements à l'intérieur de la protéine et de déduire des informations sur cet état de transition (ou un autre état intermédiaire) à partir des effets observés sur la cinétique. Les perturbations introduites au sein de la protéine sont

traditionnellement des mutations. La comparaison de la variation d'énergie de l'état intermédiaire et de l'état natif permet de déterminer si le changement effectué influe sur l'état intermédiaire. Par exemple, si la variation d'énergie due à une mutation est la même sur l'état natif et sur l'état intermédiaire, l'acide aminé muté a le même rôle structural dans ces deux états : si l'acide aminé en question est structuré de la même manière dans l'état intermédiaire et dans l'état natif, la mutation de cet acide aminé aura des effets similaires pour l'état intermédiaire et l'état natif. En revanche, si la variation d'énergie de l'état intermédiaire est négligeable par rapport à celle de l'état natif, l'acide aminé substitué n'aura pas la même structure dans ces deux états : si l'acide aminé est structuré dans l'état natif mais pas dans l'état intermédiaire, la mutation de cet acide aminé influencera davantage l'état natif que l'état intermédiaire.

Pour mesurer cet effet, on utilise un paramètre appelé  $\Phi$ . Cette valeur correspond au rapport de la variation d'énergie de l'état intermédiaire sur la variation d'énergie de l'état natif.

Les variations d'énergie sont traditionnellement induites par des mutations, cependant d'autres types de perturbations peuvent être utilisées, une variation de pH ou de température par exemple. Ici, la perturbation utilisée est la fixation du métal. Si celle-ci induit une variation d'énergie identique sur l'état natif et sur un état intermédiaire (*molten globule* ou état de transition), le site de liaison sera le même dans cet état intermédiaire et dans l'état natif ;  $\Phi$  sera proche de 1. *A contrario*, si la variation d'énergie induite sur l'état intermédiaire est négligeable par rapport à celle induite sur l'état natif, le site de liaison sera peu structuré dans l'état intermédiaire par rapport à l'état natif et  $\Phi$  sera proche de 0.

Pour calculer ces différences d'énergie, il faut choisir un état de référence dont l'énergie n'est pas modifiée par la perturbation introduite. L'état complètement dénaturé est généralement,

choisi comme état de référence. Dans nos expériences, le repliement est étudié de l'état *molten globule* à l'état natif ; l'énergie de la protéine dans son état complètement dénaturé est inaccessible. L'état natif sous sa forme apo (1 mM EDTA) ou holo (30 mM Ca<sup>2+</sup>) est donc pris comme état de référence.

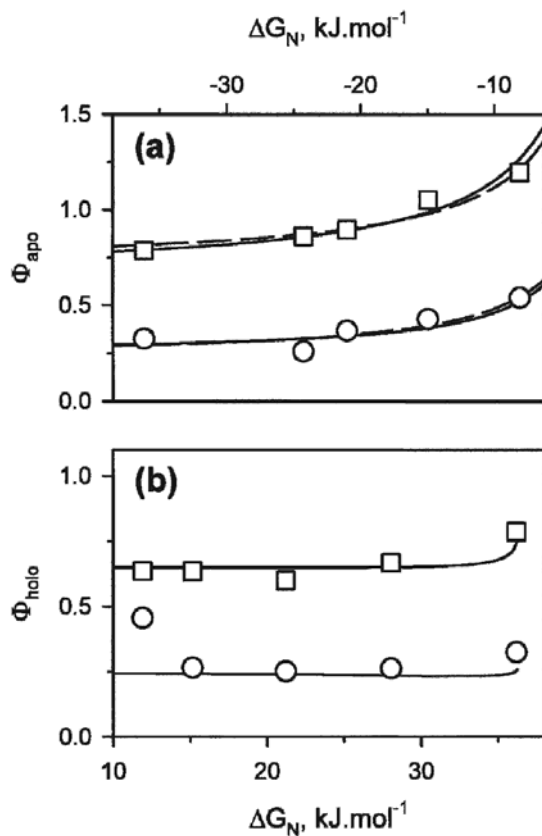
### Calcul des valeurs $\Phi$

Les variations d'énergie nécessaires pour le calcul des valeurs  $\Phi$  sont déterminées en utilisant les constantes de dissociation grâce à l'équation suivante<sup>151</sup>

$$\Delta G = -RT \ln(1 + [M]/K_{diss})$$

Cette équation donne la différence d'énergie entre une protéine liée à un métal et une protéine non liée. Elle peut donc être utilisée telle quelle lorsque la forme apo est choisie comme référence. Lorsque la forme holo est prise comme référence, cette équation est utilisée deux fois : tout d'abord pour calculer la différence d'énergie entre la protéine liée au métal et la protéine non liée, puis pour calculer la différence d'énergie entre la protéine non liée et la protéine liée au calcium (forme de référence). La différence entre les deux énergies calculées sera alors égale à la différence d'énergie entre une protéine liée à un métal et la forme de référence (protéine liée au calcium)

Les rapports des  $\Delta G$  des états intermédiaires sur les  $\Delta G$  de l'état natif sont les valeurs  $\Phi$  représentées sur la figure II.8.



**Fig II.8 :** Valeurs de  $\Phi$  obtenues pour le molten globule ( $\circ$ ) et l'état de transition ( $\square$ ).

a) La forme apo est utilisée comme référence. Les lignes discontinues correspondent au paramétrage qui prend en compte l'erreur induite par le choix de la forme apo comme référence.

b) La forme holo est utilisée comme référence.

Les lignes continues correspondent au paramétrage effectué en utilisant les régressions linéaires de la figure 7.b.

Cette figure est divisée en deux parties : la figure II.8a présente les  $\Phi$  calculés en prenant la forme apo de la protéine comme référence, la figure II.8b présente les  $\Phi$  calculés en prenant la forme holo comme référence. Dans les deux cas, les  $\Phi$  sont représentés en fonction de la variation d'énergie de l'état natif  $\Delta G_N$ . Lorsque la forme apo est utilisée, la fixation de l'ion sur l'état natif stabilise la protéine, son énergie est donc diminuée et le  $\Delta G_N$  est négatif. En revanche, lorsque la forme holo est utilisée, la fixation d'un ion autre que le calcium sur l'état natif déstabilise cet état. En effet, le calcium est l'ion le mieux adapté pour la protéine et c'est celui qui la stabilise le mieux. L'énergie de la protéine dans son état natif sera donc augmentée et le  $\Delta G_N$  sera positif.

Pour chaque figure, les  $\Phi$  calculés pour l'état de transition et l'état *molten globule* sont représentés. Les perturbations induites pour calculer ces  $\Phi$  sont dues à la fixation des différents métaux (et l'absence de fixation de métal lorsque la forme holo est prise comme référence).



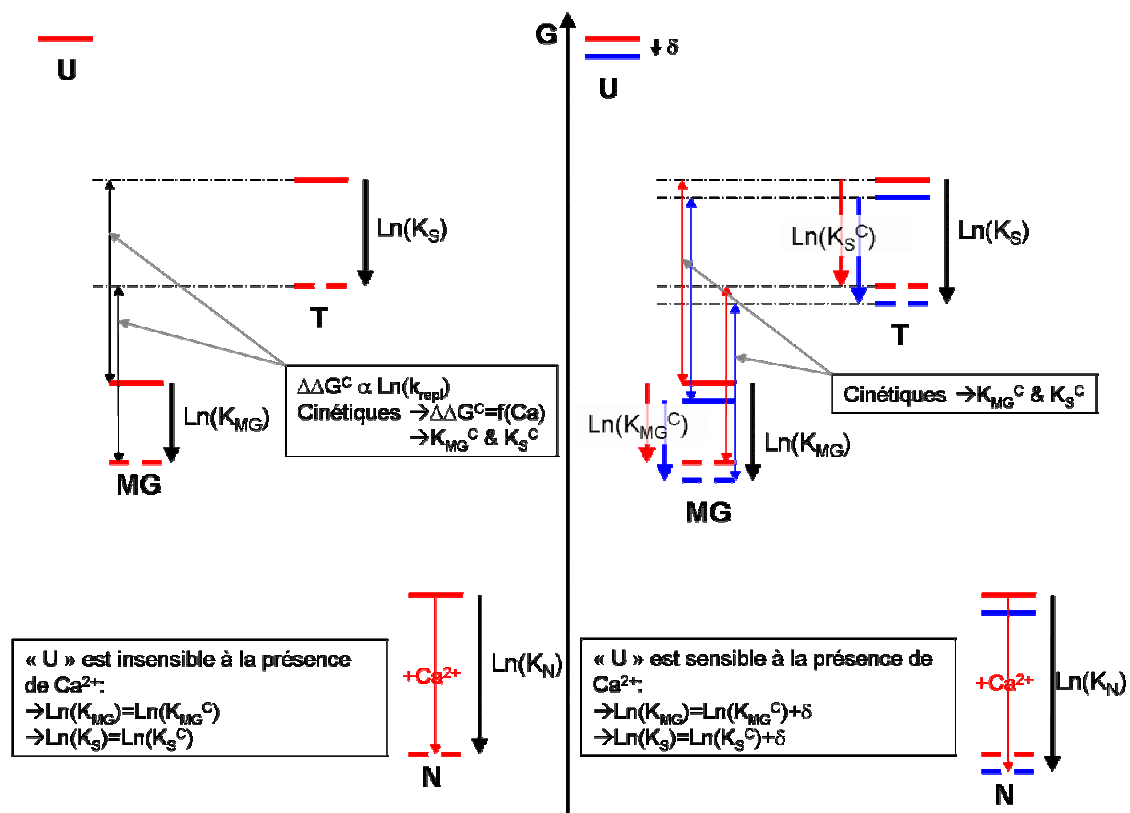
Sur la figure II.8a, les  $\Phi$  calculés pour l'état de transition sont compris entre 0,75 et 1,2, ils sont croissants avec le  $\Delta G_N$ . Les  $\Phi$  calculés pour l'état MG sont eux aussi croissants avec le  $\Delta G_N$  mais les valeurs sont moins élevées : ils sont compris entre 0,25 et 0,5. Sur la figure II.8b, les valeurs de  $\Phi$  sont constantes pour des faibles perturbations puis légèrement croissantes pour le dernier point qui correspond aux expériences réalisées en présence d'EDTA. La valeur de  $\Phi$  est d'environ 0,7 pour l'état de transition et de 0,25 pour l'état *molten globule*.

Dans le cas idéal, les valeurs de  $\Phi$  sont comprises entre 0 et 1 et constantes sur une large plage de  $\Delta G$  (ref 71). Ce n'est pas le cas lorsque la forme apo est prise comme état de référence (figure II.8a). Cette déviation des valeurs idéales peut être attribuée au fait que l'énergie de l'état de référence varie lorsque les ions métalliques sont introduits en solution. Les différentes méthodes de calcul utilisées pour déterminer les constantes de dissociation ne prennent pas toutes en compte cette variation de l'énergie de l'état déplié.

Les constantes de dissociation entre l'état *molten globule* et les métaux, et entre l'état de transition et les métaux sont déterminées par des expériences de cinétique. Elles dépendent donc uniquement de la variation de la hauteur de la barrière d'énergie que la protéine doit franchir pour se replier. Les changements d'énergie qui peuvent avoir lieu sur l'état déplié de la protéine n'ont pas d'effet sur la hauteur de cette barrière d'énergie et n'ont pas d'influence sur les constantes de dissociation.

La constante de dissociation entre le métal et l'état natif est déterminée par des expériences à l'équilibre, en paramétrant les  $A$  obtenus pour différentes concentrations en métal. Elle reflète la stabilisation de l'état natif lorsque l'ion métallique est lié. Comme la stabilité de la protéine

est définie par rapport à l'énergie de l'état déplié, des éventuels changements d'énergie de l'état déplié sont pris en compte dans le calcul de cette constante de dissociation.



**Fig II.9 :** Variation d'énergie et calcul de  $\Phi$ .

La partie à gauche représente le cas où l'énergie de l'état déplié ne varie pas en présence de métaux. La partie à droite représente le cas où les métaux modifient l'énergie de l'état déplié.

U est l'état déplié, MG l'état molten globule, T l'état de transition et N l'état natif.

Les traits pleins sont les niveaux d'énergie en absence de métaux, les traits en pointillés sont les niveaux d'énergie en présence de métaux.

Les traits en rouge ne prennent pas en compte l'effet des métaux sur l'énergie de l'état dépliées, les traits en bleu prennent cet effet en compte.

La figure II.9 schématise les différences d'énergie utilisées pour calculer les  $\Phi$  et la manière dont elles ont été obtenues. Cette figure est divisée en 2 parties. Dans la partie gauche, la perturbation qui modifie les niveaux d'énergie des états natifs, *molten globule* et de transition

ne modifie pas l'énergie de l'état déplié. Sur la partie droite, la présence de métaux a un effet sur l'état déplié.

La barrière d'énergie que la protéine doit franchir pour se replier est représentée par les flèches fines dans la figure II.9. C'est sa variation avec la concentration en calcium qui permet de déterminer les constantes de dissociation du métal avec les états *molten globule* et de transition. Cette variation de la hauteur d'énergie est accessible par les expériences de cinétique et les constantes de dissociation sont notées  $K^C$  ( $K$  cinétique).

Lorsque la perturbation qui modifie la hauteur de la barrière d'énergie ne modifie pas l'énergie de l'état déplié, les variations d'énergie calculées à partir des  $K^C$  reflètent bien la stabilisation des états *molten globule* et de transition (Figure II.9 partie gauche).

Lorsque la perturbation qui modifie les énergies des états *molten globule* et de transition modifie aussi l'état déplié (Figure II.9 partie droite), la connaissance de la variation de la hauteur de la barrière d'énergie ne suffit plus pour estimer la stabilisation des états intermédiaires. Les variations d'énergie calculées à partir des  $K^C$  (flèches de couleur sur la figure) reflètent uniquement les effets sur la barrière d'énergie sans tenir compte des effets sur l'état déplié. Pour prendre en considération ces effets, un terme supplémentaire doit être ajouté dans le calcul des énergies. Sur la figure II.9, Les différences d'énergie qui prennent en compte ces effets sont représentées par des flèches noires.

La constante de dissociation entre les métaux et l'état natif est calculée directement à partir des courbes de liaison et elle tient compte de la variation d'énergie de l'état déplié. Les variations d'énergie calculées à partir de ces constantes reflètent donc bien des variations d'énergie « réelles ».

L'énergie de l'état déplié varie en présence de métal. Ces effets sont non spécifiques et tous les métaux ont le même effet sur l'état déplié. Lorsque les  $\Phi$  sont calculés pour des solutions en présence de métal et que la référence est la forme apo, ou lorsque le  $\Phi$  est calculé sur la

solution d'EDTA et que la référence est la protéine en présence de calcium, l'énergie de l'état déplié n'est pas la même pour l'état de référence et l'état considéré. Cela explique les valeurs anormales de  $\Phi$  dans la figure II.8a. Dans cette figure, la forme apo est prise comme référence et les  $\Phi$  sont calculés pour les solutions de protéine en présence des différents métaux. Cela explique aussi que le dernier point des deux courbes de la figure II.8b soit légèrement supérieur aux autres. Ce dernier point correspond au  $\Phi$  calculé en présence d'EDTA en prenant comme référence la protéine en présence de calcium.

Les effets non-spécifiques des métaux sur l'état déplié ainsi que les  $\Phi$  réels peuvent être calculés en paramétrant les  $\Phi$  représentés dans la figure II.8.

### Paramétrage

La variation d'énergie due aux effets non spécifiques est évaluée en paramétrant les valeurs  $\Phi$  obtenues avec la forme apo comme référence. Ce paramétrage permet aussi d'estimer les valeurs  $\Phi$  « réelles ».

Lorsque la forme apo est prise comme référence, les effets non spécifiques sont ressentis pour tous les métaux, il faudra donc les prendre en compte pour le calcul des valeurs  $\Phi$ . Or, les constantes de dissociation ont été calculées sur les états intermédiaires (MG et T) à partir des expériences de cinétique. Ils ne prennent pas en compte les effets non spécifiques sur l'état déplié. Le  $\Phi$  calculé est donc égal à :

$$\Phi_c = \frac{\Delta G_I - \Delta G_I^{n-spec}}{\Delta G_N}$$

Ainsi, pour avoir accès aux valeurs  $\Phi$  réelles, on doit prendre en compte les effets non spécifiques du métal pour les états intermédiaires. Les courbes de  $\Phi$  calculées sont paramétrées pour retrouver les  $\Phi$  réels et les effets non spécifiques des métaux sur l'état déplié.

$$\Phi_r = \Phi_c + \frac{\Delta G_I^{n-spec}}{\Delta G_N}$$

Ce paramétrage est représenté par une ligne discontinue sur la figure II.8a.

Le paramétrage des valeurs de  $\Phi$  par cette équation donne  $\Delta G_T^{n-spec} = -4,3 \pm 0,6$  kJ/mol et  $\Phi_r = 0,70 \pm 0,05$  pour l'état de transition et  $\Delta G_{MG}^{n-spec} = -2,6 \pm 0,7$  kJ/mol et  $\Phi_r = 0,22 \pm 0,04$  pour l'état *molten globule*.

Les  $\Delta G^{n-spec}$  sont du même ordre de grandeur pour l'état de transition et l'état *molten globule*.

Les  $\Phi$  réels déduits de ces paramétrages correspondent bien au  $\Phi$  calculés avec la forme holoprise comme référence (figure II.8b).

Les valeurs  $\Phi$  calculées peuvent aussi être obtenues en utilisant les relations entre  $\ln(K)$  représentées dans la figure II.7b.

La variation d'énergie due à la liaison du métal sur la protéine est liée à la constante de dissociation par l'équation<sup>151</sup>  $\Delta G = -RT \ln(1 + [M]/K_{diss})$ .

Lorsque la forme apo est choisie comme référence, les valeurs  $\Phi$  suivront donc l'équation suivante :

$$\Phi = \frac{\ln(1 + [M]/K_I)}{\ln(1 + [M]/K_N)}$$

Lorsque la forme holo est choisie comme référence, les différences d'énergie sont calculées par rapport à la protéine ayant liée le calcium. Les valeurs de  $\Phi$  suivront cette fois-ci l'équation :

$$\Phi = \frac{\ln(1 + [M]/K_{I-Ca}) - \ln(1 + [M]/K_I)}{\ln(1 + [M]/K_{N-Ca}) - \ln(1 + [M]/K_N)}$$

Les  $K_{I-Ca}$  et  $K_{N-Ca}$  sont connues. Les relations entre les  $\ln(K_I)$  et  $\ln(K_N)$  sont linéaires, elles sont représentées dans la figure II.7b. Les deux équations précédentes peuvent donc être tracées en fonction de  $\Delta G_N$  qui dépend uniquement de  $K_N$ . Les courbes obtenues correspondent bien aux  $\Phi$  calculés (traits pleins, figure II.8).

### Conclusions sur l'analyse des valeurs $\Phi$

Comme expliqué précédemment, l'analyse des valeurs  $\Phi$  permet d'obtenir des informations sur les états intermédiaires. Dans notre cas, elles permettent de quantifier la structuration du site de liaison des métaux.

Pour l'état de transition, la valeur  $\Phi$  est d'environ 0,7, le site de liaison dans cet état ressemble donc à celui de l'état natif. En effet, les stabilisations des états natifs et de transition par la liaison d'un métal sont similaires.

Avec un  $\Phi$  proche de 0,22, l'état molten globule possède un site de liaison beaucoup moins structuré que l'état natif puisque la liaison d'un métal sur la protéine stabilise beaucoup moins l'état *molten globule* que l'état natif.

## II.3. Etude des paramètres thermodynamiques de la réaction de repliement

Pour avoir une description plus complète, les paramètres thermodynamiques de la réaction ou plus particulièrement la variation de ces paramètres en présence de métaux sont étudiés.

La grandeur thermodynamique la plus utilisée en biologie est l'enthalpie libre, aussi appelée énergie de Gibbs, et notée G. Cette grandeur est utilisée lorsque les expériences se déroulent à température et pression constante, ce qui est le cas dans beaucoup d'expériences de biophysique. C'est cette grandeur qui a été utilisée jusqu'à présent pour décrire les énergies.

L'enthalpie, notée H, est utilisée généralement lorsque la pression est constante mais que la température varie.

L'entropie, notée S est une mesure du désordre. Cette grandeur thermodynamique est difficilement accessible. Elle peut être estimée si l'enthalpie libre et l'enthalpie sont connues.

Les «  $\Delta\Delta$  » sont utilisés dans cette étude thermodynamique pour pouvoir s'affranchir du  $\kappa$  de l'expression de  $\Delta G$  à partir de la théorie de l'état de transition.

$$k_{repl} = \kappa \left( \frac{kT}{h} \right) \exp \left( - \frac{\Delta G}{RT} \right)$$

$$\Delta\Delta G = RT \ln(k_{max} / k_{apo})$$

Les  $\Delta\Delta H$  sont ensuite déterminés et ils sont utilisés pour estimer les  $\Delta\Delta S$ .

### II.3.1. Variation d'énergie de Gibbs

L'énergie de Gibbs ( $\Delta G$ ) de la réaction correspond à la hauteur de la barrière d'énergie que la protéine doit franchir pour pouvoir se replier. Cette hauteur, ou plus particulièrement la variation de cette hauteur en présence d'ion métallique est facilement accessible grâce aux constantes de vitesse de repliement déterminées par fluorescence. En effet, la variation d'énergie est calculée à l'aide de la formule suivante :

$$\Delta\Delta G = RT \ln(k_{\max}/k_{\text{apo}})$$

Avec  $k_{\max}$ , la constante de vitesse de repliement en présence de métaux et  $k_{\text{apo}}$  la constante de vitesse de repliement de la forme apo.

Les variations de la hauteur de la barrière d'énergie déterminées par cette méthode pour chacun des métaux sont reportées dans la figure II.11. Cette figure sera décrite plus tard.

Pour avoir une description thermodynamique plus complète, la variation d'enthalpie induite par la fixation du métal est déterminée.

### II.3.2 Variation d'enthalpie – courbes d'Eyring

La variation d'enthalpie des réactions en présence de métaux ne peut pas être calculée en utilisant les résultats obtenus jusqu'à présent. En effet, comme expliqué ci-dessus, l'enthalpie est utilisée lorsque la température varie. Pour calculer cette grandeur thermodynamique, les courbes d'Eyring sont utilisées. Le repliement est suivi à différentes températures en présence et en absence de métaux, et les constantes de vitesse sont reportées en fonction de la



température. Un traitement des courbes permet d'obtenir le  $\Delta H$  de chacune des réactions. Le  $\Delta\Delta H$  peut alors être obtenu en soustrayant le  $\Delta H$  en absence de métal au  $\Delta H$  en présence de métal.

### Principe des courbes d'Eyring

La théorie de l'état de transition nous indique que

$$k_{repl} = \kappa \left( \frac{kT}{h} \right) \exp\left( -\frac{\Delta G}{RT} \right)$$

Or,

$$\Delta G = \Delta H - T\Delta S$$

Donc,

$$k_{\max} = \kappa \left( \frac{kT}{h} \right) \exp\left( \frac{-\Delta H + T\Delta S}{RT} \right) = \left( \frac{kT}{h} \right) \exp\left( \frac{-\Delta H}{RT} \right) \exp\left( \frac{\Delta S}{R} \right)$$

D'où,

$$\ln\left( \frac{k_{\max}}{T} \right) = \ln\left( \kappa \frac{k}{h} \right) - \left( \frac{\Delta H}{RT} \right) + \left( \frac{\Delta S}{R} \right)$$

Ainsi, en mesurant les constantes de vitesse de repliement à différentes températures et en reportant les  $\ln(k_{\max}/T)$  sur un graphe en fonction de  $1/T$ , on obtient une droite de pente  $-\Delta H/R$ . La figure représentant les  $\ln(k_{\max}/T)$  en fonction de  $1/T$  est appelé courbe d'Eyring.

### Courbes d'Eyring

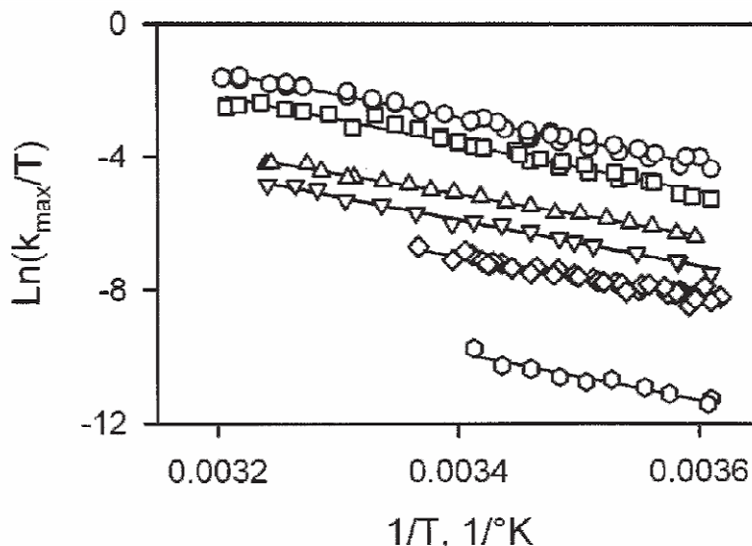
#### *Conditions expérimentales*

Les expériences de cinétique sont réalisées de la même manière que celles qui ont permis de déterminer les constantes de vitesse de repliement en présence de différentes concentrations de métaux (Partie II.1.1). Mais dans ce cas, la concentration en ion métallique est constante et égale à 30 mM, et plusieurs températures sont utilisées. Lorsque l'EDTA est utilisée pour étudier la forme apo, sa concentration sera de 1 mM.

La température minimum est 277 K. Les températures maximums varient en fonction de l'ion en solution : pour la forme apo la température maximum est 293K, elle est égale à 298 K pour  $Mg^{2+}$ , 308 K pour  $Ba^{2+}$  et  $Mn^{2+}$  et 312 K pour  $Ca^{2+}$  et  $Sr^{2+}$ . Au-delà de ces températures, la protéine ne se replie pas.

#### *Résultats*

Les courbes d'Eyring sont représentées sur la figure II.10 pour les cinq métaux d'intérêts et pour la forme apo. Sur cette figure, les  $\ln(k_{\max}/T)$  sont représentés en fonction de  $1/T$ . A chaque symbole correspond un métal. Les valeurs de  $\ln(k_{\max}/T)$  les plus élevées sont celles mesurées en présence de calcium, viennent ensuite les différents métaux. Pour la forme apo, les  $\ln(k_{\max}/T)$  sont beaucoup plus faibles qu'en présence de métaux.



**Fig II.10 : Graphes d'Eyring**

Ces graphes sont tracés à partir d'expériences de repliement effectuées à différentes températures en présence de différents métaux et d'EDTA: (○) Ca<sup>2+</sup>, (□) Sr<sup>2+</sup>, (△) Mn<sup>2+</sup>, (▽) Ba<sup>2+</sup>, (◇) Mg<sup>2+</sup>, (○) EDTA

Pour chaque métal, une régression linéaire est effectuée et est représentée sur la figure par une ligne pleine. D'après les explications précédentes, la pente de la droite obtenue est égale  $-\Delta H/R$ . Les valeurs de l'enthalpie d'activation pour chacun des métaux sont ainsi déterminées. Elles sont reportées dans la table II.1. Elles varient entre 50 et 60 kJ/mol.

Une grande différence est observée dans les constantes de vitesse de repliement, et donc dans les  $\ln(k_{\max}/T)$  : le repliement de la forme apo est environ 1000 fois plus lent que le repliement en présence de calcium. Sur la figure, les  $\ln(k_{\max}/T)$  varient entre -1 et -3 alors qu'en présence d'EDTA, les  $\ln(k_{\max}/T)$  varient entre -9 et -11. Cependant, malgré cette grande différence, les pentes des droites formées par ces points sont assez proches. Ceci indique que les variations d'enthalpie de la barrière d'énergie dues à la fixation du métal sont relativement faibles.

La variation d'enthalpie due à la fixation du métal est calculée en soustrayant les valeurs mesurées en présence de métal à la valeur en absence de métal. Les résultats obtenus sont reportés dans la figure II.11. Ils seront commentés plus longuement par la suite.

### II.3.3 Variation d'entropie

Les études précédentes permettent de calculer l'influence de la liaison du métal sur la réaction de repliement, au niveau de l'énergie de Gibbs et de l'enthalpie. La connaissance de ces deux fonctions thermodynamiques permet de mesurer l'entropie de la réaction grâce à la relation :

$$\Delta G = \Delta H - T\Delta S$$

Donc

$$\Delta\Delta G = \Delta\Delta H - T\Delta\Delta S$$

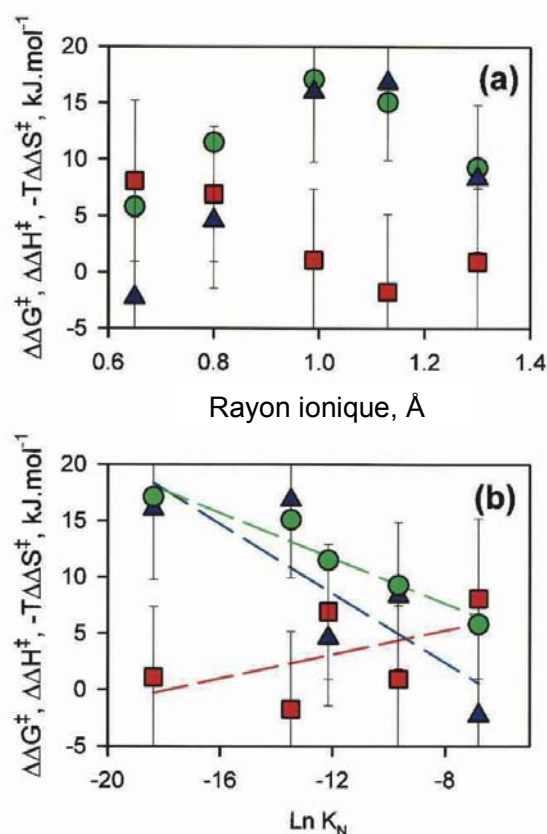
D'où

$$\Delta\Delta S = \frac{\Delta\Delta H - \Delta\Delta G}{T}$$

Les valeurs de  $\Delta\Delta S$  sont reportées sur la figure II.11.

### II.3.4 Résultats - Interprétation

La figure II.11 représente les variations des différentes grandeurs thermodynamiques de la barrière d'énergie que la protéine doit franchir pour se replier. Les variations d'énergie de Gibbs sont représentées par des ronds verts, les variations d'enthalpie par des carrés rouges et les variations d'entropie par des triangles bleus.



**Fig II.11 :** Paramètres thermodynamiques de la barrière d'énergie à franchir pour le repliement.

Les variations d'énergie de Gibbs  $\Delta\Delta G$  (●), d'enthalpie  $\Delta\Delta H$  (■) et d'entropie  $\Delta\Delta S$  (▲) (ou plus exactement  $-T\Delta\Delta S$ ) de la barrière d'énergie sont représentées en fonctions :  
a) du rayon des ions en question.  
b) du logarithme de leur constante d'affinité avec l'état natif de la protéine.

Dans la figure II.11a ces grandeurs sont représentées en fonction du rayon ionique du métal utilisé. Les courbes obtenues sont des courbes en cloche centrées sur le calcium. La variation d'énergie de Gibbs augmente avec le rayon ionique jusqu'à 1 Å, puis elle diminue au-delà ; elle est comprise entre 5 et 18  $\text{kJ}\cdot\text{mol}^{-1}$ . La variation d'énergie de Gibbs maximum est mesurée pour le calcium.

La variation de  $-T\Delta\Delta S$  a le même comportement, la valeur de  $\Delta\Delta S$  minimum est obtenue pour des rayons ioniques proches de 1 Å. La variation d'entropie est comprise entre -3 et 18 kJ/mol.

La variation d'enthalpie a un comportement différent : lorsque le rayon ionique augmente, elle diminue pour atteindre des valeurs minimales pour le calcium et le strontium, et elle augmente ensuite. Elle est comprise entre -2 et 8 kJ/mol.

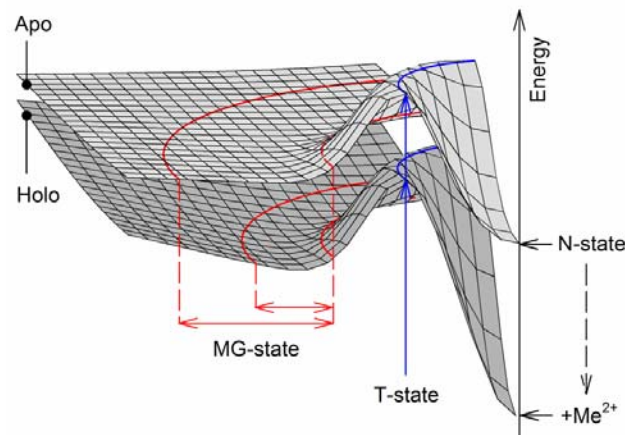
Dans la figure II.11b, ces mêmes paramètres thermodynamiques sont représentés en fonction du logarithme de la constante de dissociation entre le métal et l'état natif de la protéine. Le logarithme de cette constante de dissociation est lié à la stabilisation de l'état natif par les ions. Cette représentation permet donc de mettre en évidence les paramètres thermodynamiques les plus significatifs pour la fixation de l'ion. Les symboles utilisées pour représenter les différentes grandeurs thermodynamiques sont les mêmes que ceux de la figure II.11a.

Une régression linéaire est effectuée sur chacune des grandeurs thermodynamiques pour souligner leur évolution avec la fixation à la protéine. D'après ces régressions linéaires, et malgré une grande dispersion dans ces données, l'entropie (en bleu) est le paramètre thermodynamique qui varie le plus en présence de métaux. Ces résultats suggèrent que l'effet des métaux sur la réaction de repliement est principalement de nature entropique.

Ces paramètres thermodynamiques permettent d'estimer l'effet des métaux sur le paysage d'énergie de la protéine et donc sur son repliement.

La figure II.12 représente une superposition de deux paysages d'énergie simplifiés. Le paysage de plus haute énergie est celui de la forme apo, le paysage de plus basse énergie est celui en présence de métaux. Sur ces paysages, l'état de plus basse énergie est l'état natif. Les configurations microscopiques qui correspondent à l'état *molten globule* sont délimitées par

deux traits rouges. L'état de transition se trouve au sommet de la bosse entre l'état *molten globule* et l'état natif.



**Fig II.12 :** Effets des ions métalliques sur le paysage d'énergie de la protéine.

Deux paysages d'énergie, en absence et en présence de métaux, sont superposés.

Les paramètres thermodynamiques déterminés précédemment permettent de schématiser:

En bleu : Variation de la hauteur de la barrière d'énergie.

En rouge : configurations accessible dans l'état molten globule.

Le paysage en présence de métaux a une énergie plus faible que celui de la forme apo. Cette différence est due à la stabilisation de la protéine lorsque celle-ci fixe un ion métallique. Cet effet est particulièrement visible sur l'état natif pour lequel la fixation d'un ion métallique a un effet très stabilisateur. La figure II.11 montre que l'entropie est le paramètre qui varie le plus lorsqu'un ion métallique se fixe sur la protéine. Cette diminution d'entropie est visible sur ces paysages d'énergie en observant l'état *molten globule*.

Sur le paysage d'énergie de la forme apo, l'état *molten globule* possède une grande entropie conformationnelle. De nombreuses configurations microscopiques correspondent à un état *molten globule*. Le nombre de ces configurations correspond à l'aire entre les deux traits rouges.

Sur le paysage d'énergie de la forme holo, l'aire entre les deux traits rouges est beaucoup plus petite. Le nombre de configurations microscopiques qui correspondent à l'état *molten globule* est beaucoup moins élevé qu'en absence de métaux, et donc l'entropie conformationnelle de cet état est plus petite. Ceci explique les variations d'entropie observées expérimentalement et d'après les pentes des régressions linéaires tracées sur la figure II.11b, cet effet est celui qui a le plus d'importance dans la fixation des métaux à la protéine et l'accélération du repliement.



## **II.4. Dénaturation thermique**

La partie précédente a permis d'explorer l'influence des métaux sur le repliement et sur la barrière d'énergie que doit franchir la protéine pour se replier. Une étude de l'influence du calcium sur la stabilité de l'état natif permet de déterminer plus précisément l'influence de la liaison des métaux sur le paysage d'énergie. Pour étudier la stabilité de l'état natif, la protéine est dépliée, dénaturée. Cette dénaturation est effectuée ici par augmentation de la température.

### **II.4.1 Dénaturation thermique suivie par dichroïsme circulaire**

#### **Introduction**

La dénaturation est le processus par lequel une macromolécule biologique perd sa structure tridimensionnelle. Plusieurs conditions physiques ou chimiques peuvent provoquer la dénaturation. Les deux principales méthodes sont l'utilisation d'agent dénaturant et la dénaturation thermique.

Dans le cas de l'utilisation de dénaturant (guanidine ou urée par exemple), la concentration de dénaturant nécessaire à la déstructuration de la protéine permet d'en estimer la stabilité. Plus la protéine est stable, plus elle est difficile à dénaturer et plus la concentration en agent dénaturant doit être grande.

La chaleur permet aussi de déstructurer les protéines. Une élévation de température provoque la rupture des interactions qui stabilisent la structure native des protéines. La structure native se désagrège ; la température à laquelle ce phénomène a lieu permet de déterminer des paramètres thermodynamiques caractéristiques des états peuplés par la protéine.

Dans le cas de l' $\alpha$ -lactalbumine, la dénaturation thermique est préférée pour éviter que des agents dénaturants interfèrent dans l'interaction de la protéine avec le calcium.

Les différentes techniques qui permettent de détecter les structures des protéines peuvent être aussi utilisées pour détecter la destruction de ces structures. Les spectroscopies d'absorbance et de dichroïsme circulaire sont majoritairement utilisées du fait de leur relative facilité de mise en oeuvre. Lorsque la température augmente et que les structures sont détruites, le signal spectroscopique associé à ces structures tend à disparaître.

Pour la dénaturation de l' $\alpha$ -lactalbumine, le dichroïsme circulaire est préféré car il présente l'avantage de détecter la structure secondaire ou tertiaire en fonction des longueurs d'onde utilisées.

Les températures de fusion des structures secondaire et tertiaire ainsi que les variations d'enthalpie sont déduites du paramétrage de ces courbes de dénaturation. Elles donnent des informations sur le paysage d'énergie de la protéine.

### Conditions expérimentales

Les protéines sont dans une solution de Tris 100 mM à pH 7.5. La concentration de calcium varie entre 30  $\mu$ M et 30 mM ; on utilise de l'EDTA à 1 mM pour la forme apo. La dénaturation est suivie à 280 nm avec une concentration en protéine de 0.3 mg/mL dans l'UV proche. Le signal dans l'UV lointain est plus intense et une concentration de 0,02 mg/mL est utilisée ; dans ce cas la longueur d'onde est de 230 nm.

La température varie entre 5 et 95°C à la vitesse de 1°C par minute.

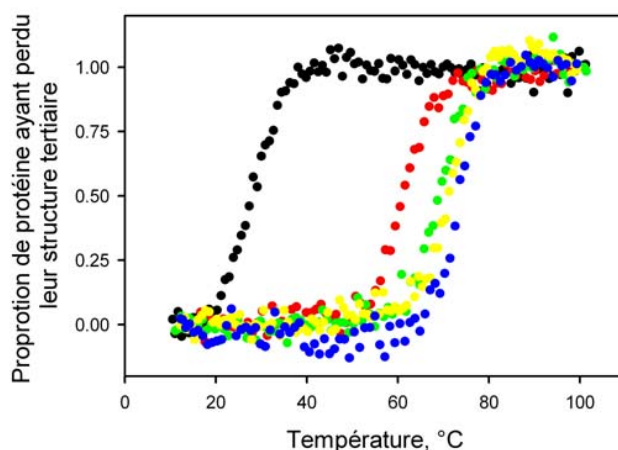
## Résultats

### *UV proche*

Le dichroïsme circulaire dans l'UV proche (250 – 320 nm) donne accès à la structure tertiaire des protéines. La longueur d'onde employée pour suivre le phénomène est de 280 nm. La mesure de dichroïsme circulaire à cette longueur donne un signal négatif caractéristique de la présence de structure tertiaire. Il disparaît au fur et à mesure que la température augmente et que la protéine se détruit. Des courbes présentant le signal de dichroïsme circulaire en fonction de la température sont ainsi tracées.

Pour permettre une lecture plus facile des résultats obtenus, ces courbes sont normalisées pour avoir accès à la quantité de structure tertiaire détruite, et donc à la proportion de protéine qui a perdu sa structure tertiaire.

Ces courbes de dénaturation sont représentées sur la figure II.13 pour différentes concentrations de calcium.



**Fig II.13 : Dénaturation thermique suivie par dichroïsme circulaire dans l'UV proche.**

Ces courbes correspondent à la proportion de protéines qui ont perdu leur structure tertiaire en fonction de la température. Différentes concentrations de Ca<sup>2+</sup> sont utilisées : (●) EDTA, (●) 30 μM Ca<sup>2+</sup>, (●) 300 μM Ca<sup>2+</sup>, (●) 3 mM Ca<sup>2+</sup> et (●) 30 mM Ca<sup>2+</sup>.

En présence d'EDTA (points noirs sur la figure), la structure tertiaire de la forme apo de la protéine est très rapidement dénaturée. A 30°C, la courbe atteint une valeur proche de 1, ce qui indique que le signal de dichroïsme circulaire est nul et qu'il n'y a plus de structure tertiaire dans la solution.

Dès qu'on introduit un peu de calcium dans la solution (points rouges : 30 μM de calcium), la température de dénaturation s'élève aux alentours de 60°C. Elle augmente légèrement avec la concentration en calcium pour atteindre des valeurs entre 70 et 75 °C.

La température de dénaturation correspond à la température pour laquelle la moitié des protéines est dénaturée tandis que l'autre moitié est encore structurée. Les courbes de dénaturation peuvent être paramétrées. Ce paramétrage permet de déterminer les températures de dénaturation et les ΔH de ces dénaturations. Les ΔH déterminés par cette méthode sont les ΔH de Van't Hoff.

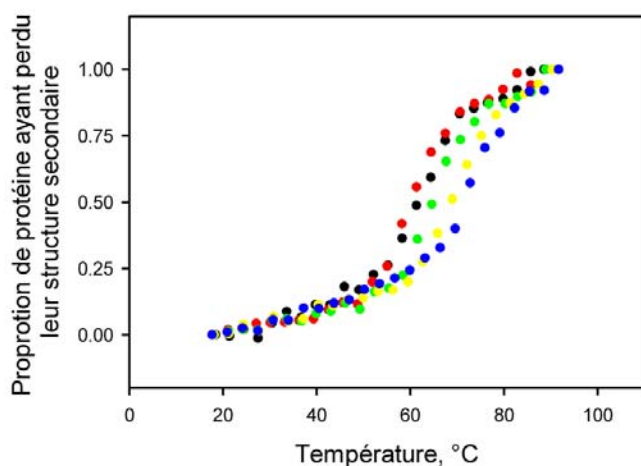
L'équation permettant le paramétrage de ces courbes est la suivante<sup>71</sup> :

$$\%Dénat = \frac{\exp\left(-\Delta H_{vH} \frac{\frac{1}{T} - \frac{1}{T_m}}{R}\right)}{1 + \exp\left(-\Delta H_{vH} \frac{\frac{1}{T} - \frac{1}{T_m}}{R}\right)}$$

Les paramètres calculés par cette méthode sont reportés sur la figure II.15 qui sera décrite par la suite.

#### *UV lointain*

Le dichroïsme circulaire dans l'UV lointain (190 - 250nm) permet de détecter la structure secondaire des protéines. La longueur d'onde employée pour suivre le phénomène est de 230 nm. Pour cette longueur d'onde, le signal de dichroïsme circulaire de l' $\alpha$ -lactalbumine est négatif. Lorsque la température augmente, la structure secondaire est détruite et ce signal disparaît. Les courbes de dénaturation sont normalisées pour permettre une lecture plus facile des résultats. Elles sont représentées dans la figure II.14.



**Fig II.14 : Dénaturation thermique suivie par dichroïsme circulaire dans l'UV lointain.**

Ces courbes correspondent à la proportion de protéines qui ont perdu leur structure secondaire en fonction de la température. Différentes concentrations de  $\text{Ca}^{2+}$  sont utilisées : (●) EDTA, (●)  $30\mu\text{M Ca}^{2+}$ , (●)  $300\mu\text{M Ca}^{2+}$ , (●)  $3\text{mM Ca}^{2+}$  et (●)  $30\text{mM Ca}^{2+}$ .

En absence de calcium, la dénaturation de la structure secondaire est très différente de la dénaturation de la structure tertiaire. En effet, alors que la structure tertiaire est complètement détruite au delà de  $30^\circ\text{C}$  (Figure II.13), la structure secondaire est intacte jusqu'à  $50^\circ\text{C}$  environ.

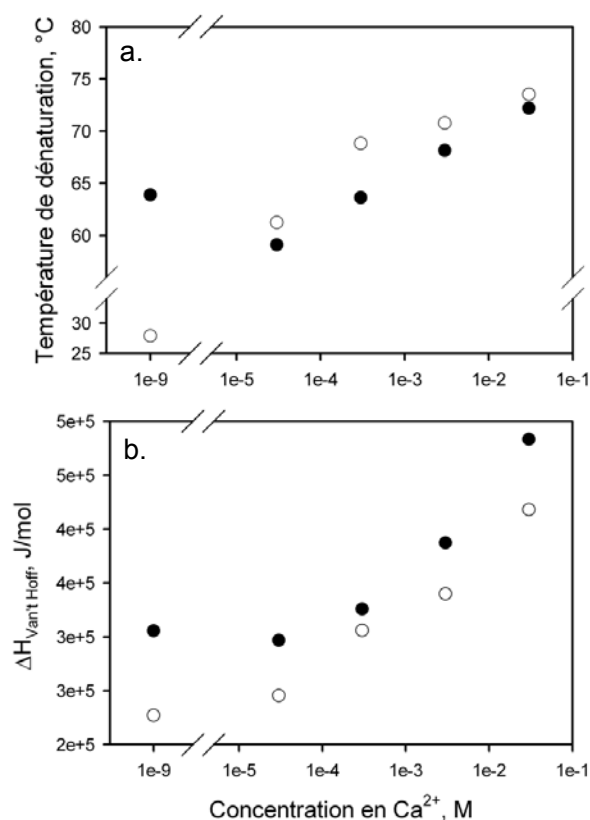
En revanche, en présence de calcium, les températures de dénaturation de la structure tertiaire et de la structure secondaire sont proches et augmentent avec la concentration en calcium.

Les courbes enregistrées dans l'UV lointain présentent une dérive de la ligne de base assez importante. Avant le début de la dénaturation de la structure secondaire, le signal augmente déjà légèrement et il continue d'augmenter après la dénaturation totale. Cette dérive de la ligne de base est estimée et prise en compte pour le paramétrage de ces courbes.

Les paramètres déterminés par cette technique sont reportés sur la figure II.15.

## Température de dénaturation et $\Delta H$ de Van't Hoff

La figure II.15 représente les températures de dénaturation et les  $\Delta H$  de Van't Hoff déterminées grâce aux paramétrages des courbes de dénaturation. Les paramètres mesurés en présence d'EDTA sont représentés sur la figure à une concentration en calcium égale à  $10^{-9}$  M de calcium.



**Fig II.15 :** Températures de dénaturation (a.) et  $\Delta H$  de Van't Hoff (b.) pour différentes concentrations de calcium.

Déterminées en paramétrant les courbes de dénaturation thermique enregistrées par dichroïsme circulaire dans l'UV proche (●) et dans l'UV lointain (○).

Les données issues des expériences effectuées dans l'UV proche sont représentées par des ronds pleins et celles issues des expériences effectuées dans l'UV lointain sont représentées par des ronds vides.

Les paramètres déterminés à partir des deux séries d'expériences sont proches et ont un comportement similaire en présence de calcium.

Les températures de dénaturation sont représentées sur la figure II.15a. Elles augmentent avec la concentration en calcium et varient entre 59 et 74°C.

Les  $\Delta H$  de Van't Hoff sont représentés sur la figure II.15b. Elles augmentent elles aussi avec la concentration en calcium et varient entre 300 kJ/mol et 550 kJ/mol.

En présence d'EDTA, les paramètres ressemblent à ceux déterminés à faible concentration de calcium, exceptée la température de dénaturation mesurée dans l'UV proche qui est beaucoup plus faible.

Lorsque la température d'une solution de protéine augmente, la chaleur apportée au cours de cette augmentation de température déstructure la protéine. Cette déstructuration a été suivie par dichroïsme circulaire dans la partie précédente. Un autre moyen de suivre la dénaturation est de mesurer directement la quantité de chaleur absorbée par la protéine pour sa déstructuration.

#### **II.4.2. Dénaturation thermique suivie par calorimétrie différentielle à balayage (differential scanning calorimetry – DSC)**

##### Principe

La DSC consiste à mesurer la quantité de chaleur fournie à un échantillon pour élever sa température. Ces mesures sont réalisées en mode différentiel : deux cellules de mesure sont utilisées. La solution de protéine à étudier est placée dans la première, la deuxième étant réservée à la solution tampon. Ces deux cellules sont en contact thermique avec un thermostat et leur température est contrôlée. La température des deux solutions est augmentée progressivement. Dans notre cas, elle augmente à la vitesse de 1°C par minute. Les puissances utilisées pour augmenter la température des échantillons sont mesurées. La puissance utilisée



par la solution tampon est soustraite à celle de la solution de protéine. Cette opération permet de distinguer les effets uniquement dus aux protéines. La courbe obtenue représente donc la puissance utilisée pour augmenter la température des protéines. Cette différence de puissance  $\Delta P$  peut être traduite en terme de variation de capacité calorifique  $\Delta C_p$  à pression constante :

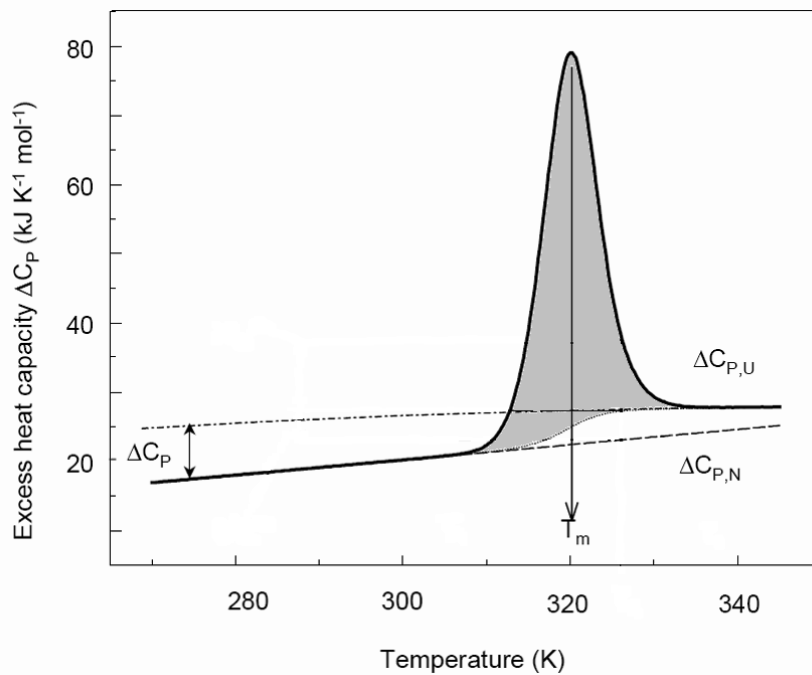
$$\Delta P = \Delta C_p \frac{dT}{dt}$$

Ainsi, un traitement mathématique adéquat, permet d'obtenir les courbes de  $\Delta C_p$  à partir des courbes de  $\Delta P$ . Ce sont ces courbes  $\Delta C_p$  qui sont utilisées pour calculer les paramètres thermodynamiques.

Dans le cas de la dénaturation d'une protéine suivie par DSC, deux phénomènes sont à l'origine de l'absorption de chaleur par l'échantillon :

- L'énergie nécessaire à l'augmentation de température de l'échantillon est puisée dans le milieu ambiant sous forme de chaleur.
- L'énergie nécessaire à la destruction de la structure de la protéine est, elle aussi, puisée dans le milieu sous forme de chaleur.

La figure II.16, adaptée de la référence 153, représente une courbe typique de dénaturation suivie par DSC. Elle permet d'illustrer le principe de lecture d'une courbe de DSC.



**Fig II.16 :** Exemple de courbe de calorimétrie différentielle à balayage.

Plusieurs zones doivent donc être distinguées dans la courbe. Une grande bande est visible aux alentours de 320 K. Cette bande centrale correspond à la chaleur absorbée par la protéine pour qu'elle se dénature. Sa position indique la température de dénaturation. L'aire de la bande, grisée sur la figure II.16, correspond au  $\Delta H$  de la réaction. En dehors de cette bande, les protéines ne subissent pas de modifications majeures et la chaleur absorbée par l'échantillon sert uniquement à élever sa température.

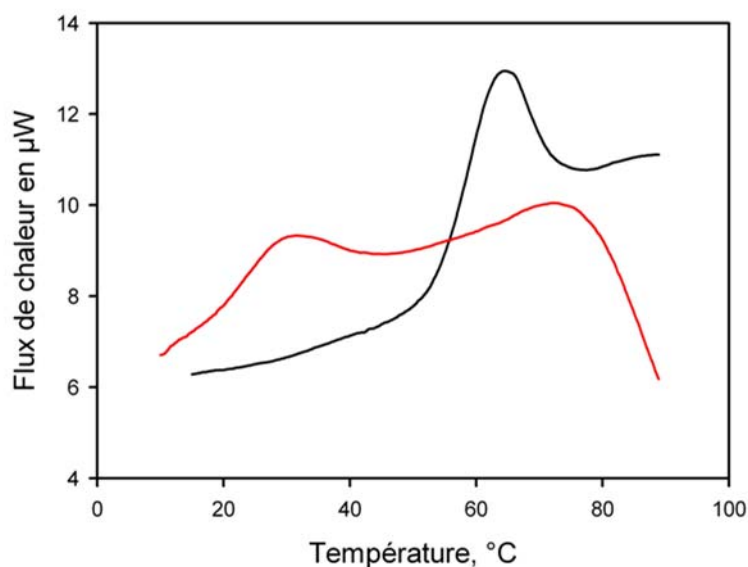
La courbe reflète alors la capacité calorifique des protéines. Avant la bande, pour  $T < T_{\text{denat}}$ , la capacité calorifique mesurée est celle de l'état natif  $\Delta C_{p,N}$ . Après la bande, pour  $T > T_{\text{denat}}$ , la capacité calorifique mesurée est celle de l'état dénaturé  $\Delta C_{p,U}$ . La différence entre ces deux capacités calorifiques  $\Delta C_p = \Delta C_{p,U} - \Delta C_{p,N}$  sera calculée.

## Conditions expérimentales

Les expériences sont réalisées dans des solutions de Tris 0,1 M à pH 7,5. La concentration de calcium varie entre 10  $\mu$ M et 100 mM. La concentration en EDTA est de 1 mM pour l'étude de la forme apo. La concentration en protéine est de 1 mg/mL (70  $\mu$ M). La température varie à la vitesse de 1°C par minute de 15°C à 90°C en présence de calcium et de 10°C à 90°C en présence d'EDTA. L'appareil utilisé pour ces mesures est un microcalorimètre Nano DSC 3.

## Résultats

Des exemples de courbes obtenues par DSC sont représentés sur la figure II.17. La courbe en noir a été enregistrée en présence de 100  $\mu$ M de calcium et la courbe en rouge en présence d'1 mM d'EDTA.



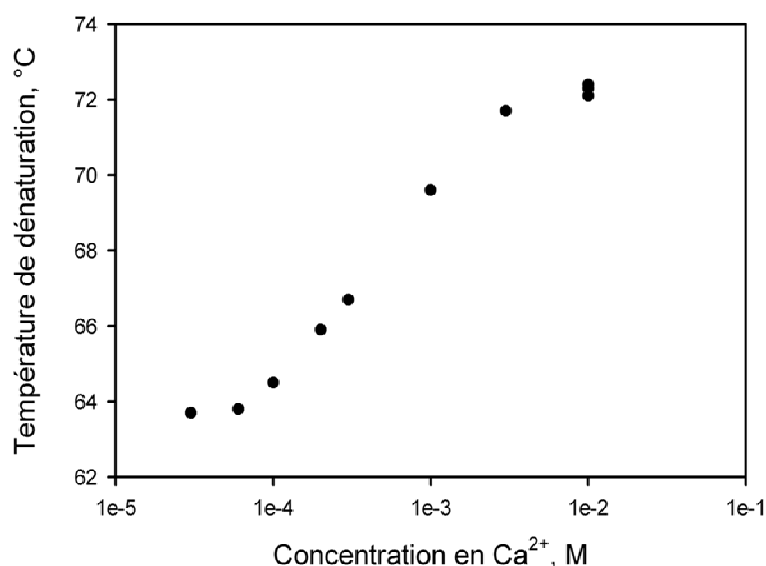
**Fig II.17 :** Courbes de dénaturation enregistrées en calorimétrie différentielle à balayage.

La courbe en noir est enregistrée en présence de 100  $\mu$ M.  
La courbe en rouge est celle enregistrée en présence d'EDTA.

La courbe enregistrée en présence de calcium correspond bien à une courbe de dénaturation thermique décrite dans l'exemple précédent (figure II.16). La courbe d'EDTA a une allure plus atypique, elle sera commentée par la suite.

Les expériences sont réalisées en présence de différentes concentrations de calcium. Pour chacune des concentrations, on mesure la température de dénaturation, le  $\Delta C_p$  entre la protéine native et la protéine dénaturée, et le  $\Delta H$  de la dénaturation.

Les températures de dénaturation déterminées en DSC sont représentées sur la figure II.18.

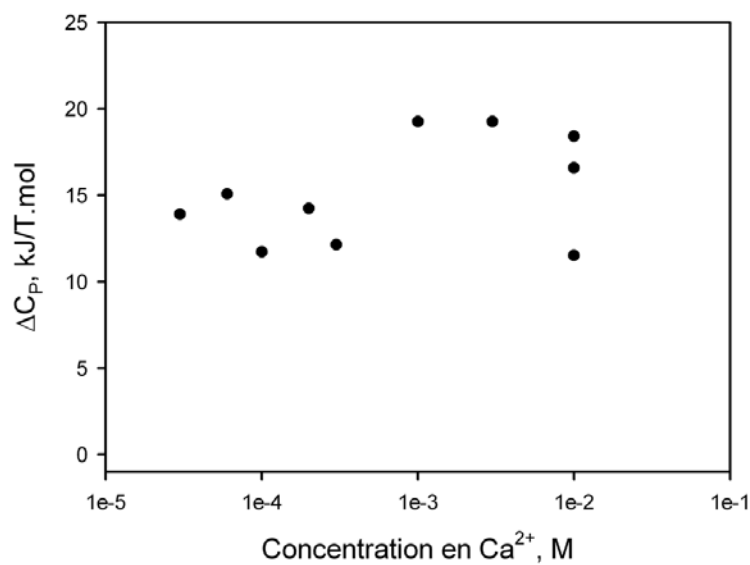


**Fig II.18 :** Température de dénaturation déterminée par calorimétrie différentielle à balayage pour différentes concentrations en calcium.

Les points représentant ces températures en fonction de la concentration en calcium forment une sigmoïde. Pour des faibles concentrations, les températures de dénaturation sont constantes et égales à environ 64°C. Elles augmentent ensuite pour des concentrations comprises entre 10  $\mu$ M et 3 mM, puis atteignent un plateau pour des concentrations supérieures. La température de dénaturation pour les hautes concentrations est d'environ

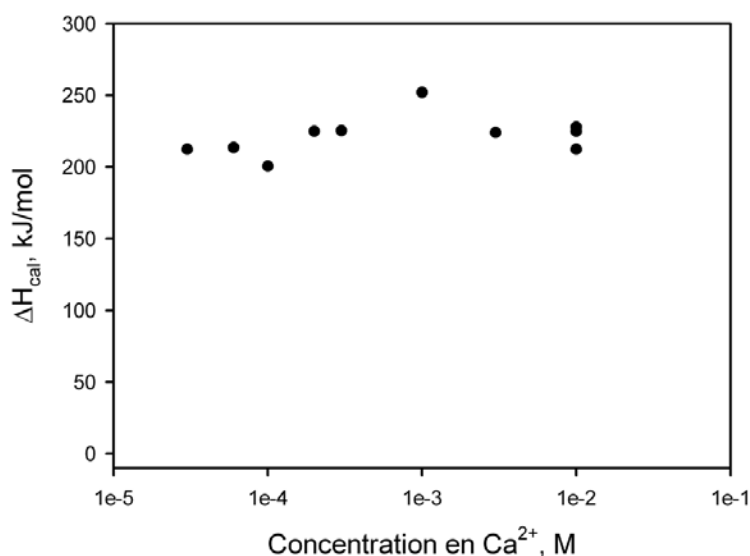
72°C. Elles sont semblables à celle déterminées en suivant la dénaturation thermique par dichroïsme circulaire (figure II.15).

Les  $\Delta C_p$  entre la protéine dépliée et la protéine sous sa forme native sont déterminées à partir de ces expériences (FigII.19). Elles sont positives et situées autour de 15 kJ/mol. Mais les valeurs obtenues sont très bruitées et ne permettent pas de souligner une évolution des  $\Delta C_p$  avec la concentration en calcium.



**Fig II.19 :**  $\Delta C_p$  mesurée par calorimétrie différentielle à balayage pour différentes concentrations en calcium.

Les  $\Delta H$  de dénaturations calculées par DSC sont représentées sur la figure II.20 en fonction de la concentration en calcium présent en solution.



**Fig II.20 :**  $\Delta H$  mesurée par calorimétrie différentielle à balayage pour différentes concentrations en calcium.

La  $\Delta H_{cal}$  calculée est moins bruitée que les  $\Delta C_p$  et semble être constante et égale à environ 220 kJ/mol.

La courbe enregistrée en présence d'EDTA (figure II.17) a une allure atypique et le traitement décrit ci-dessus n'a pas pu être effectué. Comme montrée dans les expériences de dénaturation suivie par dichroïsme circulaire, la dénaturation de l' $\alpha$ -lactalbumine en absence de calcium se déroule en deux temps. La protéine perd tout d'abord sa structure tertiaire aux alentours de 30°C. Cette destruction de la structure tertiaire correspond à la bande aux alentours de 28-30°C observée sur la courbe de DSC. Entre 30 et 60°C, la protéine se trouve dans un état *molten globule* avec la structure secondaire de l'état natif, mais pas de structure tertiaire stable. A 63,5°C, la structure secondaire est à son tour détruite. Une seconde bande est visible pour ces températures sur la courbe de DSC. Les deux bandes d'absorption sont présentes sur une large gamme de température et il est difficile de tracer des lignes de base

pour déterminer les  $\Delta C_p$  et les  $\Delta H$ . Les  $\Delta H_{\text{cal}}$  sont grossièrement estimées : elle est d'environ 53 kJ pour la bande à 30°C et 71 kJ pour celle à 63,5°C.

## Discussion

Ces températures sont du même ordre que celles déterminées grâce aux expériences de dénaturation thermique suivie par dichroïsme circulaire. Cela confirme que la dénaturation se déroule de manière identique, qu'elle soit suivie par dichroïsme circulaire ou par DSC.

Les  $\Delta C_p$  entre la protéine dénaturée et la protéine native sont très bruitées. Elles semblent légèrement positives. Ce qui indique que la capacité calorifique de la forme repliée est plus petite que celle de la forme dépliée.

Cela correspond bien aux résultats attendus : les travaux de Robertson et Murphy<sup>154</sup> suggèrent que la capacité calorifique d'une protéine augmente lorsque sa surface accessible au solvant augmente. Lorsqu'une protéine se dénature, sa structure passe d'une forme compacte à une forme dépliée et sa surface accessible au solvant augmente beaucoup.

Les  $\Delta H_{\text{cal}}$  mesurées par calorimétrie diffèrent des  $\Delta H$  de Van't Hoff déterminées en paramétrant les courbes de dénaturation thermique enregistrées par dichroïsme circulaire. Les  $\Delta H$  de Van't Hoff sont calculées en étudiant des équilibres entre l'état replié et déplié. Les  $\Delta H_{\text{cal}}$  sont calculées en mesurant la quantité d'énergie nécessaire à la protéine pour se déplier.

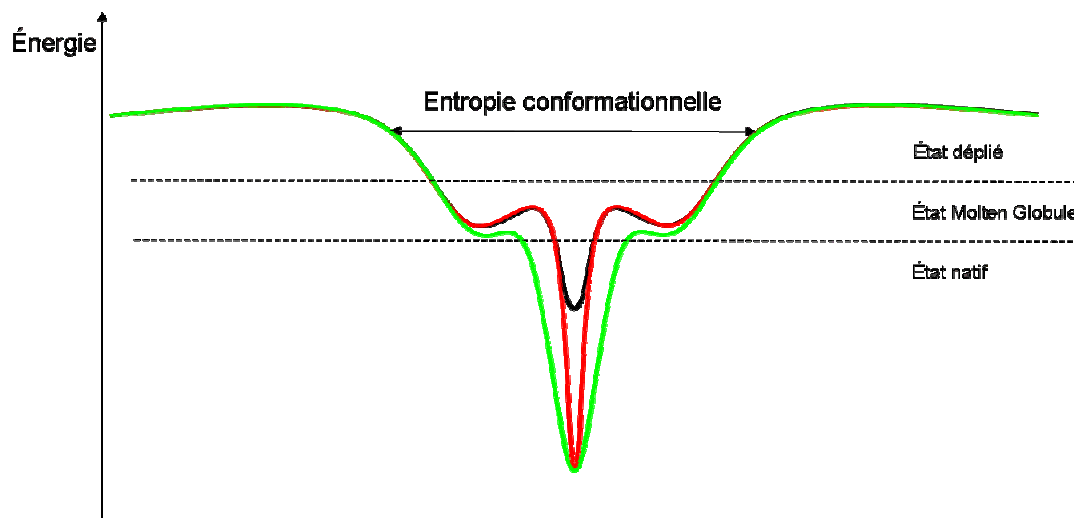
Ces  $\Delta H$  sont égaux lorsqu'il y a échange entre deux états<sup>153</sup>. Lorsque les expériences sont réalisées en présence de calcium, le dépliement est coopératif et la protéine passe directement de l'état replié à l'état déplié.

Les  $\Delta H_{\text{cal}}$  et de Van't Hoff devraient être égaux. Ce n'est cependant pas le cas. Aucune explication n'a été trouvée jusqu'à présent. Ces différences peuvent provenir du fait que l'état

*molten globule* est peuplé transitoirement au cours du dépliement. Dans ce cas, la réaction impliquerait trois états et l'égalité entre les deux variations d'enthalpie ne serait pas respectée.

### II.4.3 Discussion

Les informations déduites des expériences de dénaturation thermiques couplées avec celles obtenues en étudiant les cinétiques de repliement permettent de décrire plus précisément l'influence du calcium sur le paysage d'énergie de la protéine. L'évolution de ce paysage d'énergie est schématisée sur la figure II.21.



**Fig II.21 :** Paysage d'énergie simplifié pour différentes concentrations de calcium.

En noir : Paysage d'énergie en absence de calcium

En rouge : Paysage d'énergie en présence de 30 $\mu$ M de calcium

En vert : Paysage d'énergie en présence de 30mM de calcium

L'augmentation de température du système correspond à l'augmentation d'énergie mise à disposition de la protéine sous forme de chaleur pour se promener dans son paysage d'énergie. Les états accessibles par la protéine dépendront donc de l'énergie dont elle dispose.



Au cours d'une expérience de dénaturation thermique, la solution est supposée en équilibre thermodynamique à chaque instant. Cet état d'équilibre thermodynamique est atteint par compromis entre la minimisation de l'énergie interne à la protéine et la maximisation de l'entropie. C'est grâce à ce compromis que l'on pourra expliquer l'évolution des températures de dénaturation.

#### En absence de calcium :

Le paysage d'énergie en absence de calcium est représenté en noir sur la figure II.21.

A basse température, la protéine est dans l'état natif. Lorsque la température atteint 30°C, la protéine a suffisamment d'énergie à sa disposition pour sortir du puits de l'état natif et peupler l'état *molten globule*. Cet état a une énergie plus haute que l'état natif, mais il possède une plus grande entropie conformationnelle. Ainsi, l'énergie interne du système va augmenter, mais cette augmentation sera compensée par une augmentation de l'entropie. Entre 30 et 65°C, les protéines n'ont pas suffisamment d'énergie pour se déplier complètement, elles restent donc dans l'état *molten globule*.

Lorsque la température atteint 65°C, les protéines disposent de suffisamment d'énergie pour détruire leur structure secondaire. Lorsque la structure secondaire est détruite, il y a beaucoup moins d'interactions au sein de la protéine, beaucoup moins de contraintes et la chaîne polypeptidique peut peupler un très grand nombre de micro-états dépliés. L'état déplié correspond à l'ensemble de tous ces micro-états, ce qui explique sa très grande entropie conformationnelle. Ainsi, dès que la protéine a suffisamment d'énergie pour peupler l'état déplié, elle se dénature complètement. L'augmentation d'énergie, due au fait que les protéines dépliées possèdent une grande énergie interne, est compensée par l'augmentation d'entropie.

### En présence de calcium :

Lorsque du calcium est introduit en solution, il se fixe sur l'état natif et le stabilise. Le dépliement devient coopératif, les structures secondaires et tertiaires sont détruites en même temps et l'état *molten globule* n'est plus peuplé au cours de la dénaturation. Contrairement à ce qui se passe en absence de calcium, l'énergie dont la protéine dispose aux températures intermédiaires (entre 30 et 60°C) ne suffit plus pour détruire la structure tertiaire. L'énergie nécessaire pour détruire la structure tertiaire est élevée et ne peut plus être compensée par la grande entropie de l'état molten globule.

La structure tertiaire reste donc présente jusqu'à environ 65°C, température à laquelle la structure secondaire est détruite. Pour les températures supérieures à 65°C, la protéine dispose de suffisamment d'énergie pour peupler l'état déplié. Cet état est de haute énergie, mais il possède une très grande entropie conformationnelle. L'énergie qu'il faut dépenser pour peupler cet état est compensée par sa grande entropie. Dès que la protéine aura suffisamment d'énergie à sa disposition, elle se dépliera.

Lorsque la concentration en calcium augmente, les expériences de cinétique présentées au début de cette partie montrent que cet ion ne se fixe plus uniquement sur l'état natif, mais également sur les structures proches comme l'état *molten globule* et l'état de transition. Ces états sont stabilisés par la fixation du calcium, et le paysage d'énergie est modifié comme illustré dans la figure II.21. Cette modification se traduit par une diminution d'entropie de l'état *molten globule*. Il en résulte une accélération du repliement de la protéine comme expliquée dans le début de cette partie. Cette modification se traduit aussi par une augmentation de l'entropie de l'état natif. Le puits qui délimite l'état natif s'évase, sa courbure augmente. Cet état est donc stabilisé lorsque la concentration augmente. Cet effet est

visible sur les courbes de dénaturation : lorsque la concentration augmente, la température de dénaturation augmente elle aussi.

Au-delà du millimolaire de calcium, tous les états de la protéine susceptibles d'interagir avec le calcium ont déjà lié un ion, et l'augmentation de la concentration ne modifiera plus le paysage d'énergie et la dénaturation de la protéine.

# **III. Repliement de l' $\alpha$ - lactalbumine en absence de métaux**

Dans la partie précédente, le repliement de l' $\alpha$ -lactalbumine a été suivi en présence de métaux. Ces études ont montré que la fixation d'un ion métallique sur l'état *molten globule* diminue son entropie conformationnelle. Dans cette seconde partie, la réaction de repliement est effectuée en absence de métaux. Les expériences réalisées permettront d'observer les effets induits par la grande entropie de l'état *molten globule*.

Plusieurs méthodes spectroscopiques sont utilisées pour suivre la réaction. Le dichroïsme circulaire permet de suivre le repliement pour différentes concentrations en protéine. La spectroscopie infra rouge et la RMN nécessitent, elles, des concentrations de l'ordre de 20 mg/mL. Une large gamme de concentration en protéines est donc utilisée pour l'étude de la forme apo.

La concentration en protéine d'une solution a un effet sur sa viscosité. Or, les vitesses des réactions chimiques, et en particulier les vitesses de repliement des protéines, dépendent de la viscosité de la solution. Elle sera étudiée en effectuant des expériences de rhéologie avant de s'intéresser à la réaction de repliement.

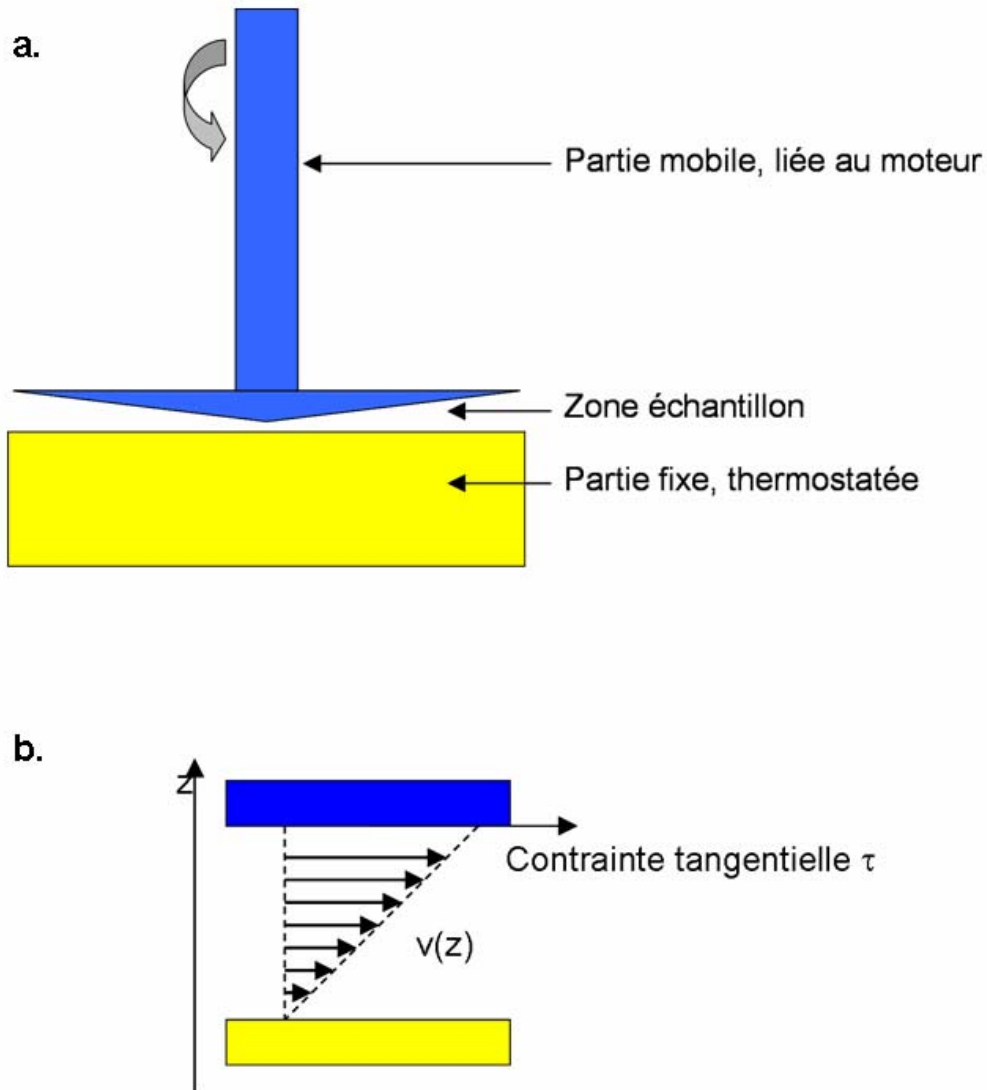
## **III.1. Rhéologie**

Des mesures de viscosité à l'équilibre sont effectuées sur la protéine dans l'état *molten globule* et dans l'état natif pour différentes concentrations en protéine. La viscosité est ensuite mesurée au cours du repliement. Cette approche permet de mettre en évidence une évolution de la viscosité pendant que la protéine se replie.

### **III.1.1. Description de l'expérience**

#### **Principe de fonctionnement du rhéomètre**

Les viscosités des solutions sont mesurées à l'aide d'un rhéomètre. La figure III.1 représente un schéma simplifié de la cellule de mesure d'un rhéomètre, et permet d'expliquer le fonctionnement de cet appareil.



**Fig III.1 : Schémas de principe du rhéomètre**

Sur cette figure, deux parties sont distinguables. Celle en jaune est la partie fixe. C'est sur cette partie que l'échantillon est déposé. Elle est plane et thermostatée dans nos expériences. La partie en bleu est la partie mobile. Elle est liée à un moteur qui possède une grande souplesse d'utilisation. Il peut être utilisé à vitesse ou couple constant : la partie mobile peut être entraînée en rotation ou effectuer des oscillations autour d'une position d'équilibre. Cette

partie mobile peut avoir plusieurs géométries. Dans la configuration utilisée ici, la géométrie de la partie mobile est conique.

Pour mesurer la viscosité d'une solution, deux millilitres sont déposés dans l'entrefer entre la partie fixe et la partie mobile. La partie mobile est entraînée en rotation par le moteur. Le fluide est alors soumis à une contrainte tangentielle  $\tau$  à l'interface avec la partie mobile, et un écoulement laminaire s'établit (figure III.1.b). Le gradient de vitesse dans l'écoulement laminaire correspond au taux de cisaillement. Pour mesurer la viscosité de la solution, le taux de cisaillement est choisi constant. Dans ce type d'écoulement, la viscosité est égale à la contrainte tangentielle divisée par le taux de cisaillement. La viscosité peut donc être déterminée car le taux de cisaillement est fixe et la contrainte tangentielle accessible à partir de la géométrie de la zone échantillon et du couple développé par le moteur. C'est la viscosité dynamique, elle s'exprime en Pascal seconde (Pa.s).

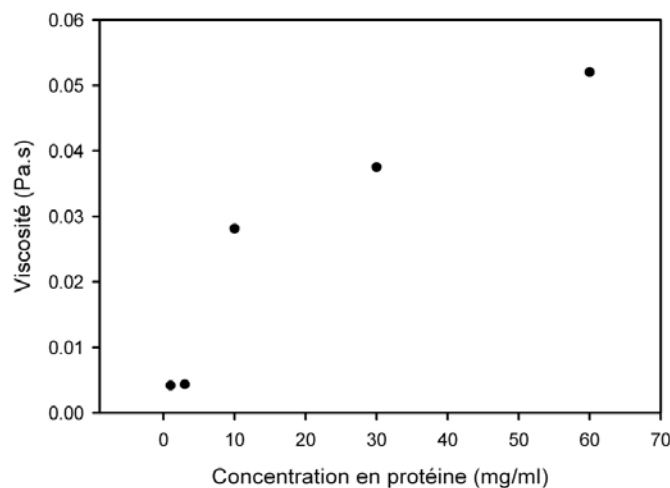
### Conditions expérimentales

Les expériences de rhéologie sont effectuées sur un rhéomètre AR-G2 (TA Instruments). Les expériences à l'équilibre sont réalisées à 5°C, les viscosités sont mesurées en utilisant des taux de cisaillement compris entre 0,1 et 3000 s<sup>-1</sup> sur des solutions de protéine dont la concentration varie entre 1 et 60 mg/mL. Les solutions de protéine dans l'état *molten globule* sont obtenues à pH acide (pH 2), les solutions de protéine dans l'état natif sont à pH 7,5. Toutes les expériences à l'équilibre ont été réalisées dans l'eau et le D<sub>2</sub>O car les expériences de repliement seront réalisées dans ces deux solvants. Les résultats étant comparables dans les deux cas, seules les expériences réalisées dans l'eau sont présentées dans la suite de cette partie.



Pour les cinétiques, le repliement est déclenché par saut de pH. Pour ce faire 10 volumes de solution de protéines à pH 2 sont mélangés avec 1 volume d'une solution de Tris 1 M, EDTA 10 mM et du HCl pour obtenir une solution finale à pH 7,5. Cette solution est rapidement déposée dans la zone échantillon thermostatée à 5°C. Le taux de cisaillement utilisé ici est de  $30 \text{ s}^{-1}$ . C'est un taux suffisamment fort pour mesurer avec précision la viscosité dynamique de la solution sans toutefois trop perturber le système.

### III.1.2. Mesures à l'équilibre



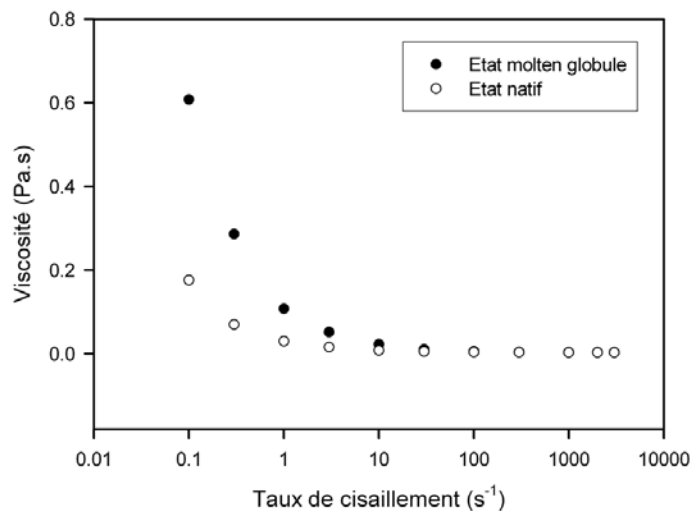
**Fig III.2 : Évolution de la viscosité d'une solution d' $\alpha$ -lactalbumine dans l'état molten globule avec la concentration en protéine**

Ces viscosités sont mesurées avec un taux de cisaillement égal à  $3 \text{ s}^{-1}$ .

La figure III.2 représente les viscosités mesurées sur les solutions de protéine dans l'état *molten globule* à  $3 \text{ s}^{-1}$ , les taux de cisaillement inférieurs à  $3 \text{ s}^{-1}$  ne permettent pas de mesurer avec précision les viscosités des solutions à faible concentration en protéine. Cette figure montre clairement que la viscosité de la solution augmente avec la concentration en protéine. Elle est inférieure à  $0,005 \text{ Pa.s}$  pour les concentrations inférieures à  $3 \text{ mg/mL}$ , et est

supérieure à 0,05 Pa.s pour une concentration égale à 60 mg/mL. En particulier, on remarque une augmentation brusque de la viscosité qui passe de 0,005 à 0,028 Pa.s, lorsque la concentration passe de 3 à 10 mg/mL.

La figure III.3 présente les viscosités mesurées pour différents taux de cisaillements sur des solutions de protéine à 60 mg/mL dans l'état natif et *molten globule*. Pour cette concentration en protéine, la viscosité peut être mesurée pour des taux de cisaillement compris entre 0,1 et 3000 s<sup>-1</sup>.



**Fig III.3 :** Viscosité des solutions de protéine à 60 mg/ml dans l'état molten globule et l'état natif pour différents taux de cisaillement

La viscosité mesurée sur les solutions diminue rapidement avec le taux de cisaillement pour atteindre un plateau aux alentours de 100 s<sup>-1</sup>.

Pour un faible taux de cisaillement, la viscosité de la solution de protéine est égale à 0,6 Pa.s dans l'état *molten globule*, et elle est de 0,18 Pa.s pour la solution de protéine dans l'état natif.

Pour des taux de cisaillement supérieur à 300 s<sup>-1</sup>, la viscosité des deux solutions est proche de 0,003 Pa.s.

Ce comportement est caractéristique d'un milieu rhéofluidifiant : plus la contrainte imposée à l'échantillon augmente, plus cet échantillon est fluide.

Un autre effet est observé sur la figure III.3. Les solutions de protéine dans l'état natif (ronds vides) sont moins visqueuses que les solutions de protéine dans l'état *molten globule* (ronds pleins). Cet effet est particulièrement visible pour les faibles taux de cisaillement.

La viscosité d'une solution de protéine dépend donc non seulement de la concentration en protéine et du taux de cisaillement auquel elle est soumise, mais aussi de l'état de la protéine en solution. Ce constat implique que la viscosité d'une solution de protéine varie lorsque la protéine se déplie.

Cependant, cette hypothèse n'est pas directement vérifiable à partir des expériences présentées précédemment. En effet, ces expériences sont réalisées à l'équilibre et les conditions expérimentales favorables à un *molten globule* sont différentes des conditions expérimentales favorables à l'état natif, plus particulièrement la force ionique. Le *molten globule* est obtenu à pH 2 alors que l'état natif est obtenu à pH 7 (100 mM Tris).

Pour évaluer plus rigoureusement l'évolution de la viscosité de la solution, la viscosité doit directement être mesurée au cours de la réaction de repliement.

### **III.1.3. Mesures en cours de cinétique**

La viscosité est mesurée pendant que la protéine se replie.

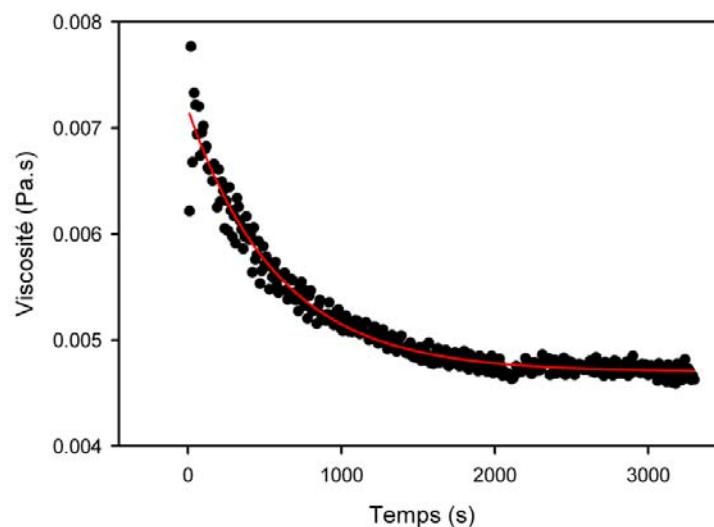
Pour ces expériences, la concentration en protéine est égale à 60 mg/mL. Cette solution à haute concentration permet de mesurer avec précision la viscosité et l'évolution de la viscosité.

Un taux de cisaillement de  $30 \text{ s}^{-1}$  est utilisé ici. Ce taux de cisaillement est suffisamment faible pour ne pas perturber le système, mais il est assez fort pour déterminer rapidement et avec précision la viscosité de la solution.

La figure III.4 montre l'évolution de la viscosité au cours du repliement.

La viscosité diminue de manière exponentielle lorsque la protéine se replie. En début de réaction, elle est comprise entre 0,007 et 0,008 Pa.s. En fin de réaction, elle est comprise entre 0,0045 et 0,005 Pa.s.

La ligne en rouge sur la figure correspond au paramétrage de la courbe par une exponentielle décroissante. La constante de temps associée à cette exponentielle est égale à  $1,7 \cdot 10^{-3} \text{ s}^{-1}$  et correspond à la constante de vitesse de repliement de la protéine sous la forme apo (voir plus loin).



**Fig III.4 : Évolution de la viscosité d'une solution de protéine à 60 mg/ml au cours du repliement.**

Points noirs : mesures de viscosité réalisées au cours de repliement par rhéométrie.  
Ligne rouge : paramétrage de la courbe par une mono-exponentielle décroissante

### III.1.4. Interprétation et discussion

Pour expliquer l'évolution de la viscosité en fonction de la concentration en protéine, il faut utiliser l'équation suivante qui permet de décrire la viscosité d'une suspension de particules<sup>155</sup>:

$$\eta = \eta_0 (1 + \nu_1 \cdot \Phi + \nu_2 \cdot \Phi^2 + \dots)$$

avec  $\eta$  la viscosité de la solution,  $\eta_0$  la viscosité du solvant,  $\Phi$  la fraction volumique occupée par les particules en solution,  $\nu_1$  un facteur dépendant de la forme des particules appelé facteur de Simha (égal à 2,5 pour une particule sphérique) et  $\nu_2$  un facteur dépendant de l'interaction entre les particules.

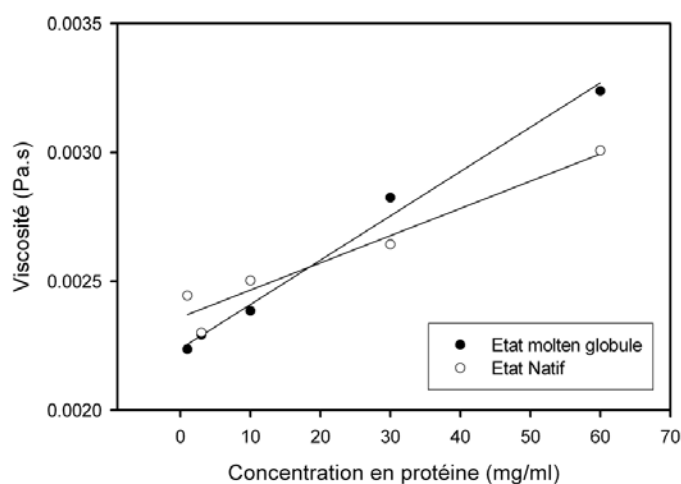
La fraction volumique  $\Phi$  occupée par les protéines en solution est calculée à partir de la concentration et du volume occupé par chaque protéine.

Les résultats expérimentaux indiquent que l'augmentation de la viscosité en fonction de la concentration en protéine n'est pas linéaire.

L'équation précédente ne peut donc pas être réduite à  $\eta = \eta_0 (1 + \nu_1 \cdot \Phi + \dots)$ , le terme  $\nu_2 \cdot \Phi^2$  ne peut pas être négligé. Ce terme non négligeable dépend de l'interaction entre les particules. L'augmentation de viscosité n'est pas due uniquement au nombre de particules en solution, mais aussi aux interactions entre les particules. Lorsque la concentration en protéine augmente, la distance moyenne entre les protéines diminue et les interactions entre ces protéines sont facilitées.

La prise en compte des interactions offre en outre l'avantage d'expliquer le caractère rhéofluidifiant des solutions. Lorsque le taux de cisaillement augmente, la contrainte imposée sur le fluide détruit ces interactions entre protéines et la solution devient plus fluide.

Pour un taux de cisaillement égal à  $3000 \text{ s}^{-1}$ , toutes les interactions entre les protéines sont détruites et l'équation liant la viscosité à la fraction volumique devient linéaire. La figure III.5 représente la viscosité de la solution mesurée à fort taux de cisaillement pour différentes concentrations en protéine.



**Fig III.5 : Viscosités mesurées à fort taux de cisaillement.**

Les viscosités sont mesurées sur des solutions d' $\alpha$ -lactalbumine dans l'état molten globule et l'état natif pour différentes concentrations. Le taux de cisaillement utilisé pour ces mesures est de  $3000 \text{ s}^{-1}$ .

Cercles pleins : Protéines dans l'état molten globule.

Cercles vides : Protéines dans l'état natif.

Droites : régressions linéaires effectuées sur les données expérimentales et utilisées pour calculer les facteurs de Simha.

La viscosité évolue de manière linéaire avec la concentration en protéine, et donc avec la fraction volumique occupée par la protéine.

L'équation liant la viscosité à la fraction volumique  $\eta = \eta_0 (1 + \nu_1 \cdot \Phi + \nu_2 \cdot \Phi^2 + \dots)$  peut être réduite à ses deux premiers termes  $\eta = \eta_0 (1 + \nu_1 \cdot \Phi + \dots)$ .

Les droites sur la figure III.5 représentent les régressions linéaires effectuées sur les données. La fraction volumique occupée par les protéines évolue de manière linéaire avec la concentration en protéine. Le coefficient de linéarité est déterminé en calculant le volume spécifique des protéines dans l'état natif et *molten globule* à partir de leur rayon hydrodynamique. Les pentes des droites de la figure III.5 permettent donc de calculer les facteurs de Simha  $v_1$ .

Le facteur de Simha renseigne sur la forme des particules en suspension. Il est minimum pour des particules sphériques. Dans ce cas, il est égal à 2.5. Les facteurs plus élevés correspondent à des formes ellipsoïdales.

Le facteur de Simha est égal à 4,95 pour l'état *molten globule* et à 3,67 pour l'état natif. Si il n'y a réellement plus d'interactions entre les protéines, elles sont toutes sous forme monomérique et les facteurs indiquent qu'elles ne sont pas tout à fait sphériques. De plus, les protéines dans l'état *molten globule* sont « moins sphériques » que celles dans l'état natif. Ce constat est en accord avec le fait que le *molten globule* ne possède pas de structure tertiaire stable, et qu'il est donc moins compact, un peu plus déplié que l'état natif. Cependant on ne peut pas exclure la possibilité qu'il subsiste encore des interactions et que les protéines soient sous forme multimérique.

Le caractère rhéofluidifiant des solutions de protéine dans l'état *molten globule* est plus important que celui observé dans les solutions de protéine à l'état natif, cela est bien visible sur la figure III.3.

La différence ne peut pas être expliquée uniquement par la différence de taille entre l'état *molten globule* et l'état natif. Elle reflète aussi une différence dans les interactions entre les protéines. Les protéines dans l'état *molten globule* interagissent plus entre elles et influencent plus la viscosité que les protéines repliées.

Sur le plan structural, ces deux états possèdent la même structure secondaire mais le *molten globule* ne possède pas de structure tertiaire stable. Les parties hydrophobes peuvent donc être accessibles au solvant et aux autres protéines, permettant ainsi la création d'interactions hydrophobes.

Les solutions sur lesquelles sont enregistrées les cinétiques de repliement sont peu ou pas agitées. Les viscosités à prendre en compte pour l'interprétation des résultats sont donc celles enregistrées à faible taux de cisaillement. Dans ces conditions, la viscosité est élevée, les interactions entre protéines sont importantes. Ces interactions permettront par la suite d'expliquer l'évolution des constantes de vitesse de repliement en fonction de la concentration en protéine. Ces constantes de vitesse de repliement sont déduites d'expériences de dichroïsme circulaire.



## **III.2. Dichroïsme circulaire**

La réaction de repliement de l' $\alpha$ -lactalbumine est étudiée par dichroïsme circulaire, RMN et spectroscopie infrarouge.

Le dichroïsme circulaire permet de détecter la présence de structures secondaire et tertiaire en solution ; cette méthode est à très basse résolution et ne donne pas d'informations précises sur la structure des protéines. Elle permet en revanche d'enregistrer les cinétiques de repliement pour une large gamme de concentration en protéine.

La technique de RMN du proton est utilisée ensuite pour détecter la présence de structure tertiaire en différents endroits de la protéine.

Puis la spectroscopie infrarouge permet de déterminer avec précision la quantité de chaque élément de structure secondaire (hélices, feuillets...). Pour cette dernière technique, les expériences doivent être réalisées dans de l'eau lourde ( $D_2O$ ). Ainsi, pour pouvoir comparer les résultats obtenus par chacune des méthodes, toutes les expériences sont réalisées dans le  $D_2O$ .

Pour le dichroïsme circulaire, les expériences peuvent être réalisées à la fois dans des tampons hydrogénés ou deutérés : les deux solvants ( $H_2O$  et  $D_2O$ ) sont utilisés pour pouvoir souligner les différences et les similitudes dans les réactions.

### III.2.1. Conditions expérimentales – Résultats

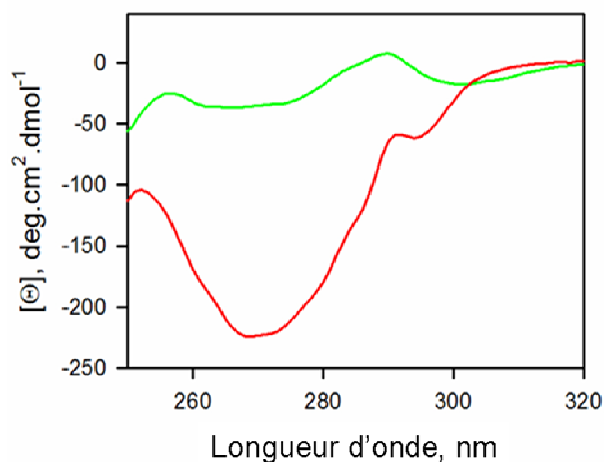
#### Principe

Le dichroïsme circulaire permet de détecter la présence de structure tertiaire dans une solution de protéine.

La figure III.6 représente des spectres de dichroïsme circulaire dans l'UV proche pour l' $\alpha$ -lactalbumine dans l'état *molten globule* et dans l'état natif.

Les protéines dans l'état *molten globule* ne possèdent pas de structure tertiaire, le signal de dichroïsme circulaire est quasi-nul.

Lorsque les protéines sont repliées, le signal de dichroïsme circulaire est négatif. Au cours du repliement, le signal de dichroïsme circulaire va donc décroître.



**Fig III.6 :** Spectres de dichroïsme circulaire dans l'UV

La courbe en rouge représente le spectre de l'état natif, la courbe en vert, celui de l'état molten globule.

Les cinétiques sont enregistrées à différentes longueurs d'onde pour différentes concentrations. L'amplitude de la variation du signal de dichroïsme circulaire varie donc d'une expérience à l'autre. Pour permettre de comparer plus facilement les différentes réactions, les signaux enregistrés sont normalisés. Les courbes obtenues représentent alors la proportion de protéines non structurées.

### Conditions expérimentales

Le repliement est déclenché par un saut de pH en mélangeant 10 volumes de solution de protéine à pH 2 (protéines dans l'état *molten globule*) avec un volume de tampon Tris à 1 M ajusté pour obtenir une solution à pH 7,5 après mélange (pH favorable à la formation de l'état natif). Le mélange est effectué à la main, d'où un temps mort d'environ 20 s pour ces expériences. La réaction est suivie à 5°C.

Les spectres sont enregistrés sur un spectromètre Jasco J-810. Le signal de dichroïsme circulaire est enregistré à 280 nm pour des concentrations en protéine inférieures à 1 mg/mL, à 295 nm pour des concentrations comprises entre 1 et 5 mg/mL, à 300 nm pour des concentrations comprises entre 5 et 20 mg/mL et à 305 nm pour des concentrations supérieures à 20 mg/mL.

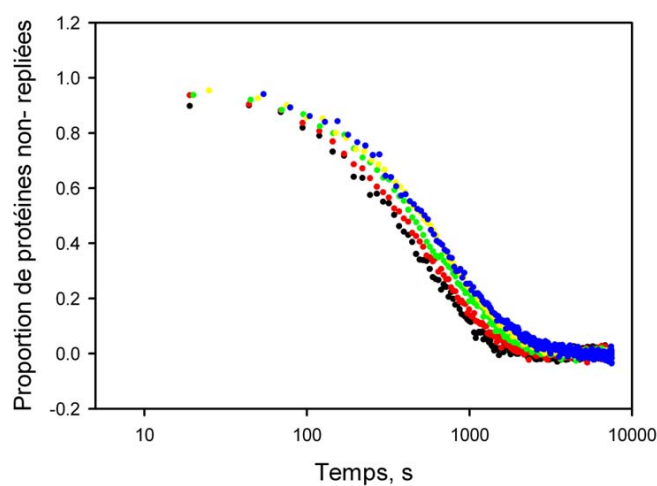
Le choix des longueurs d'onde est un compromis entre deux critères : le signal doit être suffisamment fort pour avoir un bon rapport signal sur bruit mais il ne doit pas être saturé en fin de cinétique lorsque la protéine est repliée.

Une mesure du signal de dichroïsme circulaire est enregistrée toutes les 5 s. Le repliement est suivi pendant 5 h pour des concentrations supérieures à 2 mg/mL et pendant 2 h pour des concentrations inférieures à 2 mg/mL.

## Résultats

### *Dans l'eau*

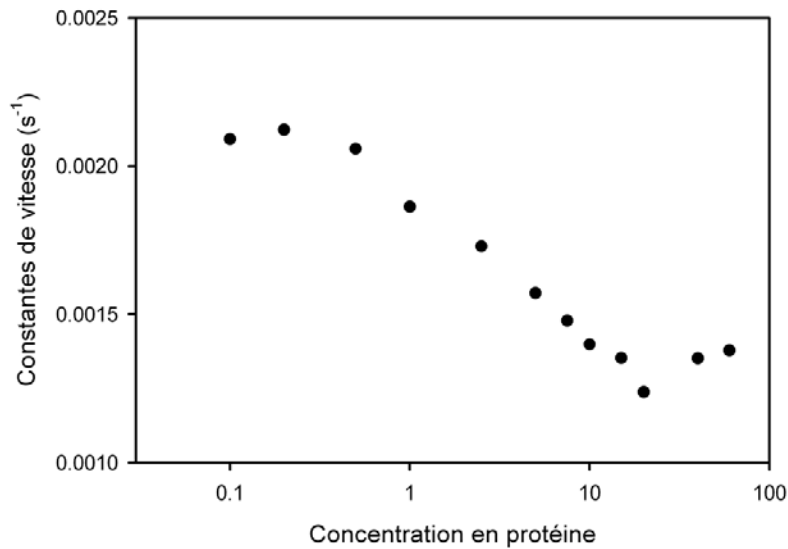
La figure III.7 représente les cinétiques enregistrées dans l'eau pour différentes concentrations en protéine.



**Fig III.7 :** Cinétiques de repliement dans l'eau enregistrées par dichroïsme circulaire

Différentes concentrations de protéines sont utilisées: (●) 0,2 mg/mL, (●) 1 mg/mL, (●) 5 mg/mL, (●) 10 mg/mL, (●) 40 mg/mL.

Quelle que soit la concentration, les cinétiques sont toutes mono-exponentielles. En revanche, la constante de vitesse de repliement varie avec la concentration en protéine. Les courbes représentées dans la figure III.7 sont paramétrées par des exponentielles décroissantes. Les constantes de temps associées à ces exponentielles donnent les constantes de vitesse de repliement de la protéine.



**Fig III.8 :** Constante de vitesse de repliement dans l'eau pour différentes concentrations en protéine.

La

figure III.8 représente ces constantes de vitesse de repliement pour différentes concentrations en protéine.

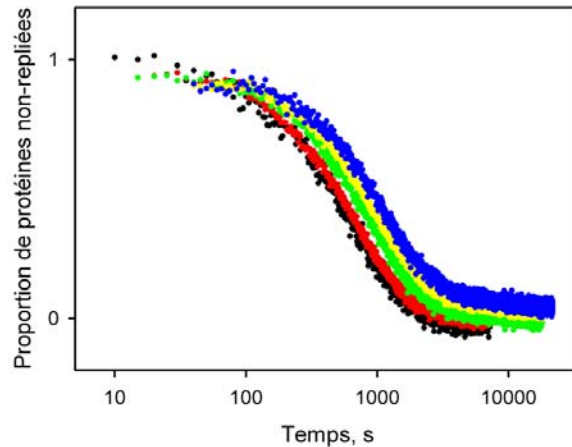
Pour les basses concentrations, en dessous de 1 mg/mL, la constante de vitesse de repliement est constante et égale à  $0,0021 \text{ s}^{-1}$ .

Lorsque la concentration augmente, la réaction est ralentie. La constante de vitesse varie de  $0,0021 \text{ s}^{-1}$  pour une concentration de 0,8 mg/mL à environ  $0,0014 \text{ s}^{-1}$  pour une concentration de 10 mg/mL.

Pour les concentrations supérieures (de 10 à 60 mg/mL), la constante de vitesse de repliement ne varie plus et est égale à environ  $0,0014 \text{ s}^{-1}$ .

Dans le D<sub>2</sub>O

La figure III.9 représente les cinétiques enregistrées dans le D<sub>2</sub>O pour différentes concentrations en protéine.



**Fig III.9 :** Cinétiques de repliement dans le D<sub>2</sub>O enregistrées par dichroïsme circulaire

Différentes concentrations de protéines sont utilisées: (●) 1 mg/mL, (●) 2,5 mg/mL, (●) 5 mg/mL, (●) 10 mg/mL, (●) 40 mg/mL.

Ces cinétiques présentent des différences importantes avec celles enregistrées dans du tampon hydrogéné. Lorsque la concentration en protéine est faible, elles sont bien mono-exponentielles. En revanche, lorsque la concentration augmente, elles deviennent bi-exponentielles, une phase lente apparaît.

Pour déterminer les constantes de vitesse associées à la réaction, il faut donc utiliser une somme de deux exponentielles :

$$I(t) = I_r \cdot \exp(-k_r \cdot t) + I_l \cdot \exp(-k_l \cdot t)$$

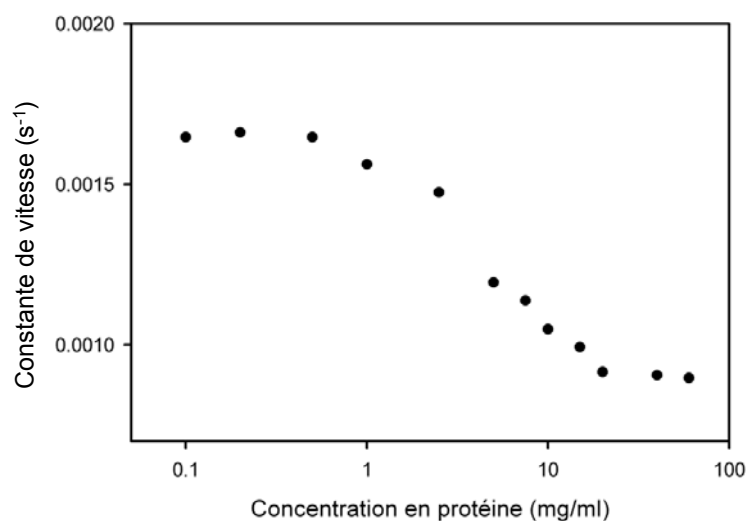
$k_r$  et  $k_l$  sont les constantes de vitesse associées à la phase rapide et à la phase lente.  $I_r$  et  $I_l$  correspondent aux amplitudes relatives de chacune des phases.

Comme les courbes sont normalisées,  $I_r + I_l = 1$ .

La phase rapide est celle qui représente la plus grande partie du changement. La phase lente n'apparaît que lorsque la concentration est élevée.

La figure III.10 représente l'évolution des constantes de vitesse de repliement associées à la phase rapide en fonction de la concentration en protéine.

Ces constantes de vitesse sont légèrement inférieures à celle enregistrées dans l'eau, mais elles évoluent de manière similaire avec la concentration. Pour les basses concentrations (jusqu'à 1 mg/mL), elles sont constantes et égales à  $1,65 \cdot 10^{-3} \text{ s}^{-1}$ . Entre 1 et 20 mg/mL, la constante de vitesse de repliement diminue. Pour des concentrations supérieures, elle redevient constante. Pour les hautes concentrations, la constante de vitesse associée à la phase rapide est égale à  $9 \cdot 10^{-4} \text{ s}^{-1}$ .



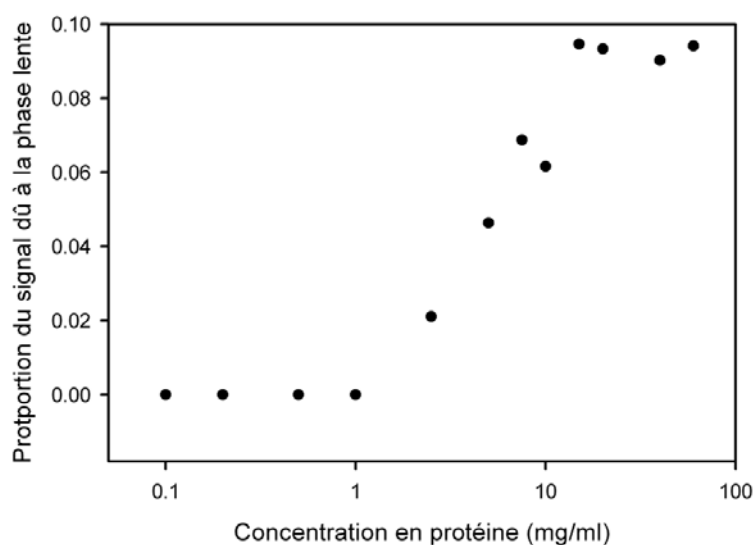
**Fig III.10 :** Constante de vitesse de la phase rapide du repliement dans le D<sub>2</sub>O pour différentes concentrations en protéine.

La

figure III.11 représente la proportion de la phase lente dans la cinétique en fonction de la concentration :  $I_1=f(c)$ .

Pour les faibles concentrations en protéine, les cinétiques sont mono-exponentielles, la phase lente n'est pas détectée ( $I_1 = 0$ ). A partir de 2,5 mg/mL, la seconde phase apparaît. Sa proportion augmente progressivement avec la concentration, pour atteindre un plateau à 15 mg/mL. Pour les hautes concentrations en protéine, l'amplitude relative de la phase lente représente près de 10% du signal.





**Fig III.11 :** Proportions de la phase lente de cinétique pour différentes concentrations en protéine.

Les constantes de vitesse associées à la phase lente sont difficilement déterminables car cette phase ne représente qu'une petite partie du signal total. Pour les hautes concentrations (au-delà de 15 mg/mL), la phase lente représente 10% du signal total ; la constante de vitesse de la phase lente reste très bruitée, mais elle semble constante à une valeur comprise entre 4 et 6  $10^{-5} \text{ s}^{-1}$ .

### III.2.2 Interprétation – Construction d'un modèle de réaction

Différences entre les réactions en D<sub>2</sub>O et H<sub>2</sub>O

Le repliement de l' $\alpha$ -lactalbumine est différent selon qu'il se passe dans le  $H_2O$  ou dans le  $D_2O$ . Dans le  $D_2O$ , le repliement est globalement plus lent et une seconde phase est détectée pour les hautes concentrations, alors que le repliement dans l'eau reste mono-exponentiel.

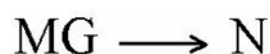
Dans l'« eau lourde », les atomes d'hydrogène sont remplacés par des atomes de deutérium. Ils sont légèrement moins électronégatifs que les atomes d'hydrogènes. Il en résulte que les liaisons O—D sont légèrement plus polarisables que les liaisons O—H et peuvent plus facilement être impliquées dans des liaisons hydrogènes. Or, les liaisons hydrogènes jouent un rôle fondamental dans de nombreuses réactions biologiques.

Ces différences entre les deux solvants sont vraisemblablement responsables de l'apparition de la phase lente dans le  $D_2O$  et du fait que la constante de vitesse de la phase rapide est légèrement plus faible que la constante de vitesse de repliement lorsque la réaction est effectuée dans l'eau.

### Modèle de réaction

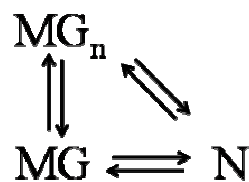
#### *Dans l'eau*

Pour les faibles concentrations en protéine (inférieures à 1 mg/mL), les cinétiques sont mono-exponentielles et la constante de vitesse de repliement ne varie pas. Un modèle simple à deux états suffit pour expliquer ces cinétiques. La protéine dans l'état *molten globule* se replie directement pour peupler l'état natif :



Lorsque la concentration en protéine augmente, la constante de vitesse diminue et le modèle précédent ne permet plus d'expliquer les cinétiques de repliement.

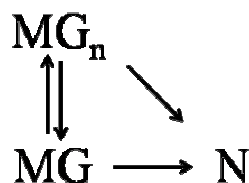
Les expériences réalisées en rhéométrie montrent que la viscosité de la solution augmente avec la concentration en protéine, en particulier à partir de quelques milligrammes de protéines par millilitre de solution. Cette augmentation de viscosité de la solution a été attribuée à des interactions entre des protéines. Ces interactions sont d'autant plus présentes que la distance moyenne entre les protéines diminue, donc lorsque la concentration augmente. Ces interactions doivent être prises en compte pour l'interprétation de l'évolution des constantes de vitesse de repliement. Un troisième état est introduit dans le modèle. Cet état correspond à un état *molten globule* qui interagit avec les autres protéines. Il sera noté  $MG_n$ . Dans le cas le plus général, le modèle à trois états peut être représenté de la manière suivante :



Les données expérimentales permettent de simplifier ce schéma.

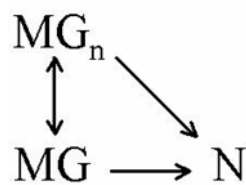
Les flèches qui pointent sur les états  $MG$  et  $MG_n$  depuis l'état natif représentent le passage de l'état natif aux états  $MG$  et  $MG_n$ . Or, en fin de réaction, toutes les protéines sont dans l'état natif, les passages de l'état natif aux états  $MG$  et  $MG_n$  peuvent donc être négligés.

Le schéma de réaction devient :



Ce schéma ne permet toujours pas d'expliquer correctement les données. Si les cinétiques suivaient ce schéma de réaction, elles seraient bi-exponentielles avec deux constantes de temps invariantes. Or les cinétiques sont mono-exponentielles avec une constante de temps qui varie avec la concentration.

Pour expliquer ces données expérimentales, l'équilibre entre les états MG et  $MG_n$  doit être considéré comme un équilibre rapide. Le schéma de réaction devient alors :



Ce modèle permet bien de décrire l'évolution de la constante de vitesse avec la concentration : à basse concentration, les protéines sont toutes dans l'état MG et se replient pour peupler l'état natif. Lorsque la concentration augmente, les protéines se rapprochent et peuvent interagir les unes avec les autres, ce qui leur permet de peupler l'état  $MG_n$ .

Les protéines dans l'état  $MG_n$  se replient moins rapidement que celles dans l'état MG, donc la constante de vitesse observée diminue au fur et à mesure que la concentration augmente et que l'état  $MG_n$  se peuple.

A haute concentration, la constante de vitesse associée à la réaction de repliement n'évolue plus. Cela s'explique par le fait que toutes les protéines dans l'état *molten globule* sont susceptibles d'interagir avec leurs voisines. Ainsi, toutes les protéines non repliées seront dans l'état  $MG_n$ .

A haute concentration, le modèle peut être réduit au suivant :



### *Dans le D<sub>2</sub>O*

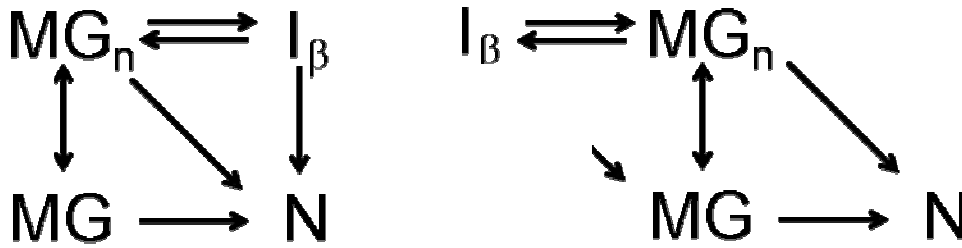
La constante de vitesse de la phase rapide dans le D<sub>2</sub>O évolue de la même manière que la constante de vitesse de repliement dans l'eau. Le modèle présenté précédemment permet donc d'expliquer l'évolution de la constante de vitesse associée à la phase rapide. En revanche, il ne permet pas d'expliquer l'apparition de la phase lente pour les hautes concentrations en protéine. Pour expliquer l'apparition de cette seconde phase, un quatrième état doit être introduit. Cet état sera noté I<sub>β</sub> pour des raisons qui seront éclaircies par la suite.

Pour expliquer le fait que cet état permet l'apparition d'une nouvelle phase, il ne doit pas être en équilibre rapide avec un autre état du modèle. En effet, les protéines sont piégées transitoirement dans cet état au cours de la cinétique et se replient ensuite lentement.

Le paramétrage des cinétiques a montré que la phase lente apparaît pour les concentrations supérieures à 2 mg/mL. La proportion du signal associée à la phase lente augmente pour des concentrations comprises entre 2 et 20 mg/mL et devient constante au-delà. L'amplitude relative de la phase lente évolue donc dans les mêmes gammes de concentration que la constante de vitesse de la phase rapide. Or, l'évolution de la constante de vitesse de la phase rapide est liée à la population de l'état MG<sub>n</sub>. L'état supplémentaire introduit dans le modèle est donc vraisemblablement formé à partir de l'état MG<sub>n</sub>.

Les données expérimentales nous renseignent sur l'état à partir duquel est accessible l'état I<sub>β</sub>, elles nous apprennent aussi que les protéines doivent pouvoir être piégées dans cet état. En revanche, il est impossible de savoir si l'intermédiaire peuplée au cours du repliement est sur ou en dehors du chemin de repliement.

Les considérations précédentes permettent de proposer deux modèles :



Ces deux modèles expliquent bien les résultats obtenus jusqu'à présent :

- A basse concentration (inférieur à 1 mg/mL), l'état  $MG_n$  n'est pas peuplé ; l'état  $I_\beta$  ne pourra donc pas être peuplé et la protéine se replie directement de l'état *molten globule* à l'état natif. Les cinétiques sont mono-exponentielles avec une vitesse de repliement constante.
- Pour les concentrations « intermédiaires » (entre 2 et 20 mg/mL environ), dès le début de la réaction, un équilibre s'établit entre les états  $MG$  et  $MG_n$ . Cet équilibre dépend de la concentration en protéine : l'état  $MG_n$  est d'autant plus peuplé que la concentration en protéine est importante. L'équilibre entre  $MG$  et  $MG_n$  induit la diminution de la constante de vitesse de la phase rapide. Comme l'état  $MG_n$  se peuple, l'état  $I_\beta$  peut à son tour se peupler. Une partie des protéines se replie directement depuis les états  $MG$  et  $MG_n$  ; elles sont à l'origine de la phase rapide détectée au cours des cinétiques. Une autre partie peuple l'état  $I_\beta$ , se replie beaucoup plus lentement et elles sont à l'origine de la phase lente.
- Pour les concentrations supérieures à 20 mg/mL, l'équilibre entre les états  $MG$  et  $MG_n$  est très favorable à  $MG_n$ . Ainsi dès le début de la réaction, les protéines se trouvent toute dans l'état  $MG_n$ . Une partie se replie alors rapidement, une autre peuple l'état  $I_\beta$  avant de se replier lentement. La constante de vitesse de la phase rapide et la proportion de chaque phase n'évoluent plus.

## Conclusion

A partir des résultats obtenus en dichroïsme circulaire et en rhéométrie, des modèles de réaction sont construits pour expliquer le repliement dans l'eau et dans le D<sub>2</sub>O. Par la suite, le repliement sera uniquement étudié dans le D<sub>2</sub>O. D'une part, parce que la spectroscopie infrarouge ne peut être utilisée qu'en D<sub>2</sub>O et, d'autre part, parce que la réaction est plus complexe dans le D<sub>2</sub>O. De plus, les états peuplés au cours de la réaction de repliement dans le D<sub>2</sub>O s'avéreront avoir des propriétés intéressantes.

Des expériences supplémentaires doivent être réalisées pour déterminer lequel des deux modèles décrit le mieux la réaction et pour avoir plus d'informations structurales sur les états introduits dans les modèles.

### **III.3. Spectres de corrélation 2D**

Les expériences de dichroïsme circulaire effectuées dans la partie précédente renseignent sur la quantité de structure tertiaire formée. Les cinétiques de repliement peuvent être déduites de ces mesures, mais aucune information précise sur la structure tertiaire n'est accessible. Pour cela, la spectroscopie par résonance magnétique nucléaire est utilisée.

La structure secondaire de l'état *molten globule* est très proche de celle de l'état natif, mais il existe tout de même de légères différences. Elles peuvent être détectées en utilisant la spectroscopie infrarouge à transformée de Fourier. Ces techniques permettent d'approfondir l'étude du repliement et de préciser la structure des états intermédiaires introduits dans le modèle.

Pour les expériences de dichroïsme circulaire, l'ellipticité est mesurée à une certaine longueur d'onde à intervalles de temps réguliers. Pour les expériences de spectroscopie RMN et infrarouge, un spectre entier est enregistré à intervalles de temps réguliers. La quantité d'informations apportées par ces expériences est donc bien plus grande que dans le cas du dichroïsme circulaire. Pour extraire ces informations, des spectres de corrélation bidimensionnels sont tracés à partir de ces séries de spectres unidimensionnels. Dans cette partie, le principe de construction et de lecture de ces spectres est expliqué. Les expériences de RMN et d'infrarouge et les spectres 2D tracés à partir des spectres 1D enregistrés par ces techniques sont présentés dans les parties suivantes.



### III.3.1. Principe

La spectroscopie de corrélation 2D a été principalement développée par Noda à la fin des années 1980 (refs 156, 157). Les spectres de corrélation 2D sont construits à partir d'une série de spectres 1D enregistrés sur un système soumis à une perturbation. La perturbation induit des changements dans la signature spectrale du système. La spectroscopie de corrélation 2D permet de corréler les changements qui ont lieu aux différentes longueurs d'onde dans le spectre 1D.

Dans notre cas, la perturbation en question est le saut de pH. Le saut de pH va déclencher le repliement de la protéine. Une série de spectres est enregistrée pendant cette réaction. Les changements occasionnés aux différentes longueurs d'onde pendant le repliement de la protéine sont corrélés entre eux grâce aux spectres de corrélation 2D.

Ces spectres sont calculés avec l'équation suivante :

$$\Phi(\nu_1, \nu_2) + i\Psi(\nu_1, \nu_2) = \frac{1}{\pi T} \int_0^{\infty} TF(y_{\nu_1})(\omega) \cdot TF^*(y_{\nu_2})(\omega) \cdot d\omega$$

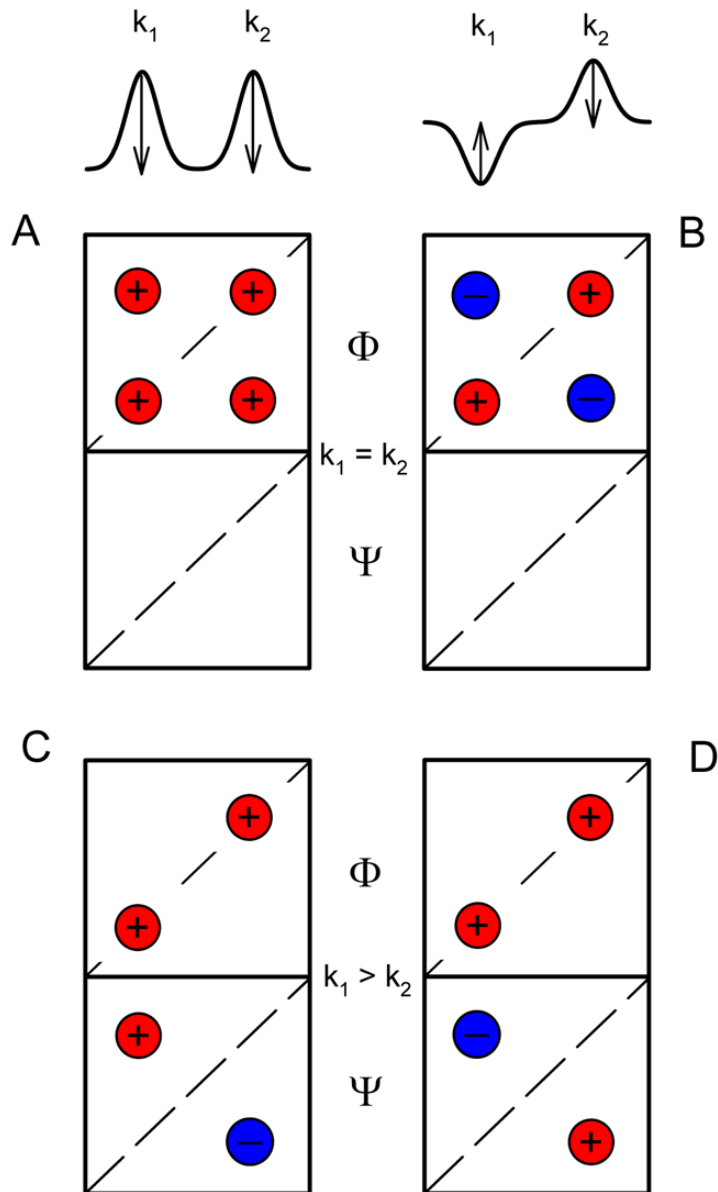
$y_\nu$  est l'évolution de l'intensité au cours du repliement à la longueur d'onde  $\nu$ .

Les deux fonctions  $\Phi$  et  $\psi$  obtenues correspondent respectivement au spectre de corrélation synchrone et au spectre de corrélation asynchrone. Un pic aux coordonnées  $(\nu_1, \nu_2)$  du spectre de corrélation synchrone indique que les événements en  $\nu_1$  et  $\nu_2$  sont corrélés ; un pic aux coordonnées  $(\nu_1, \nu_2)$  du spectre de corrélation asynchrone renseigne sur l'ordre dans lequel se déroulent les événements en  $\nu_1$  et  $\nu_2$ .

### III.3.2. Exemples

Des exemples de spectre de corrélation sont représentés en figure III.12.

Dans cette figure, les spectres de corrélations 2D sont tracés à partir de série de spectres 1D qui présentent deux bandes d'absorption, aux longueurs d'ondes  $\nu_1$  et  $\nu_2$ . Des exemples de spectres 1D sont représentés au dessus des spectres 2D. Pour les spectres A et C, les deux bandes sont positives. Pour les spectres B et D, la bande à la longueur d'onde  $\nu_1$  est négative et la bande à la longueur d'onde  $\nu_2$  est positive. Les constantes de vitesse de disparition des bandes sont égales à  $k_1$  pour la bande en  $\nu_1$  et à  $k_2$  pour la bande en  $\nu_2$ .



**Fig III.12 : Exemples de spectre de corrélation 2D.**

Pour les spectres A et B, les évènements sont corrélés, les bandes disparaissent à la même vitesse ( $k_1=k_2$ ).

Dans ce cas, quatre pics sont détectés sur le spectre de corrélation synchrone :

- Deux pics sont visibles sur la diagonale en pointillé. Ces pics sont situés aux coordonnées  $(v_1, v_1)$  et  $(v_2, v_2)$ . Ils indiquent que les évènements en  $v_1$  et  $v_2$  sont

corrélés avec eux-mêmes et n'apportent pas d'informations. Ces pics particuliers sont nommés autotopics.

- Les pics hors diagonales renseignent sur la corrélation synchrone des événements. Ici, deux pics hors diagonales sont détectés, en  $(\nu_1, \nu_2)$  et  $(\nu_2, \nu_1)$ . Ces pics indiquent que les événements en  $\nu_1$  et  $\nu_2$  se déroulent simultanément, ce qui est bien le cas dans l'exemple ( $k_1=k_2$ ). Le signe des pics hors diagonale dépend des signes des bandes dans le spectre 1D. Deux bandes de même signe dans le spectre 1D donnent un pic de corrélation positif ; dans la figure III.12A, les deux bandes sont positives, les pics hors diagonale sont positifs. Deux bandes de signe opposé dans le spectre 1D donnent un pic de corrélation négatif ; dans la figure III.12.B, la bande à la longueur d'onde  $\nu_1$  est négative et la bande à la longueur d'onde  $\nu_2$  est positive ; les pics aux longueurs d'ondes  $(\nu_1, \nu_2)$  et  $(\nu_2, \nu_1)$  sont dans ce cas négatifs.

Aucun pic n'apparaît sur le spectre asynchrone car tous les événements sont corrélés.

Dans le second cas (fig C et D), la bande en  $\nu_1$  évolue plus rapidement que la bande en  $\nu_2$  ( $k_1 > k_2$ ).

- Chacune des bandes est corrélée avec elle-même et les autotopics sont toujours présents sur le spectre asynchrone. En revanche, aucun pic hors diagonale n'est présent.
- Sur le spectre asynchrone, un pic positif est détecté aux coordonnées  $(\nu_1, \nu_2)$ . Les deux bandes sont de même signe, ce pic positif indique que le changement en  $\nu_1$  se déroule plus rapidement que le changement en  $\nu_2$ . Le pic aux coordonnées  $(\nu_2, \nu_1)$  est négatif, il indique que l'événement en  $\nu_2$  est plus lent que l'événement en  $\nu_1$  car les deux bandes sont de même signe. La figure D est construite à partir de série de spectres 1D dans lesquels les bandes ont des signes opposés. Le pic négatif aux coordonnées  $(\nu_1, \nu_2)$  indique dans ce cas que l'événement en  $\nu_1$  se déroule avant l'événement en  $\nu_2$ .

De la même manière, le pic positif en  $(\nu_2, \nu_1)$  montre que l'évolution du spectre en  $\nu_2$  est plus lente que l'évolution du spectre en  $\nu_1$ .

Les règles énoncées ci-dessus permettent de déduire des propriétés de symétries des spectres 2D. Le spectre synchrone est symétrique par rapport à la diagonale tracée en pointillé sur la figure III.12. En effet, si les événements en  $\nu_1$  et  $\nu_2$  sont corrélés, les pics apparaîtront en  $(\nu_1, \nu_2)$  et en  $(\nu_2, \nu_1)$  et leur signe sera le même. Le spectre asynchrone est antisymétrique par rapport à sa diagonale. Lorsque l'évènement en  $\nu_1$  se déroule avant l'évènement en  $\nu_2$ , un pic est présent en  $(\nu_1, \nu_2)$  ; l'évènement en  $\nu_2$  se déroule après l'évènement en  $\nu_1$  et le pic en  $(\nu_2, \nu_1)$  a un signe opposé au pic en  $(\nu_1, \nu_2)$ . Ces propriétés de symétrie permettent de réduire l'interprétation des spectres 2D à la moitié de ces spectres car les mêmes informations sont présentes au-dessus et au-dessous de la diagonale.

La spectroscopie de corrélation est ici utilisée pour interpréter des données issues d'expériences de repliement de l' $\alpha$ -lactalbumine suivies par RMN et par FTIR.

## **III.4. Résonance magnétique nucléaire du proton**

La spectroscopie RMN est utilisée pour sonder l'environnement des protons de la protéine et permet ainsi de mesurer la structure tertiaire des protéines.

### **III.4.1. Description de l'expérience**

#### Principe

Les atomes d'hydrogène possèdent un spin non nul. Les spins sont souvent comparés à de petites toupies qui peuvent interagir avec un champ magnétique. Les champs magnétiques sont utilisés pour aligner ces toupies et les faire tourner. Les expériences de RMN utilisent cette propriété des spins pour obtenir des informations sur la structure des protéines.

Plusieurs méthodes peuvent être utilisées en RMN.

Pour les expériences réalisées ici, les échantillons sont placés dans un champ magnétique intense et constant,  $B_0$ , ceci pour aligner les spins des hydrogènes. A l'aide d'un second champ magnétique  $B_1$  plus faible et oscillant à la fréquence  $\nu$ , on va pouvoir perturber la distribution des états imposée par le champ  $B_0$ . Lorsque  $\nu$  est égal à la fréquence de Larmor des spins ( $\frac{\gamma}{2\pi} B_0$ ) ( $\gamma$  est le rapport gyromagnétique), ceux-ci entrent en résonance avec le champ oscillant et absorbent une partie de son énergie. C'est ce phénomène qui est à l'origine des expériences RMN réalisées et ce sont les fréquences de résonance qui sont mesurées au cours de ces expériences. L'environnement atomique et moléculaire d'un atome peut avoir

divers effets sur le champ magnétique local. Le noyau est soumis alors à un champ effectif  $B_{\text{eff}}$  plus faible que le champ  $B$  appliqué:  $B_{\text{eff}} = (1 - 10^{-6} \sigma)B$ .  $\sigma$  est appelé constante d'écran et est exprimé en partie par million (ppm) en prenant pour référence un proton d'un groupe méthyle fortement blindé (un proton d'un groupe méthyle du triméthyl-silylpropionate par exemple).

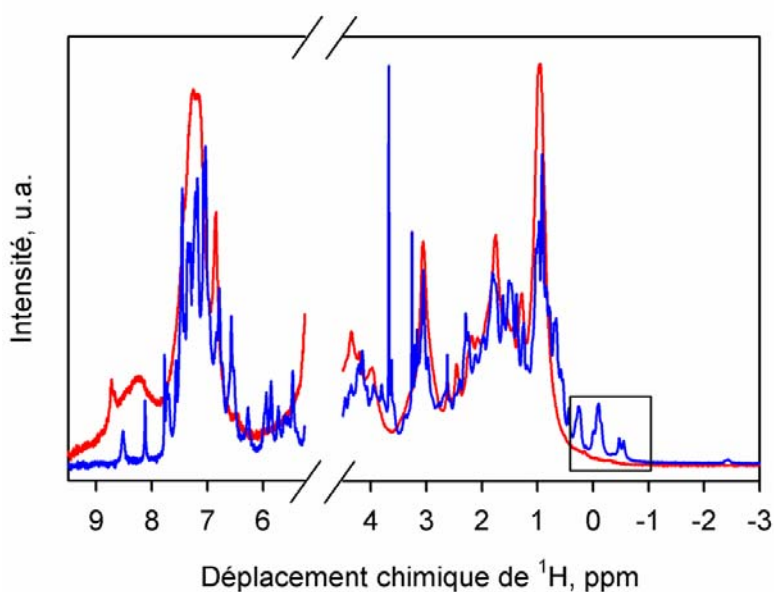
Le spectre RMN donne accès aux constantes d'écran et donc il renseigne sur le voisinage des atomes d'hydrogènes. Chaque hydrogène d'une protéine est ainsi caractérisé par un déplacement chimique (exprimé en ppm) égal à la valeur de  $\sigma$  locale qui dépend de l'environnement de l'hydrogène considéré.

### Conditions expérimentales

Les expériences présentées ont été réalisées avec un spectromètre BRUKER 500MHz. Le repliement est déclenché par un saut de pH, par mélange de 10 volumes de solution de protéine à pH 2 (protéine dans l'état *molten globule*) avec un volume d'une solution tampon de Tris 1 M pour atteindre une solution à pH 7,5 (pH favorable à la formation de l'état natif). La concentration en protéine est de 20mg/mL. Le mélange est effectué à la main, d'où un temps mort d'environ 3 min : mélange + mise en place de l'échantillon. Les expériences sont réalisées à 5°C dans du D<sub>2</sub>O. Des spectres sont enregistrés toutes les 30 s.

### III.4.2. Spectres 1D

Les spectres obtenus par RMN représentent l'énergie absorbée par la protéine en fonction du déplacement chimique. La figure III.13 représente les spectres RMN des états *molten globule* (en rouge) et natif (en bleu) de la protéine.



**Fig III.13 :** Spectre RMN des états natif et molten globule.

En bleu, spectre de l'état natif.

En rouge, spectre de l'état molten globule.

La zone encadrée correspond aux H portés par les acides aminés aromatiques, c'est cette zone qui sera utilisée dans la suite.

Le spectre de l'état *molten globule* présente des pics de résonance large. Cela est dû au fait que le *molten globule* n'a pas de structure tertiaire stable, et qu'un atome d'hydrogène de la protéine n'a pas le même environnement d'une protéine à l'autre. La constante d'écran  $\sigma$  associée à un hydrogène de la protéine n'est donc pas la même pour toutes les protéines de la solution. De plus la protéine dans l'état molten globule est très dynamique à l'échelle de la

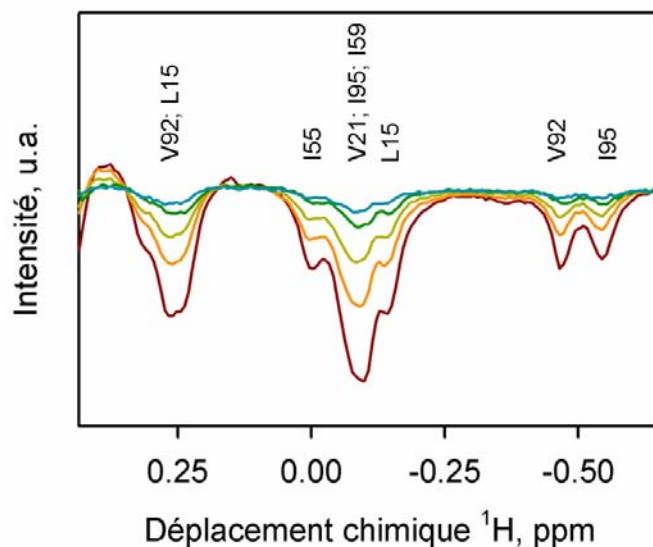


milliseconde. Les constantes d'écran sont donc dispersées, donnent des pics larges et un spectre mal résolu.

Lorsque les protéines sont dans l'état natif, elles possèdent une structure tertiaire stable. L'environnement d'un atome d'hydrogène est le même d'une protéine à l'autre ; la constante d'écran d'un hydrogène de la protéine est la même dans toutes les protéines. Lorsque le spectre 1D est tracé, les pics sont fins. Le spectre de l'état natif le montre bien sur la figure III.13.

Dans la suite, l'attention sera plus particulièrement portée sur la zone du spectre comprise entre  $-1$  et  $0,5$  ppm. Ces déplacements chimiques correspondent à la zone de résonance des protons portés par les cycles aromatiques de la protéine, elle est encadrée sur la figure III.13. Cette portion du spectre présente l'avantage de ne pas contenir de pic lorsque la protéine est dans l'état *molten globule*.

Pour souligner l'évolution du spectre RMN lorsque la protéine passe de l'état *molten globule* à l'état natif, des spectres de différences sont calculés. A chaque spectre enregistré en cours de cinétique est soustrait le spectre de fin de cinétique. Ces spectres de différence permettent de souligner l'évolution de l'intensité des pics au cours de la réaction. Ils sont représentés sur la figure III.14.



**Fig III.14 :** Spectres RMN de différence.

En rouge l'état initial, en orange à  $t=5\text{min}$ , en vert clair à  $t=10\text{min}$ , en vert foncé à  $t=25\text{min}$  et en bleu à  $t=45\text{min}$ .

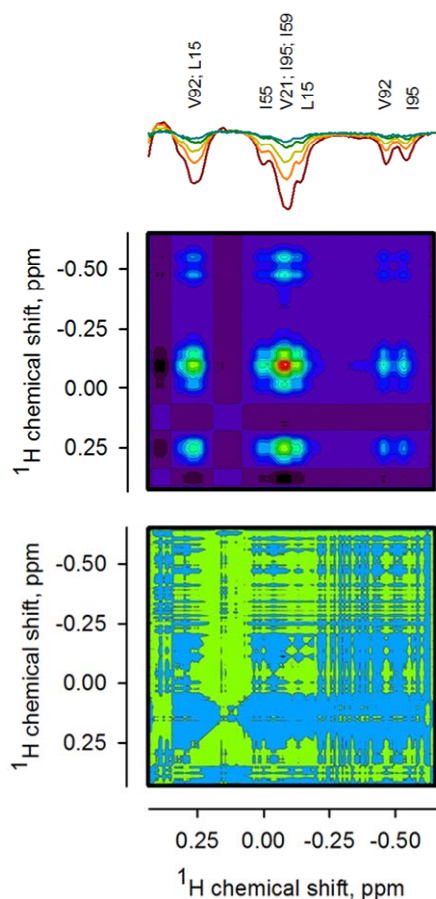
Au dessus de chaque pic d'absorption est indiqué l'acide aminé auquel appartient le proton sondé.

Plusieurs bandes sont présentes sur cette figure et certaines sont superposées. Chaque bande traduit la formation de structure tertiaire autour d'un atome d'hydrogène. Les acides aminés auxquels appartiennent les hydrogènes considérés sont indiqués au dessus de chaque bande<sup>158</sup>. Pour comparer l'évolution des différentes bandes, des spectres de corrélation sont tracés à partir de la série de spectres de différence enregistrés au cours de la cinétique.

### III.4.3. Spectres de corrélation 2D

La figure III.15 représente les spectres de corrélation synchrone et asynchrone.

Au dessus de ces spectres, les spectres de différences sont représentés pour permettre de repérer plus facilement les pics sur les spectres 2D.



**Fig III.15 : Spectres de corrélation 2D RMN.**

Les spectres de différences sont représentés au dessus des spectres de corrélation pour aider à situer les pics.

Le premier spectre 2D est le spectre de corrélation synchrone.

Le second est le spectre de corrélation asynchrone.

Sur le spectre de corrélation synchrone, des pics sont présents à tous les endroits possibles, c'est-à-dire pour toutes les coordonnées  $(v_i, v_j)$  tels que  $v_i$  et  $v_j$  correspondent chacune à la coordonnée d'un pic dans les spectres 1D.

Le spectre asynchrone ne présente aucun pic significatif et est quasiment nul.

Ces données indiquent que toutes les évolutions dans le spectre 1D se font de manière corrélée. Les bandes disparaissent toutes de façon identique dans les spectres RMN. Les cinétiques de formation de structure tertiaire autour des protons sondés sont semblables.

### III.4.4. Implications pour le modèle de réaction

Les acides aminés sondés dans ces expériences sont répartis dans toute la protéine, les résultats obtenus permettent donc de conclure que la formation de la structure tertiaire est entièrement corrélée. Ce constat permet de préciser le modèle de réaction construit à partir des résultats de dichroïsme circulaire.

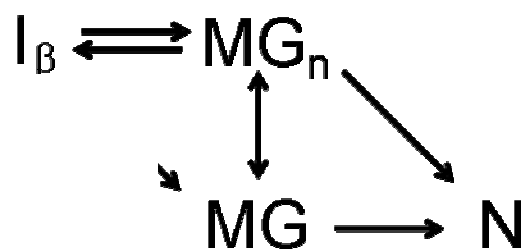
Tout d'abord, le fait que tous les éléments de structure tertiaire se forment en même temps indique que les états  $MG_n$  et  $I_\beta$  introduits dans les modèles de réaction ne possèdent pas de structure tertiaire. En effet, si ces états étaient structurés, l'apparition, puis la disparition de leur structure au cours de la réaction induirait l'apparition de pics sur le spectre de corrélation asynchrone.

Deux modèles sont proposés pour expliquer le repliement de l' $\alpha$ -lactalbumine. Dans un cas, la protéine se replie depuis les états MG et  $MG_n$ , dans l'autre, elle se replie depuis les états MG et  $I_\beta$ . Or les spectres de corrélation indiquent que tous les événements liés à la formation de la structure tertiaire sont corrélés, comme si la protéine se repliait depuis un seul état.

Sur le plan structural,  $MG_n$  est un état MG qui interagit avec ses voisins. C'est un état en équilibre rapide avec MG, donc la barrière d'énergie entre ces deux états est peu élevée. MG et  $MG_n$  sont structurellement très proches ; la manière dont la structure tertiaire se forme depuis ces deux états est donc assez semblable.

$I_\beta$  est un état dans lequel la protéine peut être piégée ; une barrière d'énergie existe donc entre  $I_\beta$  et les autres états. La structure de cet état est donc vraisemblablement assez différente de celle de l'état MG. La formation de structure tertiaire est corrélée, comme si l'état natif, qui est le seul à posséder de la structure tertiaire, se peuplait depuis un seul état.

D'après les considérations structurales exposées ci-dessus et les résultats obtenus par RMN, le premier modèle de repliement exposé dans la partie précédente semble le plus à même de décrire correctement la réaction.



## **III.5. Spectroscopie infrarouge**

Des informations sur la structure tertiaire des états présents dans le modèle de réaction ont été déduites des expériences de RMN. En revanche, peu d'informations sont disponibles en ce qui concerne la structure secondaire. Ce manque peut être pallié en utilisant la spectroscopie infrarouge à transformée de Fourier.

Cette technique spectroscopique permet de détecter avec une grande précision les structures secondaires des protéines, si bien que des différences infimes entre la structure secondaire de l'état *molten globule* et celle de l'état natif peuvent être mesurées.

### **III.5.1. Description de l'expérience**

#### **Principe**

Les liaisons covalentes des protéines peuvent absorber l'énergie des rayons infrarouges et les longueurs d'onde d'absorption sont caractéristiques de ces liaisons covalentes.

Par exemple, la bande d'absorption de la double liaison C=O qui se situe aux alentours de  $1550\text{ cm}^{-1}$  est particulièrement utile dans l'étude de la structure secondaire des protéines. En effet, les éléments de structure secondaire sont stabilisés par des liaisons hydrogènes au sein de la chaîne polypeptidique. Ces liaisons se font entre l'hydrogène porté par l'azote de la fonction amine et l'oxygène de la fonction carboxyle. Lorsqu'un oxygène est impliqué dans des liaisons hydrogènes, la bande d'absorption de la liaison C=O est décalée dans le spectre infra rouge ; ce décalage dépend du type de structure secondaire dans laquelle l'oxygène est

impliqué. Ainsi, chaque longueur d'onde d'absorption correspond à un certain type de structure secondaire.

Entre 1550 et 1750  $\text{cm}^{-1}$ , une large bande d'absorption est présente sur le spectre infrarouge de la protéine. Cette bande est appelée bande amide I et correspond à l'absorption par la liaison C=O. D'après ce qui est expliqué ci-dessus, cette bande est composée d'une somme de plusieurs bandes à différentes longueurs d'onde. Ces bandes qui composent la grande bande Amide I sont décorréées et chacune est attribuée à un type de structure secondaire à l'aide de tables, comme celle représentée en table III.1. L'amplitude d'une bande est proportionnelle à la quantité de structure secondaire à laquelle elle est associée.

Nombre d'onde ( $\text{cm}^{-1}$ )	Attribution
1621 - 1640	Feuillets $\beta$
1641 - 1647	Pelote statistique
1651 - 1657	Hélices $\alpha$
1658 - 1696	Coudes
1671 - 1679	Feuillets $\beta$

**Table III.1 : Positions caractéristiques des bandes amides I correspondant aux différents types de structure secondaire<sup>39</sup>.**

Les expériences sont réalisées dans le  $\text{D}_2\text{O}$ , car les liaisons covalentes H—O présentes dans l'eau absorbent aux mêmes longueurs d'onde que la double liaison C=O.

## Conditions expérimentales

Comme précédemment, le repliement est effectué en mélangeant 10 volumes d'une solution de protéine à pH 2 avec un volume de solution de Tris 1 M pour avoir une solution à pH 7,5-8 après mélange.

La concentration en protéine pour ces expériences est de 20 mg/mL. Les mélanges sont réalisés à l'aide d'un appareil de mélange rapide (SFM Stopped flow, Bio-logic), d'où un temps mort d'environ 100 ms. Les expériences sont réalisées à 5°C et dans le D<sub>2</sub>O.

Les mesures sont effectuées sur un spectromètre à transformée de Fourier JASCO FT/IR610.

Pour les cinétiques, un spectre est enregistré toutes les 30 s et correspond à la moyenne de 16 spectres mesurés. A l'équilibre, les spectres enregistrés correspondent à la moyenne de 1000 spectres mesurés.

### **III.5.2. Spectres 1D**

#### A l'équilibre

La figure III.16 représente la bande amide I' des spectres infrarouges de l'état *molten globule* et l'état natif.

La ressemblance des deux spectres confirme que le *molten globule* et la forme native ont une structure secondaire très proche.

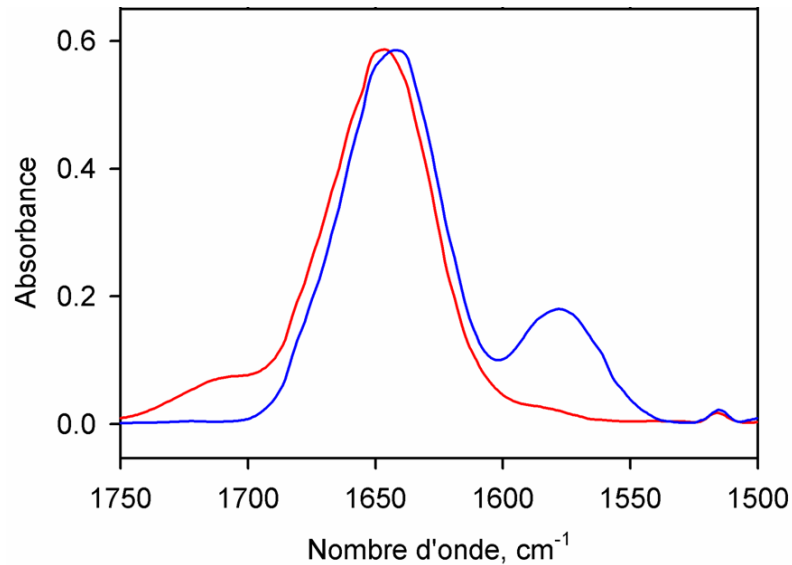
Il y a cependant une bande d'absorption qui se déplace de 1570 cm<sup>-1</sup> dans le spectre de la protéine dans l'état natif, à 1710 cm<sup>-1</sup> dans le spectre de la protéine dans l'état *molten globule*.

Cette bande est due à la protonation de certaines fonctions acides sous l'effet du pH (COO<sup>-</sup> →



COOH). En effet, il s'agit ici de mesures à l'équilibre, donc le spectre de l'état natif est enregistré à pH 7,5 alors que le spectre du *molten globule* est enregistré à pH 2.

L'effet observé ne trouve donc pas son origine dans une différence de structure secondaire.



**Fig III.16 :** Spectres d'absorption infrarouge des états natif et molten globule.

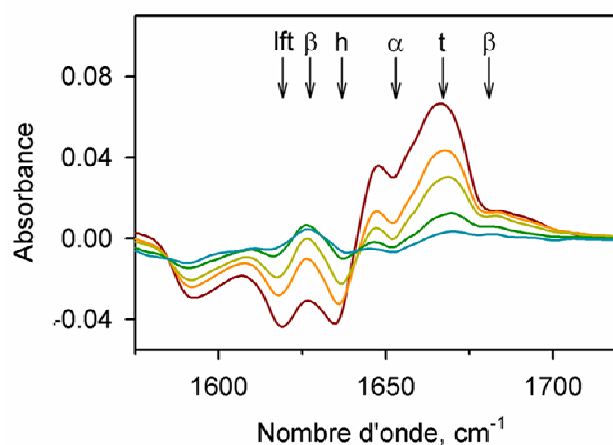
En rouge, l'état molten globule, en bleu l'état natif.

Comme les structures secondaires des deux états sont très proches, les spectres de différences sont particulièrement utiles pour souligner l'évolution du spectre infrarouge au cours de la cinétique.

### Cinétique

Des spectres infrarouges sont mesurés pendant le repliement de la protéine et les spectres de différences sont calculés en soustrayant à chaque spectre enregistré au cours du repliement le spectre en fin de repliement.

Ces spectres de différence sont représentés dans la figure III.17.



**Fig III.17 : Spectres infrarouges de différence.**

Ces spectres ont été calculés à partir de la cinétique effectuée avec 20 mg d'α-lactalbumine par mL de solution. En rouge l'état initial, en orange à  $t = 5$  min, en vert clair à  $t = 10$  min, en vert foncé à  $t = 25$  min et en bleu à  $t = 45$  min.

Les éléments de structure secondaire à l'origine des bandes d'absorption sont indiqués au dessus des spectres. lft : low frequency turn (structures secondaires mal définies) ;  $\beta$  : feuillets  $\beta$ , h : hélices  $3_{10}$ ,  $\alpha$  : hélices  $\alpha$ , t : turn (coude).

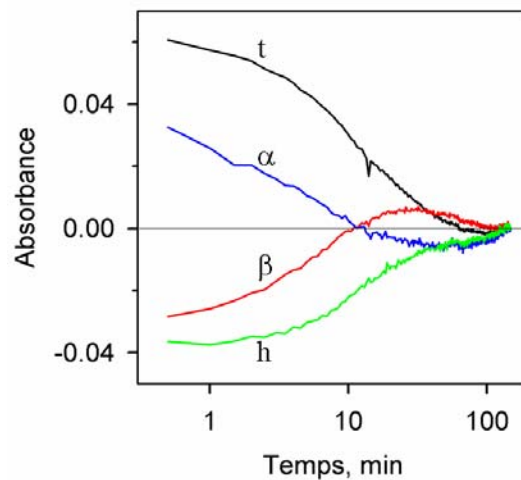
Ces spectres présentent plusieurs bandes d'absorption.

Les structures secondaires à l'origine de ces bandes sont identifiées et indiquées sur la figure III.17. Des coudes (t), des hélices  $\alpha$ , des hélices  $3_{10}$  (h), des feuillets  $\beta$  et des éléments de structures secondaires mal définis notés ici lft (low frequency turn) sont détectés dans l'α-lactalbumine<sup>39</sup>.

Sur les spectres de différences tracés en RMN, toutes les bandes sont négatives. Ce n'est pas le cas pour les spectres de différences enregistrés en infra rouge où des bandes négatives et positives sont présentes. La disparition d'une bande négative au cours du repliement est liée à l'apparition de la structure secondaire qui lui est associée. Inversement, lorsqu'une bande positive disparaît, la structure à laquelle elle est associée est détruite.

Une première idée sur la cinétique de formation de chacun des types de structure secondaire peut être déduite des spectres 1D. Pour cela, on suit l'évolution de l'intensité aux longueurs d'onde correspondant aux différentes structures secondaires.

De telles cinétiques sont représentées en figure III.18.



**Fig III.18 :** Évolution de l'intensité des bandes d'absorption au cours de la réaction de repliement.

Les types de structures secondaires à l'origine de l'absorption infrarouge sont indiqués à côté des courbes. La concentration en protéine est de 20 mg/mL.

On constate d'emblée que ces cinétiques ne sont pas simples et sont très différentes d'une structure à l'autre.

La courbe en noir sur cette figure correspond à l'évolution de l'intensité à la longueur d'onde d'absorption des coudes. Cette intensité est positive en début de réaction et décroît progressivement vers une valeur nulle. Ceci indique que les coudes disparaissent au cours de la réaction.

La courbe en vert montre l'évolution de l'intensité à la longueur d'onde d'absorption des hélices  $3_{10}$ . Cette intensité est négative et augmente pour atteindre une valeur nulle. Des hélices  $3_{10}$  apparaissent au cours de la réaction.

Les évolutions des intensités associées aux brins  $\beta$  (en rouge) et aux hélices  $\alpha$  (en bleu) sont plus complexes.

L'intensité de la bande associée aux brins  $\beta$  est négative en début de réaction. Cette intensité augmente et devient positive au bout d'une dizaine de minutes. Elle continue d'augmenter jusqu'à la quarantième minute environ puis décroît pour atteindre une valeur nulle en fin de réaction. Pendant le repliement, des brins  $\beta$  non-natifs sont repérés ; des brins formés en début de réaction sont détruits en fin de réaction, et sont responsables de l'« overshoot » positif détecté à la quarantième minute.

La courbe en bleu, associée aux hélices  $\alpha$  présente, elle aussi, un *overshoot*. Dans ce cas, l'*overshoot* est négatif : l'intensité en début de réaction est positive, elle diminue, prend une valeur négative au bout de dix minutes, elle continue de diminuer jusqu'à la soixantième minute puis augmente pour atteindre une valeur nulle en fin de réaction. Dans ce cas cependant, cet *overshoot* n'est pas dû à la formation d'hélices non natives. Son origine peut être déterminée en observant les spectres de différences dans la figure III.17. Sur ces spectres, la bande d'absorption des hélices  $\alpha$  est superposée à la bande d'absorption des coudes. La bande d'absorption des hélices est négative, mais elle est de faible amplitude et noyée dans la bande d'absorption positive des coudes qui, elle, est de grande amplitude. Les coudes disparaissent plus rapidement que les hélices  $\alpha$  n'apparaissent. Au bout d'une cinquantaine de minutes, les coudes ont quasiment disparu, alors que les hélices  $\alpha$  continuent à apparaître.

La spectroscopie de corrélation 2D est très utile lorsque les bandes d'absorption se chevauchent comme dans le cas présent.

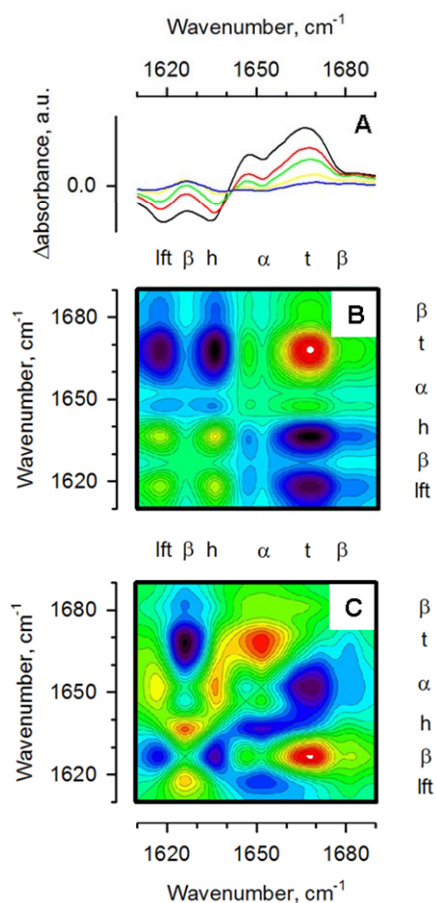
Ces spectres sont calculés et présentés dans la partie suivante. Pour les interpréter, il est important de garder en mémoire qu'au cours de la réaction, des coudes disparaissent alors que des lft et des hélices  $\alpha$  et  $3_{10}$  apparaissent. Le cas des brins  $\beta$  est plus complexe mais, sur la totalité de la réaction, les brins  $\beta$  apparaissent, les feuilletés étant plus nombreux en fin de réaction que dans l'état initial.

### III.5.3. Spectres de corrélation 2D

Les spectres de corrélation 2D sont représentés sur la figure III.19.

Sur ces spectres, les valeurs positives sont colorées en jaune et rouge, les pics négatifs sont représentés en violet et bleu.

Pour aider à situer les pics des spectres 2D, les spectres de différences sont représentés au dessus des spectres 2D et les longueurs d'ondes associées aux structures secondaires sont indiquées au dessus et à droite des spectres.



**Figure III.19 : Spectres de corrélation 2D infrarouges**

A. Spectres de différence représentés pour pouvoir situer plus facilement les pics.

B. Spectres de corrélation 2D synchrone.

C. Spectres de corrélation 2D asynchrone.

En plus des autotpics de la diagonale, des pics sont détectés aux coordonnées  $(\nu_t, \nu_{lft})$ ,  $(\nu_t, \nu_h)$  et  $(\nu_{lft}, \nu_h)$ . Ces pics indiquent que les évolutions des hélices  $3_{10}$ , des lft et des coudes se déroulent de manière synchrone. Les signes de ces pics sont en accord avec les évolutions observées dans la figure III.15 : les lft et les hélices  $3_{10}$  apparaissent au cours de la réaction, le

pic en  $(v_{lft}, v_h)$  est positif ; les coudes disparaissent au cours de la réaction, les pics en  $(v_t, v_{lft})$  et en  $(v_t, v_h)$  sont négatifs.

Sur ce spectre synchrone, aucun pic n'est détecté aux coordonnées correspondant aux brins  $\beta$  et aux hélices  $\alpha$ . Pour avoir des informations sur l'évolution de ces structures au cours du repliement, il faut s'intéresser au spectre asynchrone.

Le spectre asynchrone présente un pic négatif aux coordonnées  $(v_\beta, v_t)$  et deux pics positifs aux coordonnées  $(v_\beta, v_h)$  et  $(v_\beta, v_{lft})$ . Puisque les brins  $\beta$  apparaissent et les coudes disparaissent au cours de la réaction, et d'après les règles énoncées précédemment, les brins  $\beta$  évoluent plus rapidement que les coudes. Les coudes, les lft et les hélices  $3_{10}$  évoluent simultanément, les brins  $\beta$  évoluent donc plus rapidement que les lft et les hélices  $3_{10}$ . Ceci est confirmé par les pics aux coordonnées  $(v_\beta, v_h)$  et  $(v_\beta, v_{lft})$ . Ces pics sont, dans ce cas, positifs car, contrairement aux coudes, les hélices  $3_{10}$  et les lft apparaissent pendant le repliement.

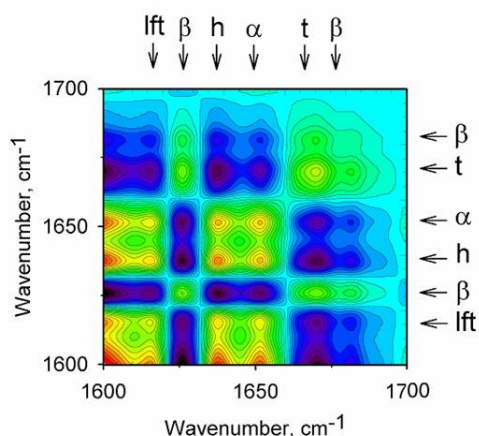
En ce qui concerne les hélices  $\alpha$ , un pic est présent aux coordonnées  $(v_\alpha, v_t)$ . Ce pic est positif. Comme les hélices  $\alpha$  apparaissent et que les coudes disparaissent, le signe du pic indique que les hélices  $\alpha$  évoluent plus lentement que les coudes. Elles évoluent donc plus lentement que les hélices  $3_{10}$  et les lft comme le confirment les pics positifs aux coordonnées  $(v_\alpha, v_h)$  et  $(v_\alpha, v_{lft})$ .

Pour résumer, les spectres de corrélation montrent que les brins  $\beta$  sont les structures secondaires qui évoluent le plus rapidement. Ensuite, et de manière simultanée, les hélices  $3_{10}$  et les lft apparaissent pendant que les coudes disparaissent. Enfin, les hélices  $\alpha$  sont les structures secondaires qui évoluent le moins rapidement.

Pour avoir une description plus complète de la réaction, le spectre de corrélation synchrone est calculé à partir des spectres enregistrés en fin de réaction, à partir de la quarantième

minute, c'est-à-dire, à partir de l'*overshoot* positif détecté dans la cinétique de formation de feuillets  $\beta$ .

Ce spectre est représenté sur la figure III.20.



**Fig III.20 :** Spectre de corrélation synchrone calculé à partir des spectres 1D enregistrés en fin de cinétique.

Sur ce spectre 2D, des pics sont visibles à tous les endroits où ils sont susceptibles de se trouver. Ceci indique que toutes les structures secondaires évoluent de manière corrélée. Ce spectre permet aussi de vérifier que les brins  $\beta$  disparaissent en fin de réaction. En effet, le pic aux coordonnées  $(\nu_{\beta}, \nu_t)$  est positif ; étant donné que les coudes disparaissent, ce pic positif indique que les brins  $\beta$  disparaissent eux aussi en fin de réaction. Les pics aux coordonnées  $(\nu_{\beta}, \nu_{lft})$ ,  $(\nu_{\beta}, \nu_h)$  et  $(\nu_{\beta}, \nu_{\alpha})$  sont négatifs car les hélices  $\alpha$  et  $3_{10}$  ainsi que les lft apparaissent en fin de réaction.

Sur la totalité de la réaction, les coudes, les hélices  $3_{10}$  et les lft évoluent de manière synchrone. Des feuillets  $\beta$  apparaissent rapidement en début de réaction et disparaissent ensuite. Les hélices  $\alpha$  apparaissent lentement. Ces résultats suggèrent qu'en fin de réaction, les feuillets  $\beta$  non-natifs se replient sous forme d'hélices  $\alpha$ .

### III.5.4. Implications pour le modèle de réaction

Les informations obtenues par spectroscopie infrarouge permettent de déduire des caractéristiques structurales sur l'intermédiaire  $I_\beta$  introduit dans le modèle de réaction.

La concentration utilisée pour ces expériences est de 20 mg/mL. Les études effectuées en dichroïsme circulaire ont permis de montrer qu'à cette concentration, au début du repliement, deux réactions sont en compétition : une partie des protéines se replie directement alors qu'une autre partie peuple un état appelé  $I_\beta$ . Les protéines de l'état  $I_\beta$  se replient beaucoup plus lentement.

La spectroscopie infrarouge indique que des brins  $\beta$  non natifs sont formés en début de réaction pour disparaître ensuite. La mise en commun des informations obtenues par les deux techniques montre une corrélation entre la population de l'état  $I_\beta$  et la présence de brins  $\beta$  non natifs. Ces brins non natifs appartiennent donc vraisemblablement à l'état  $I_\beta$ .

Sur la totalité de la réaction, les hélices  $\alpha$  sont les structures qui se forment le moins rapidement, alors que les brins  $\beta$  semblent se former le plus rapidement. Or, en fin de réaction, toutes les structures secondaires se forment de manière synchrone.

Ces résultats suggèrent qu'en début de réaction, la formation d'hélices  $\alpha$  est ralentie alors que la formation de brins  $\beta$  est accélérée. L'accélération de la formation de brins  $\beta$  s'explique par le fait que ces feuilletts sont formés dans l'état  $I_\beta$ . Pour expliquer le ralentissement de la formation d'hélices  $\alpha$ , il faut considérer que l'état  $I_\beta$  a moins d'hélices  $\alpha$  que les autres états. Ceci suggère que les brins  $\beta$  de l'état  $I_\beta$  sont formés à partir de brins structurés en hélices  $\alpha$  dans l'état natif de la protéine.



En fin de réaction, le repliement de l'état  $I_\beta$  à l'état natif s'accompagne d'une conversion entre ces deux types de structure secondaire.

Dans le modèle de réaction, l'état  $MG_n$  se forme lorsque les protéines dans l'état *molten globule* sont suffisamment proches pour interagir. L'état  $I_\beta$  est formé à partir de l'état  $MG_n$ . Le fait que les brins  $\beta$  non-natifs se forment lorsque les protéines sont suffisamment proches pour interagir suggère que ces feuillettes sont intermoléculaires, qu'ils sont formés à partir de brins appartenant à différentes protéines. Dans ce cas, l'état  $I_\beta$  est multimérique. Pour déterminer la taille de ces multimères, des expériences permettant d'estimer la taille de molécules en solution sont utilisées.

## **III.6. Chromatographie d'exclusion de taille**

### **III.6.1. Description de l'expérience**

#### Principe

Le terme chromatographie regroupe les techniques d'analyse dans lesquelles un échantillon contenant une ou plusieurs espèces est entraîné par un courant de phase mobile le long d'une phase stationnaire. Les différentes molécules contenues dans l'échantillon se déplacent à des vitesses différentes et sont ainsi distinguables. Dans le cas de la chromatographie d'exclusion de taille, un liquide contenant les espèces à distinguer est entraîné dans une colonne remplie de grain poreux. Les petites molécules peuvent pénétrer à l'intérieur des trous des grains poreux et sont ralenties. Les grosses molécules, elles, sont rapidement éluées. De cette manière les molécules peuvent être distinguées en fonction de leur taille.

#### Conditions expérimentales

Les expériences de chromatographie ont été réalisées sur une colonne HPLC SW300 (300 mm × 7,5 mm) (Waters). Cette colonne a été calibrée en utilisant des kits de calibration de basse et haute masse moléculaire (Amersham). Avant chaque expérience, la colonne a été équilibrée en utilisant un volume de tampon correspondant à 5 fois le volume de la colonne.

Le tampon utilisé est du Tris 100 mM à pH 7,5 pour les expériences sur la forme native, repliée.

Pour le *molten globule*, le tampon est du citrate de sodium 20mM à pH 3,5. A cette valeur de pH, la protéine est toujours dans l'état *molten globule* en absence de calcium.

Les expériences sont réalisées à 20°C avec un débit égal à 0,5 mL/min. L'élution de la protéine est suivie par absorption UV à 280 nm.

### III.6.2. Résultats

Les mesures de chromatographie par exclusion de taille ne peuvent pas être effectuées pendant la cinétique : d'une part, les temps d'élution sont trop élevés pour pouvoir mesurer une cinétique, d'autre part, la concentration en protéine n'est pas contrôlée, la technique utilisée implique des interactions avec les grains poreux qui sont dans la colonne et ces interactions peuvent perturber la cinétique. On utilise donc cette technique pour vérifier que les états initiaux et finaux de la protéine sont monomériques et, de ce fait, que les multimères transitoirement formés au cours de la cinétique ont bien disparus en fin de réaction.

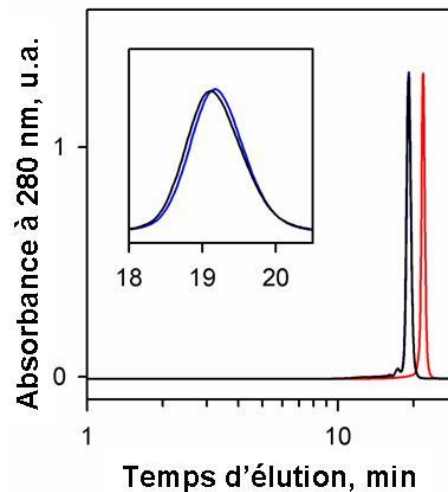
Trois solutions de protéine sont analysées par chromatographie d'exclusion de taille : une solution de protéine à pH 2 dans l'état *molten globule*, une solution de protéine à pH 7,5 dans l'état natif, et une solution de protéine prélevée après repliement suivi par infrarouge.

La figure III.21 représente l'absorption à 280 nm en fonction du temps d'élution pour les trois solutions : en rouge, la solution de protéine dans l'état *molten globule*, en noir, la solution de protéine après repliement depuis l'état *molten globule*, et en bleu, la solution de protéine dans l'état natif.

Les deux solutions de protéine à l'état natif présentent des pics superposés pour un temps d'élution d'environ 19 minutes. L'insert dans la figure est un agrandissement de la courbe

pour des temps d'élution compris entre 18 et 20,5 minutes. Cet insert nous permet de vérifier qu'il y a bien deux pics superposés correspondant à l'état natif.

Le pic en rouge correspondant à l'état *molten globule* apparaît pour un temps d'élution égal à environ 22 minutes.



**Fig III.21 :** Profil d'élution de solution d' $\alpha$ -lactalbumine dans différents états.

En rouge : protéine dans l'état molten globule.

En bleu : protéine dans l'état natif.

En noir : protéine dans l'état natif après repliement.

L'insert dans la figure représente un agrandissement des deux courbes de l'état natif entre 18 et 21 min d'élution. Il permet de montrer que la courbe noire et la courbe bleue sont superposées.

Le pic d'élution correspondant à l'état *molten globule* apparaît légèrement plus tard que ceux correspondant à l'état natif. Ceci est certainement dû au fait que la solution de protéine et le tampon utilisé pour équilibrer la colonne sont à pH 7,5 pour l'état natif, alors qu'ils sont à pH 3,5 pour l'état *molten globule*. Les interactions entre la matrice et la protéine peuvent légèrement varier en fonction du pH, et entraîner une différence dans le profil d'élution.

Aucun pic n'est détecté avant les pics d'élution principaux pour chacune des courbes, donc aucune des solutions ne possède d'oligomères ; toutes les protéines sont sous forme monomérique en début et en fin de réaction.

## **III.7. Diffusion quasi élastique de lumière**

Pour détecter la présence de multimères formés pendant la cinétique, des mesures de diffusion quasi élastique sont réalisées pendant que la protéine se replie.

### **III.7.1. Description de l'expérience**

#### Principe

La diffusion quasi élastique de lumière est aussi appelée diffusion dynamique de lumière ou spectroscopie par corrélation de photon.

Pour réaliser ces expériences, de la lumière issue d'un laser est envoyée sur une solution de protéine. Les protéines en solution diffusent alors cette lumière et la lumière diffusée est enregistrée grâce à des capteurs CCD.

Les protéines sont soumises au mouvement brownien en solution, l'intensité de la lumière diffusée enregistrée par les capteurs varie donc au cours du temps. Elle varie d'autant plus vite que le mouvement brownien est important.

La variation de l'intensité de diffusion au cours du temps permet d'estimer le mouvement brownien et d'avoir accès au coefficient de diffusion. Pour des particules sphériques, le coefficient de diffusion est donné par la formule de Stokes-Einstein :

$$D = \frac{k_B T}{6\pi\eta a}$$

Où  $k_B$  est la constante de Boltzmann,  $T$  la température,  $\eta$  la viscosité du solvant et  $a$  le rayon de la particule.

Ainsi, lorsque la particule est petite, le coefficient de diffusion est grand, et la particule soumise au mouvement brownien bouge rapidement. La corrélation des intensités de diffusion obtenues à différents temps permet de savoir de quelle manière le profil de diffusion varie, et donc de déterminer les coefficients de diffusion des protéines en solution. A partir de la formule de Stokes-Einstein et en supposant les particules sphériques, ces expériences de diffusion quasi élastiques permettent de connaître la taille des particules en solution.

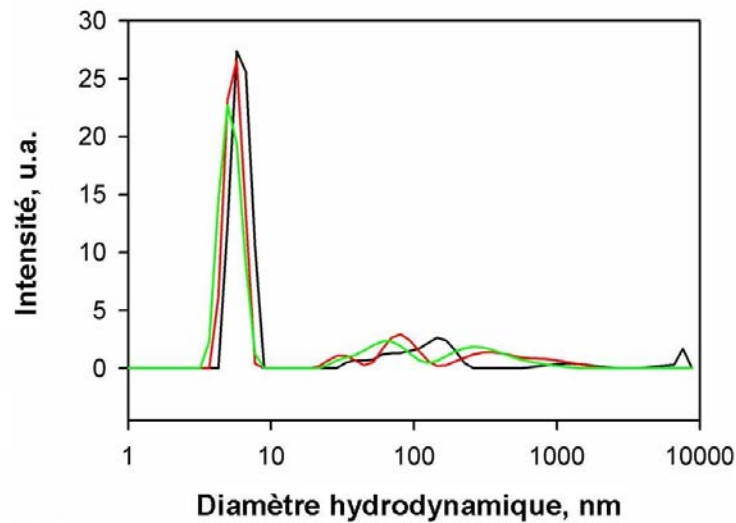
### Conditions expérimentales

Comme précédemment, les expériences sont déclenchées par saut de pH et sont effectuées à 5°C dans le D<sub>2</sub>O. Les mesures sont réalisées sur un appareil Zetasizer NanoZS (Malvern instruments). La concentration en protéine pour ces expériences est de 40 mg/mL.

## **III.7.2. Résultats et interprétation**

### Résultats

La figure III.22 montre les courbes obtenues en diffusion de lumière en début de cinétique (en noir), en milieu de cinétique (en rouge) et à l'état final (en vert). Pour toutes les courbes obtenues, un seul grand pic de diffusion est observé aux petits diamètres hydrodynamiques. Sur chacune des courbes, des pics secondaires de plus faible intensité et de moindre résolution sont visibles aux grands diamètres hydrodynamiques.



**Fig III.22 : Profils de diffusion enregistrés par DLS.**

En noir : Début de cinétique.  
 En rouge : Au bout de 40 min.  
 En vert : Fin de cinétique.

### Interprétation

La connaissance de la viscosité de la solution étudiée est nécessaire au traitement des données de diffusion quasi élastique de lumière. Les expériences de DLS ont été réalisées avant les expériences de rhéologie, et la variation de viscosité au cours du repliement nous était encore inconnue à l'époque. Pour ces expériences, la viscosité a été supposée constante au cours du repliement. Les valeurs des rayons hydrodynamiques calculés par cette technique sont donc erronées. Pour mesurer la viscosité en même temps que les expériences de DLS et donc avoir accès aux valeurs réelles des rayons hydrodynamiques, des particules de taille connue doivent être introduites dans l'échantillon. La viscosité est alors accessible à partir des valeurs du rayon apparent de ces particules. Malgré les erreurs dans les rayons de giration, des informations peuvent être déduites.

Les courbes représentées ici sont des courbes d'intensité. La hauteur des pics de diffusion dépend du nombre et de la taille des protéines en solution. Plus les protéines sont grosses, plus la lumière diffusée est intense. Les pics secondaires détectés sur les courbes correspondent à des structures de grande taille (entre 60 et 700 nm). Or, l'intensité des pics secondaires est beaucoup plus faible que celle du pic principal correspondant à des petites tailles. Ces états sont donc très peu peuplés et ne peuvent correspondre aux états multimériques détectés pendant la cinétique. De plus, ces pics secondaires sont détectés non seulement au cours, mais également, en début et en fin des cinétiques, lorsque les états sont monomériques. Ces pics secondaires ne sont donc pas significatifs et sont vraisemblablement dus à des impuretés dans la solution : ce type d'appareil a une sensibilité de l'ordre de 0,01 mg/mL ; comme la concentration en protéine est très élevée, 40 mg/mL, la moindre impureté sera visible.

Ainsi, un seul pic correspondant à l' $\alpha$ -lactalbumine est détecté sur les courbes de diffusion de lumière, y compris sur la courbe enregistrée au bout de 40 minutes de cinétique, au moment où l'état  $I_\beta$  est le plus peuplé. Les expériences de diffusion quasi élastique de lumière ne permettent donc pas de détecter la formation de grands multimères au cours du repliement.



## **III.8. Étude de la séquence de l' $\alpha$ -lactalbumine**

Les données de spectroscopie infrarouge montrent qu'une portion de la protéine est structurée en brin  $\beta$  dans l'état  $I_\beta$  et en hélice  $\alpha$  dans la forme native. Une étude de la séquence met en évidence une portion de la protéine possédant des propriétés physico-chimiques intéressantes qui pourraient expliquer les phénomènes observés au cours de la cinétique.

### **III.8.1. Méthodes utilisées**

#### Prédiction de structure secondaire

Les prédictions de structure secondaire effectuées ici sont réalisées par un logiciel du pôle bioinformatique lyonnais de l'Institut de Biologie et Chimie des Protéines. Le programme fait une prédiction en utilisant un consensus de trois méthodes : MLRC, DSC, PHD.

#### Profil d'hydrophobicité

Le programme « protscale » est disponible sur le serveur ExPASy de la société suisse de bioinformatique. Il permet de déterminer, entre autre, le profil d'hydrophobicité d'une protéine.

Chaque acide aminé est plus ou moins hydrophobe. Pour tracer un profil d'hydrophobicité, on analyse la séquence de la protéine et on associe à chaque acide aminé un nombre caractéristique de son hydrophobicité. Ce nombre est déterminé en utilisant l'hydrophobicité de l'acide aminé et de ses voisins.

La méthode utilisée ici est la méthode de Kyte & Doolittle avec une fenêtre de 9 acides aminés.

### Prédiction TANGO

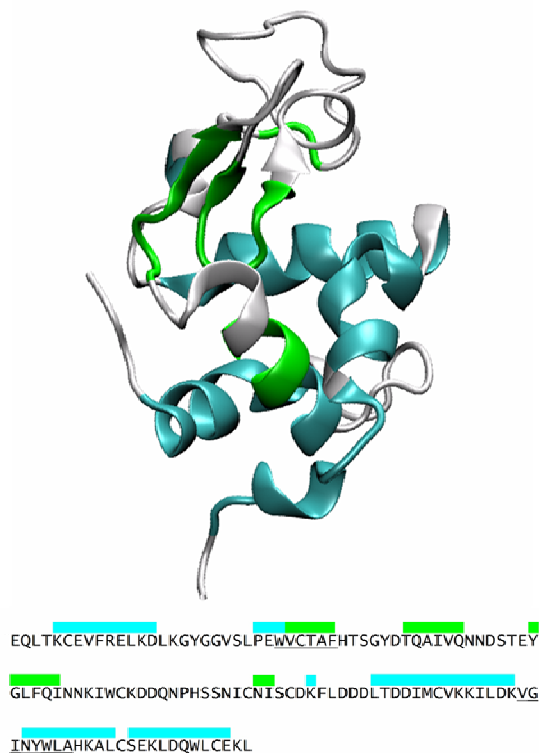
L'algorithme TANGO permet de prédire les portions de la protéine susceptibles d'être impliquées dans l'agrégation des protéines. Cette méthode a été développée par Serrano et ses collaborateurs à l'EMBL<sup>159</sup>, elle utilise la prédiction de structure secondaire, la stabilité et les conditions physico-chimiques (pH, concentration, force ionique...).

Le programme TANGO est disponible sur le web (<http://tango.crg.es/>). Les paramètres ont été ajustés pour correspondre aux conditions expérimentales : T= 5°C, force ionique : 0,1 M. Le seuil de détection a été laissé à 1.

## **III.8.2 Résultats**

La prédiction de structure secondaire réalisée sur la séquence de l' $\alpha$ -lactalbumine est représentée sur la figure III.23.

Cette figure montre une assez bonne concordance entre la prédiction de structure secondaire et la structure tridimensionnelle résolue par diffraction. Une hélice cependant est prédite en brin  $\beta$ . Cette hélice présente, en plus, un caractère hydrophobe, elle est enfouie au cœur de la protéine, et l'algorithme TANGO indique que cette hélice a une propension à former des fibres amyloïdes.



**Fig III.23 :** Structure de l' $\alpha$ -lactalbumine et prédictions réalisées sur la séquence.

Les portions de la structure colorées en bleu correspondent aux parties prédites en hélices  $\alpha$ . Les portions de la structure colorées en vert correspondent aux parties prédites en feuillets  $\beta$ .

Ces portions sont indiquées sur la séquence par des rectangles colorés situés aux dessus des lettres codant les acides aminés. Les portions hydrophobes sont celles dont les lettres sont surlignées en rouge. Les portions susceptibles d'être impliquées dans la formation de fibres amyloïdes selon l'algorithme TANGO sont soulignées.

Cette hélice est un bon candidat à la formation du feuillet  $\beta$  intermoléculaire de la forme  $I_\beta$ . En effet, elle peut être considérée comme un brin  $\beta$  « frustré » qui ne peut trouver de brin partenaire pour former un feuillet. Le brin « frustré » est hydrophobe, et séquestré à l'intérieur de la protéine lorsque la protéine est repliée.

Au début de la réaction, la protéine est dans l'état *molten globule* (MG) et ne possède pas de structure tertiaire stable. L'hélice hydrophobe est donc accessible au solvant et aux parties hydrophobes des autres protéines. Lorsque la concentration en protéine est suffisamment élevée, les parties hydrophobes des protéines peuvent s'attirer et entraîner ainsi la formation de multimères de protéine ( $MG_n$ ). Une fois les parties hydrophobes en contact, les brins  $\beta$  frustrés ne sont plus isolés et peuvent former un feuillet intermoléculaire. La protéine se trouve alors dans l'état  $I_\beta$ . La formation du feuillet stabilise le multimère et cet état sera transitoirement peuplé au cours du repliement.

### **III.9. Synthèse, Discussion**

Toutes les techniques utilisées ici ont permis d'esquisser un modèle de réaction et d'obtenir des informations sur les états présents dans les modèles de réaction. D'après les résultats obtenus, le modèle qui décrit le mieux la réaction est un modèle à 4 états.

En début de réaction, les protéines sont dans un état *molten globule*. Lorsque la concentration est suffisamment élevée, ces protéines peuvent interagir entre elles. Il en résulte une augmentation de la viscosité et une diminution de la constante de vitesse de repliement. Pour tenir compte de ces interactions dans le modèle de réaction, on utilise un état noté  $MG_n$  en équilibre rapide avec l'état *molten globule*.

Si la concentration est trop faible pour que les protéines dans l'état *molten globule* interagissent entre elles, elles se replient directement et peuplent l'état natif. Lorsque la concentration augmente et que les protéines dans l'état *molten globule* peuvent interagir, elles peuvent soit se replier et trouver leur forme native, soit mettre à contribution une partie de leur séquence pour former un feuillet beta intermoléculaire. Elles se retrouvent alors dans un état noté  $I_\beta$ . Ces feuillets seront détruits lorsque la protéine se replie.

#### Description structurale

Les expériences de rhéologie montrent que les solutions de protéines dans l'état *molten globule* sont beaucoup plus visqueuses que les solutions de protéines dans l'état natif. L'augmentation de viscosité est attribuée aux interactions entre les protéines. Les protéines interagissent donc plus dans l'état *molten globule* que dans l'état natif. Les deux formes de la protéine possèdent la même structure secondaire. En revanche, l'état *molten globule* ne possède pas de structure tertiaire stable, ce qui permet aux parties hydrophobes de la protéine

d'être accessibles au solvant ainsi qu'aux parties hydrophobes des protéines voisines. Les interactions à l'origine de l'augmentation de viscosité semblent donc être de nature hydrophobe.

La partie la plus hydrophobe est une hélice enfouie au sein de la protéine lorsque celle-ci est repliée. Cette protéine possède en outre la particularité d'être prédite en brin  $\beta$  par les logiciels de prédiction de structure secondaire. Ces hélices sont donc considérées comme des brins  $\beta$  « frustrés » qui ne trouvent pas de partenaire pour former des feuillets et qui, séquestrées dans la protéine, se résolvent à se replier sous forme d'hélices. Lorsque les interactions hydrophobes entre les protéines sont rendues possibles ces brins « frustrés » trouvent un partenaire et forment un feuillet  $\beta$ . La formation de ce feuillet  $\beta$  stabilise les protéines, elles sont transitoirement piégées dans cet état avant de se replier.

On peut noter cependant la présence d'une cystéine dans l'hélice considérée. Cette cystéine est impliquée dans un pont disulfure avec un cystéine de l'avant dernière hélice. Ce pont disulfure a tendance à rigidifier la structure, à maintenir la portion d'intérêt au sein de la protéine. Cependant, au vu des résultats expérimentaux et des prédictions effectuées sur la séquence, cette hélice semble la plus à même de former les brins  $\beta$  non natifs détectés au cours du repliement.

#### Lien avec les fibres amyloïdes

Les feuillets  $\beta$  intermoléculaires sont des structures caractéristiques des fibres amyloïdes. De plus, l' $\alpha$ -lactalbumine est une protéine susceptible de former des fibres amyloïdes à partir de son état *molten globule*<sup>160</sup>.

Les fibres amyloïdes se forment par un processus de nucléation croissance. Des protéines commencent par se rassembler pour former des multimères peu stables. Ces multimères

grandissent et se structurent peu à peu. Cette structuration va peu à peu stabiliser ces multimères. A partir d'une certaine taille et d'un certain degré de structuration, ces multimères deviennent des noyaux sur lesquels les protéines pourront directement se fixer dans une configuration de feuillets  $\beta$  intermoléculaires pour faire grandir la fibre.

Les intermédiaires détectés au cours de la cinétique possèdent des feuillets  $\beta$  intermoléculaires, qui font penser à des fibres. Cependant, aucun oligomère de grande taille n'est détecté. L'état  $I_\beta$  ne semble pas être un état de fibres amyloïdes. De plus, les fibres amyloïdes sont des structures très stables : des protéines impliquées dans leur formation peuvent difficilement s'en détacher pour se replier. En revanche, l'état  $I_\beta$  détecté au cours de la réaction pourrait être un état préfibrillaire dans lequel les protéines ne sont pas suffisamment stabilisées pour pouvoir former un noyau.

Les expériences réalisées dans cette partie permettent donc d'observer la compétition entre la formation de fibre et la réaction de repliement. Une telle compétition a été suggérée mais n'avait jusqu'à présent pas été observée expérimentalement.

# **IV. Formation de fibres amyloïdes**

Le repliement de la forme apo a été étudié par de nombreuses méthodes spectroscopiques dans la partie précédente. Les résultats obtenus permettent de proposer un modèle de réaction à quatre états. L'état  $I_{\beta}$  est responsable de la phase lente détectée en cours de cinétique. Il est riche en feuillets  $\beta$  intermoléculaires qui sont des structures caractéristiques des fibres amyloïdes et il a été considéré comme un précurseur de ces fibres. L'étude de la séquence de l' $\alpha$ -lactalbumine met en évidence une portion de la protéine qui présente des caractéristiques intéressantes : elle est hydrophobe et prédite en feuillets  $\beta$ . Lorsque la protéine est sous forme native, cette portion est enfouie au cœur de la structure et se présente sous forme d'hélice  $\alpha$ . En revanche, lorsque l' $\alpha$ -lactalbumine est dans un état *molten globule*, elle ne possède pas de structure tertiaire stable. La portion hydrophobe est donc accessible au solvant et aux autres protéines. C'est cette absence de structure tertiaire qui permet d'après les observations effectuées, de former les feuillets  $\beta$  intermoléculaires et éventuellement des fibres.

A partir de l'état *molten globule*, la protéine peut donc se replier ou former des fibres amyloïdes. Dans les expériences précédentes, les conditions expérimentales sont favorables à la forme native des protéines. Bien qu'un état pré-fibrillaire ( $I_{\beta}$ ) soit observé, les protéines finissent par se replier car la forme native est la plus stable. Pour empêcher le repliement et observer la formation de fibres, les protéines sont placées à pH 2. Le pH acide est favorable à la forme *molten globule*. Dans ce cas, la protéine ne se replie pas : elle ne possède pas de structure tertiaire stable et l'hélice hydrophobe reste accessible.

Les réactions de formation de feuillets  $\beta$  intermoléculaires au cours du repliement (à pH 7) et de formation de fibres amyloïdes à pH 2 se font toutes deux depuis l'état *molten globule*. Cependant, elles ne peuvent être comparées directement car les conditions expérimentales sont différentes. La charge de la protéine, en particulier, change avec le pH. Ces charges sont à l'origine des interactions électrostatiques entre les chaînes polypeptidiques. Pour compenser la variation de charges, plusieurs concentrations en NaCl sont utilisées. Le sel



permet d'écranter les charges électrostatiques et de moduler ainsi les interactions entre protéines.

## **IV.1. Formation de fibres amyloïdes à partir de l' $\alpha$ -lactalbumine**

### **IV.1.1. Microscopie électronique**

#### Principe

Pour les expériences de microscopie électronique en transmission, les fibres sont déposées sur une plaque mince puis colorées à l'aide de molécules qui absorbent les électrons.

La plaque est alors bombardée par un faisceau d'électrons. Certains sont absorbés par l'échantillon, d'autres passent à travers. Derrière la plaque, des capteurs permettent de restituer une image par transmission des fibres déposées.

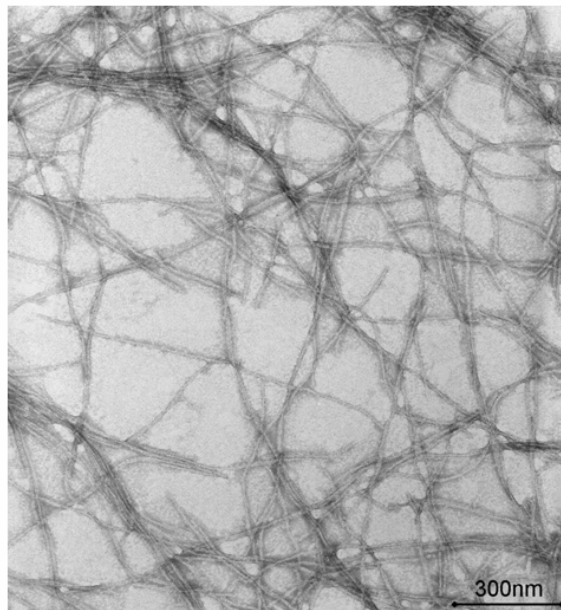
#### Conditions expérimentales

Les fibres sont formées en utilisant une solution d' $\alpha$ -lactalbumine à 10 mg/ml en présence de 100 mM de NaCl. La solution est ajustée à pH 2, placée à 40 °C et agitée légèrement pendant plusieurs heures.

Pour la microscopie électronique, une goutte de solution de fibres à 1 mg/mL est déposée sur la grille de cuivre enduit d'une résine (Formware). L'échantillon est évaporé pendant 10 à 30 minutes. Les fibres se sédimentent sur la grille. Elles sont ensuite rincées avec une solution d'uranyle acétate à 2% pour les colorer. Les clichés sont enregistrés sur un microscope électronique Phillips TECNAI 12 Biowin à 80 kV.

## Résultats

La figure IV.4 présente un cliché de microscopie électronique. Cette photo montre des filaments d'un diamètre de 5 à 10 nm, de plusieurs micromètres de longueur, avec un cœur clair et des bords sombres. L'allure et les dimensions observées correspondent bien à celles de fibres amyloïdes.



**Fig IV.4 :** Fibres amyloïdes formées par l' $\alpha$ -lactalbumine observées par microscopie électronique.

### **IV.1.2. Formation de fibres suivie par fluorescence de la Thioflavine T**

#### Principe

L'utilisation de thioflavine T pour détecter les fibres est une technique largement répandue, et utilisée tant en biophysique qu'en histologie<sup>161,162</sup>.

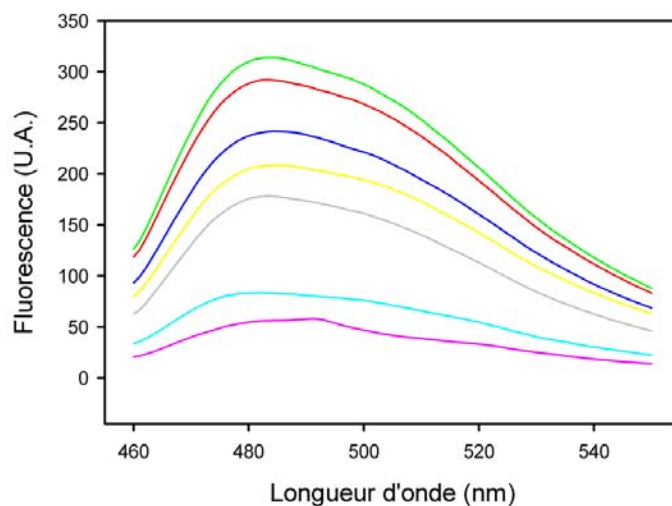
Cette molécule a la particularité de se fixer sur les feuillets  $\beta$  intermoléculaires présents dans les fibres amyloïdes<sup>163</sup>. Cette fixation modifie son spectre de fluorescence : l'intensité de fluorescence entre 450 et 560 nm augmente linéairement avec la concentration en fibre, un maximum étant détecté à 480 nm. Cette méthode sera utilisée pour suivre les cinétiques de formation de fibres.

### Matériel et méthode

Pour suivre les cinétiques, les protéines, à 10 mg/mL, sont placées dans des solutions contenant différentes concentrations de NaCl (0, 30, 60 et 100 mM). Les solutions sont ajustées à pH 2 puis placées à 40°C et agitées légèrement. Un échantillon est prélevé toutes les 3 h et dilué dans une solution de thioflavine pour obtenir une concentration de 10  $\mu$ M Tht et 10  $\mu$ M protéine. La fluorescence du mélange obtenu est mesurée sur un spectrofluorimètre FP-6500 (Jasco), la longueur d'onde d'excitation est 440 nm et la fluorescence est mesurée à 480 nm.

### Résultats

Des spectres de fluorescence enregistrés pendant la cinétique de formation de fibres sont représentés sur la figure IV.1. L'intensité de fluorescence de la thioflavine T augmente bien au cours de la réaction. Les spectres de fin de réaction sont caractéristiques de la thioflavine T en présence de fibres amyloïdes : forte intensité entre 450 et 560 nm avec un maximum à 480 nm.

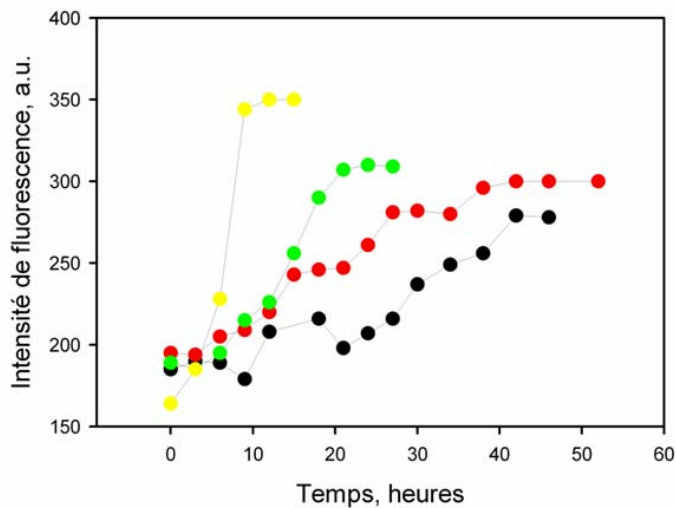


**Fig IV.1 :** Spectres de fluorescence de la Thioflavine T enregistrés pendant la cinétique de formation de fibres.

Les spectres sont enregistrés toutes les 3 heures:

- |            |             |
|------------|-------------|
| — 0 heure  | — 12 heures |
| — 3 heures | — 15 heures |
| — 6 heures | — 18 heures |
| — 9 heures |             |

Pour représenter les cinétiques de formation des fibres suivies par fluorescence de la thioflavine T, la valeur de l'intensité de fluorescence à 480 nm est représentée en fonction du temps de réaction. Les cinétiques enregistrées pour différentes concentrations de NaCl sont représentées sur la figure IV.2.



**Fig IV.2 : Cinétiques de formation de fibres suivies par augmentation de fluorescence de la thioflavine**

Différentes concentrations de NaCl sont utilisées: (●) 0mM, (●) 30mM, (●) 60mM, (●) 100mM.

Les mesures effectuées sur les différentes solutions de protéine montrent toutes une augmentation de l'intensité de fluorescence. Cela indique que des fibres sont présentes dans tous les échantillons en fin de réaction. Elles sont formées plus ou moins rapidement selon la concentration en sel. Plus la concentration en sel est forte, plus la réaction est rapide. Pour 100 mM de NaCl, les fibres sont formées en une dizaine d'heures ; pour 60 mM, en 20 h ; et lorsque la concentration descend en dessous de 30 mM, la réaction dure plus de 40 h.

Trois phases peuvent être distinguées dans ces cinétiques. Dans un premier temps, le signal n'évolue pas ou peu. Cette phase est appelée phase de latence. Ensuite, l'intensité augmente brusquement. C'est la phase de croissance. Enfin, le signal de fluorescence n'évolue plus, la réaction de formation de fibre est terminée.

Ces trois phases sont présentes quelque soit la concentration en sel, mais les durées et autres paramètres associés à chacune de ces phases varient d'une concentration à l'autre.

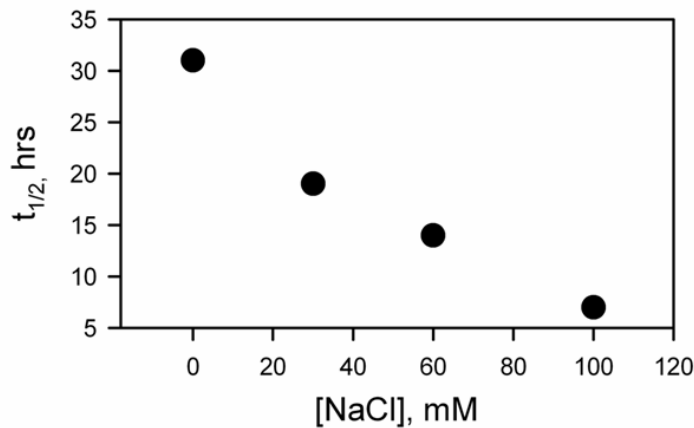
La durée de la phase de latence diminue fortement lorsque la concentration en sel augmente. Elle est de plus de 10 h pour les concentrations inférieures à 30 mM, et est quasiment inexistante lorsque la concentration est égale à 100mM.

La vitesse d'élongation correspond à la vitesse à laquelle la fluorescence augmente pendant la phase de croissance. Plus la concentration en sel est élevée, plus l'augmentation de fluorescence pendant la phase de croissance est rapide, donc plus la vitesse d'élongation est grande. Cela est bien visible sur la figure IV.2 : la courbe enregistrée à 100 mM de NaCl varie plus rapidement dans sa phase de croissance que celle enregistrée à 60 mM. De même, la vitesse d'élongation en présence de 60 mM de sel est plus grande que celle en présence de 30 mM, et elle-même, plus grande que celle en absence de sel.

Les données enregistrées ici ne permettent pas de déterminer avec certitude la durée de la phase de latence et la vitesse d'élongation. Cependant, ces deux phénomènes évoluent dans le même sens avec la concentration en sel : lorsque la concentration en sel augmente, la durée de la phase de latence diminue et la vitesse d'élongation augmente. Dans les deux cas, l'ajout de sel accélère la réaction.

Pour mettre en évidence cette accélération, on utilise le temps de demi-réaction. Ce paramètre correspond au temps pour lequel la moitié des protéines est sous forme de fibres amyloïdes. Il est déterminé à l'aide des courbes d'augmentation de fluorescence présentées dans la figure IV.2.

Les temps de demi-réaction en présence des différentes concentrations de NaCl sont représentés dans la figure IV.3.



**Fig IV.3 :** Temps de demi réaction de la réaction de formation de fibres pour différentes concentrations en NaCl.

Cette figure montre bien que le temps de demi-réaction diminue avec la concentration en sel. Il est égal à 31 h en absence NaCl ; il passe à 19 h en présence de 30 mM de NaCl, 14 h pour 60 mM, et 6 h pour 100mM. La réaction de formation de fibre est accélérée lorsque la force ionique augmente.

### Discussion

Les trois phases observées correspondent aux trois phases du phénomène de nucléation-croissance décrit dans l'introduction. En début de réaction, pendant la phase de nucléation, les protéines commencent à s'assembler en petites structures, mais ces structures ne sont pas stables et ne constituent pas des fibres amyloïdes, elles ne permettent pas d'augmenter le signal de fluorescence de la thioflavine T. A la fin de la phase de latence, les oligomères deviennent suffisamment stables et structurés pour former des noyaux. Ces noyaux sont des petites fibres amyloïdes, sur lesquelles les protéines peuvent se fixer directement en configuration de feuilletts  $\beta$  croisés. Les fibres « poussent » sur ces noyaux et s'allongent très



rapidement. C'est ce qui explique les changements importants dans l'intensité de fluorescence de la thioflavine pendant la phase de croissance. Ce phénomène se poursuit jusqu'à ce qu'il n'y ait plus de protéine monomérique en solution : le signal n'évolue plus (troisième phase de la cinétique).

Les phénomènes de multimérisation comme la formation de fibres dépendent beaucoup de l'interaction entre les protéines ; les protéines doivent se rencontrer et interagir pour pouvoir s'assembler. Il n'est donc pas étonnant de voir varier les cinétiques lorsque la concentration en sel, et donc les répulsions électrostatiques entre les protéines, changent<sup>164</sup>.

Les différences observées lorsque la concentration en sel varie peuvent avoir plusieurs origines. Elles peuvent provenir, comme évoqué ci-dessus, d'une modification des interactions entre protéines. L'ajout de sel permet d'écranter les charges des protéines et d'atténuer la répulsion électrostatique. En absence de cette répulsion, les chaînes polypeptidiques interagissent plus facilement. L'accélération de la réaction peut aussi provenir d'un effet sur la protéine elle-même et non plus sur les interactions entre molécules. Un effet du sel sur la stabilité ou la conformation de la protéine est imaginable. Par exemple, la zone hydrophobe de la protéine pourrait être alors plus largement exposée ce qui modifierait les interactions hydrophobes ; la conversion hélice  $\alpha \rightarrow$  brin  $\beta$ , nécessaire à la formation de fibres, pourrait aussi être facilitée.

Dans la partie suivante, des clichés de microscopie électronique et des expériences d'infrarouge sont réalisés sur les fibres obtenues. Cela permet de s'assurer qu'il s'agit bien de fibres amyloïdes. Dans la suite, l'action du sel sur les protéines et leurs interactions sont approfondies.

### IV.1.3. Spectroscopie infrarouge

La spectroscopie infrarouge est utilisée ici car elle permet de détecter les structures secondaires, en particulier les feuillets  $\beta$  intermoléculaires présents dans les fibres amyloïdes.

#### Conditions expérimentales

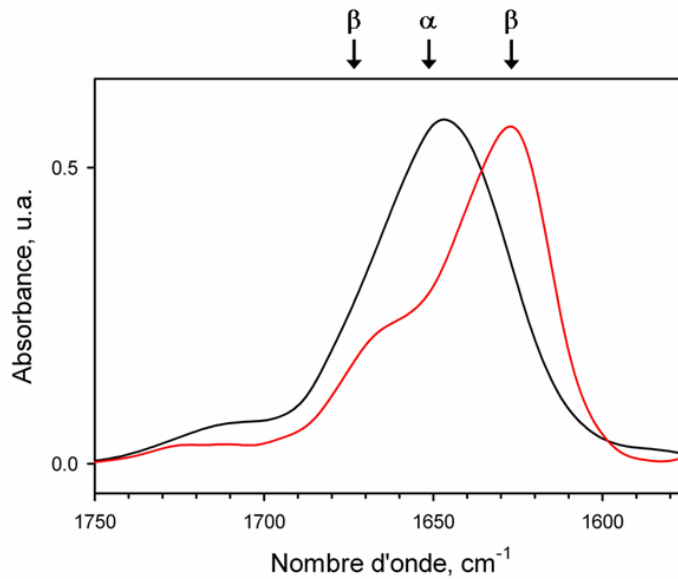
Des spectres d'absorption infrarouge sont enregistrés sur une solution de protéines dans l'état *molten globule* et sur une solution de fibres. Dans les deux cas, la concentration en protéine de ces solutions est de 20 mg/mL et les solutions sont dans le D<sub>2</sub>O.

La solution de protéines dans l'état *molten globule* est préparée juste avant l'expérience. La solution de fibres amyloïdes a été obtenue en plaçant 20 mg/mL de protéine dans une solution de D<sub>2</sub>O à pH 2. Le mélange est porté à 40°C et agité pendant plusieurs heures. Des fibres sont alors formées et peuvent être analysées en infrarouge.

Le spectromètre utilisé pour les mesures est un Jasco FT/IR615. L'absorption est mesurée dans la bande Amide I', entre 1500 et 1800 cm<sup>-1</sup>.

#### Résultats

Les spectres enregistrés sont représentés sur la figure IV.5. Comme expliqué précédemment, ces spectres sont composés de plusieurs bandes d'absorption qui correspondent chacune à un type de structure secondaire. Par exemple, les feuillets  $\beta$  absorbent entre 1621 et 1641 cm<sup>-1</sup> et entre 1671 et 1679 cm<sup>-1</sup> ; les hélices  $\alpha$  absorbent entre 1651 et 1657 cm<sup>-1</sup>.



**Fig IV.5 :** Spectre d'absorption infrarouge de l' $\alpha$ -lactalbumine dans l'état molten globule et dans l'état de fibre amyloïde.

— : Molten globule, — : Fibres amyloïdes

Les longueurs d'onde d'absorption des hélices  $\alpha$  et des feuilletts  $\beta$  sont indiquées au dessus de la figure.

Les spectres infrarouges sont différents pour l'état *molten globule* et les fibres. Ces deux états ne possèdent donc pas la même structure secondaire. Pour mieux caractériser les différences entre les deux structures, les longueurs d'onde d'absorption des hélices  $\alpha$  et des feuilletts  $\beta$  sont repérées sur le spectre.

Pour l'état *molten globule* (en noir), le spectre est principalement composé d'une large bande associée aux hélices  $\alpha$ . Une analyse plus détaillée du spectre infrarouge de l'état *molten globule* ainsi que les longueurs d'ondes associées aux différents types de structures secondaires sont accessibles dans l'article de Troullier et ses collaborateurs (ref 39).

Pour les fibres (en rouge), la bande d'absorption principale est décalée vers les faibles longueurs d'onde et une seconde bande apparaît aux alentours de  $1670\text{ cm}^{-1}$ . Ces longueurs d'onde sont caractéristiques des feuilletts  $\beta$ . La protéine sous forme de fibre possède donc moins d'hélices  $\alpha$ , mais plus de feuilletts  $\beta$  que la protéine dans l'état *molten globule* (43,2 %

d'hélices  $\alpha$  et 27,1 % de feuillets  $\beta$  pour le molten globule contre 20,3 % d'hélice  $\alpha$  et 60,8 % de feuillets  $\beta$  pour les fibres<sup>160</sup>).

Les expériences de spectroscopie infrarouge montrent que les protéines sous forme de fibres sont riches en feuillets  $\beta$ . Cette observation corrobore l'hypothèse de fibres amyloïdes caractérisées par leur structuration en feuillets  $\beta$  croisés.

## **IV.2. Effet du sel dans les solutions d' $\alpha$ -lactalbumine dans l'état**

### ***molten globule***

Les cinétiques de formation de fibres varient avec la concentration en NaCl. Pour déterminer l'origine des différences observées, l'effet du sel sur l'état initial de la réaction, le *molten globule* à pH 2, est étudié. Dans un premier temps, des expériences de dichroïsme circulaire sont réalisées. Les spectres dans l'UV lointain sont utilisés pour vérifier que le sel ne modifie pas la structure secondaire de la protéine. Les dénaturations thermiques suivies par dichroïsme circulaire permettent de mesurer la stabilité de la protéine dans l'état *molten globule* pour différentes concentrations en NaCl. Des expériences de diffusion de rayons X aux petits angles sont aussi effectuées sur les protéines dans l'état *molten globule* à pH 2 pour déterminer leurs tailles et mesurer les interactions qui ont lieu entre elles.

#### **IV.2.1 Dénaturation thermique du *molten globule***

Pour déterminer si le sel a un effet sur la stabilité de l'état *molten globule*, la dénaturation thermique de l' $\alpha$ -lactalbumine est suivie par dichroïsme circulaire en présence de différentes concentrations de sel.

##### **Conditions expérimentales**

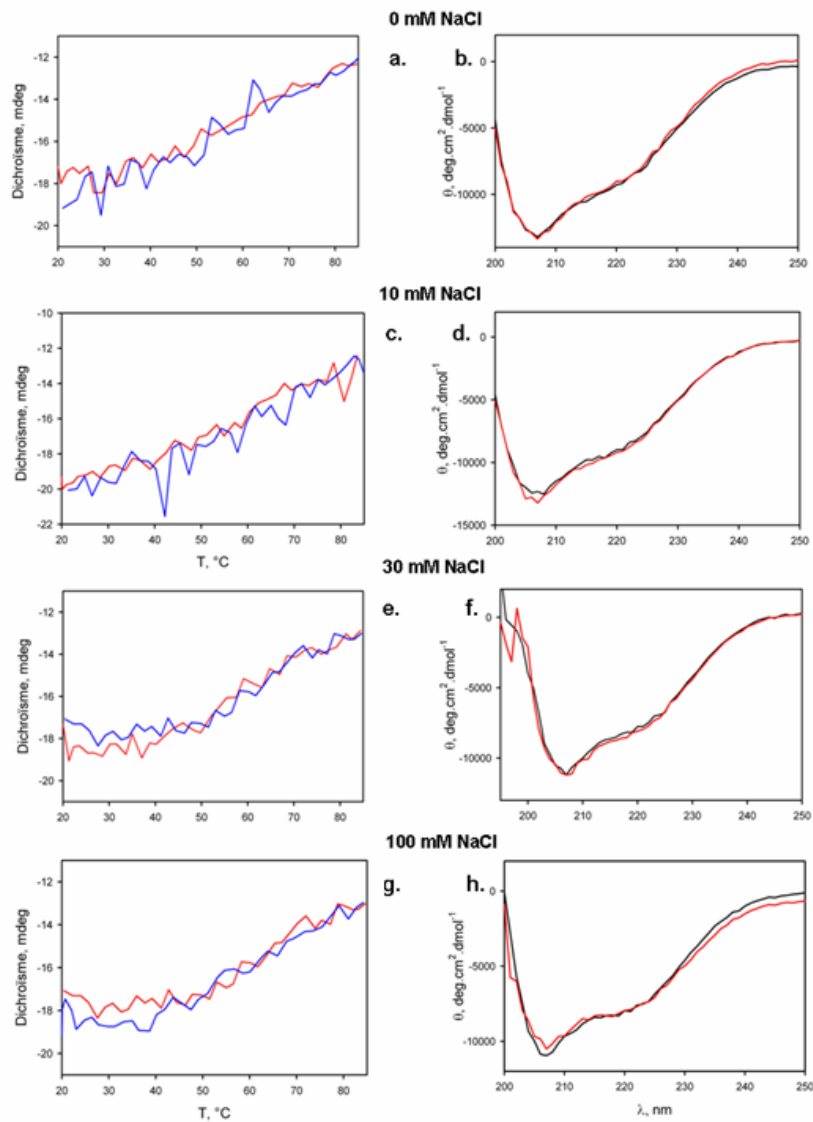
Les expériences sont effectuées sur un spectromètre CD J-810 (Jasco), dans des cuves en quartz de trajet optique 1 cm. Les dénaturations thermiques sont réalisées à l'aide d'un

appareil pellettier Biologic. La température varie entre 20 et 90°C à la vitesse de 2,5°C par minute. Les concentrations en sel utilisées pour ces expériences sont de 0, 10, 30 et 100mM. Les spectres dans l'UV lointain sont enregistrés entre 200 et 250 nm.

## Résultats

L'état *molten globule* ne possède pas de structure tertiaire stable, le spectre dans l'UV proche ne présente pas de signal. La dénaturation thermique de la protéine est donc suivie dans l'UV lointain, ce qui permet de suivre la dénaturation des structures secondaires.

La figure IV.6 présente les données enregistrées en dichroïsme circulaire. Cette figure est composée de huit parties. Les quatre parties à gauche correspondent aux résultats obtenus pour les expériences de dénaturation/renaturation. Les quatre parties à droite présentent les spectres de dichroïsme circulaire dans l'UV lointain avant et après dénaturation/renaturation. Chaque ligne correspond à une concentration en sel. Les deux schémas de la première ligne sont enregistrés en absence de sel. Les expériences de la deuxième ligne sont réalisées en présence de 10 mM de NaCl, puis 30 mM pour la troisième ligne et 100 mM pour la quatrième.



**Fig IV.6 : Expériences de dichroïsme circulaire effectuées sur le molten globule à pH2.**

Les expériences sont réalisées en présence de différentes concentrations en sel : 0mM (a. et b.), 10mM (c. et d.), 30mM (e. et f.) et 100mM (g. et h.). Les figures à gauche (a., c., e. et g.) présentent les expériences de dénaturation et renaturation. — : dénaturation, — : renaturation. Les figures à droites (b., d., f., et h) représentent les spectres de la protéine avant (—) et après (—) dénaturation-renaturation.

Intéressons-nous tout d'abord aux expériences réalisées en absence de sel. La structure secondaire est mesurée par dichroïsme circulaire dans l'UV lointain. Le spectre obtenu est représenté en noir dans la figure b. Le signal enregistré est négatif, il présente un minimum

aux alentours de 209 nm et un léger épaulement vers 222 nm. C'est un spectre caractéristique d'une structure riche en hélice  $\alpha$ . Lorsque la protéine perd sa structure, le spectre dans l'UV lointain prend une forme caractéristique des pelotes statistiques : il est nul entre 250 et 210 nm et devient négatif en dessous de 210 nm. La mesure de dichroïsme circulaire entre 250 et 210 nm permet donc de suivre la destruction des structures secondaires. Au fur et à mesure, que la protéine perd sa structure, le signal à ces longueurs d'onde va progressivement augmenter pour atteindre une valeur nulle lorsque la protéine est entièrement dénaturée.

La figure IV.6.a présente la mesure de dichroïsme circulaire à 222 nm pour différentes températures. Les mesures sont réalisées au cours de la dénaturation de la protéine (lorsque la température augmente, en rouge sur la figure) et au cours de la renaturation (lorsque la température diminue, en bleu sur la figure). A basse température, le signal de dichroïsme se situe aux alentours de -18 mdeg. Lorsque la température augmente, le signal augmente ; en particulier à partir de 25-30°C. Le signal continue d'augmenter au delà de 80°C, pour les plus hautes températures mesurées ; la dénaturation de la protéine semble donc être incomplète. De plus, pour ces températures, le dichroïsme circulaire est d'environ -12 mdeg ; il est loin d'être nul et cela confirme que la structure secondaire n'est pas totalement détruite. L'augmentation du signal et le fait que la protéine n'est pas complètement dénaturée à 85 °C indique que la dénaturation est faiblement coopérative. Lorsque la température redescend, le signal de dichroïsme circulaire décroît, indiquant que la protéine se restructure.

Malgré le bruit présent dans les résultats, les courbes en rouge et en bleu semblent superposées, ce qui indique que la protéine se renature correctement lorsque la température redescend. Pour vérifier cette hypothèse, un spectre de dichroïsme circulaire est mesuré dans l'UV lointain après renaturation. Il est représenté en rouge sur la figure VI.6.b. Ce spectre est très peu différent de celui enregistré avant la dénaturation (en noir) et permet de conclure que la protéine se renature correctement.



Les expériences sont aussi réalisées en présence de sel et les résultats sont représentés dans les figures IV.6.c à h.

Les spectres de dichroïsme circulaire sont tous semblables. Ils présentent un minimum vers 209 nm et un épaulement vers 222 nm, caractéristiques de protéines riches en hélices  $\alpha$ . Les spectres enregistrés après renaturation sont identiques au bruit près à ceux enregistrés avant la dénaturation et ceci quelque soit la concentration en sel.

En ce qui concerne la dénaturation/renaturation suivie par dichroïsme circulaire, les données sont semblables pour les différentes concentrations de sel. A 20°C, le signal est d'environ -18 mdeg ; il augmente avec la température à partir d'environ 30°C et continue d'augmenter à 80°C. Pour ces hautes températures, le signal est d'environ -12 mdeg, la protéine n'a donc pas perdu la totalité de sa structure secondaire.

Les expériences de dénaturation permettent d'estimer la stabilité de la protéine. La manière avec laquelle les structures sont détruites en fonction de la température est caractéristique de cette stabilité. Dans le cas présent, les dénaturations thermiques sont identiques quelque soit la concentration en sel ; il n'influence donc pas la stabilité de l'état *molten globule*.

Ainsi, les spectres de dichroïsme circulaire et les dénaturations thermiques sont identiques quelle que soit la concentration en sel. Le sel n'a donc pas d'effet notable sur les protéines considérées comme isolées. Pour expliquer son influence sur les cinétiques de formation de fibres, il faut étudier son effet sur les interactions entre protéines.

La diffusion de rayons X aux petits angles (SAXS pour *small angle X-ray scattering*) permet de mesurer les rayons de giration des protéines et leurs facteurs d'interactions. Ce facteur caractérise les interactions entre protéines et le second coefficient du Viriel, utilisé pour quantifier ces interactions, peut en être déduit.

Avant de s'intéresser aux résultats obtenus en SAXS, il est intéressant de revenir sur les différents types d'interactions qui peuvent exister entre les protéines.

#### **IV.2.2. Interactions protéine-protéine**

Les protéines sont des objets complexes et leurs surfaces sont constituées d'acides aminés qui peuvent être chargés ou non et qui interagissent avec le solvant. Des forces de van der Waals et des répulsions électrostatiques peuvent exister entre les protéines ; des forces d'hydratation et hydrophobique jouent aussi un rôle dans leur solubilité. Il est donc difficile de construire un modèle qui traduise parfaitement toutes les forces qui existent entre elles.

Les interactions entre particules sont traditionnellement décrites à l'aide d'un potentiel d'interaction. Ce potentiel est défini à partir d'un système constitué de deux particules éloignées d'une distance  $x$ . En faisant varier  $x$  et en mesurant l'énergie du système, une courbe peut être tracée. Cette courbe est le potentiel d'interaction. Il indique si les forces sont plutôt répulsives ou attractives, et à quelle distance elles sont ressenties. Le potentiel d'interaction le mieux compris et le plus utilisé pour les suspensions colloïdales est le potentiel DLVO<sup>165,166</sup>.

##### Théorie DLVO

La théorie DLVO sur les interactions entre particules d'un colloïde explique la stabilité colloïdale en terme d'équilibre entre les interactions de van der Waals attractives et des répulsions électrostatiques.

Le potentiel d'interaction entre deux particules est défini dans le cadre de la théorie DLVO comme la somme de trois potentiels :

- le potentiel de répulsion sphère dure (de volume exclu),
- le potentiel électrostatique,
- le potentiel de van der Waals.

Le potentiel de répulsion sphère dure traduit le fait que des protéines ne peuvent pas se superposer. Lorsqu'elles s'approchent trop l'une de l'autre, les nuages électroniques des protéines se superposent. Il en résulte une forte répulsion qui les empêche de se rapprocher davantage. Vis-à-vis de cet effet, les protéines sont considérées comme des sphères dures.

Les protéines possèdent des charges électrostatiques et les interactions qui en résultent sont prises en compte dans le potentiel électrostatique.

Le potentiel de van der Waals provient de l'attraction qui existe entre deux dipôles. Ces dipôles peuvent être permanents (présents dans la molécule) ou induits (lorsque les molécules sont polarisables).

Ce modèle DLVO, initialement développé pour des solutions de colloïdes ne prend pas en compte les forces hydrophobes. Ces forces deviennent importantes lorsque la distance entre les protéines est de quelques nanomètres. Dans ce cas, le modèle DLVO ne permet pas de décrire les interactions entre les protéines.

### Forces hydrophobes

Les forces hydrophobes prennent de l'importance pour les interactions entre les protéines lorsque ces protéines sont distantes de quelques nanomètres<sup>167</sup>. Pour vérifier la pertinence de ces forces dans notre cas, la distance entre les protéines est calculée.

Soit  $c$  la concentration en mg/mL de la solution de protéines. Le nombre de protéines par litre est égal à  $c \cdot \mathcal{N} / M$  avec  $\mathcal{N}$  le nombre d'Avogadro et  $M$  la masse molaire de la protéine. Or, un litre de solution occupe  $10^{-3} \text{ m}^3$ . Le volume moyen par protéine est de  $10^{-3} \cdot M / c \cdot \mathcal{N}$ , et la distance moyenne entre les centres des protéines est donc de  $(10^{-3} \cdot M / c \cdot \mathcal{N})^{1/3}$  soit :

- 13 nm pour une concentration de 10 mg/mL,
- 10,5 nm pour 20 mg/mL,
- 8,4 nm pour 40 mg/mL, et
- 7,3 nm pour 60 mg/mL.

Ces distances sont celles entre les centres des protéines. Sachant que les protéines ont un rayon de giration proche de 2 nm, il faut soustraire 4 nm à chacune des distances calculées précédemment pour obtenir les distances entre les bords des protéines. Elles varient dans ce cas, de 9 nm pour une concentration de 10 mg/mL à 3,3 nm pour une concentration de 60 mg/mL. La distance entre les protéines est donc de quelques nanomètres, et les forces hydrophobes ne peuvent plus être négligées dans la description des interactions<sup>167</sup>.

La partie hydrophobe de l' $\alpha$ -lactalbumine a été mise en évidence dans l'étude de la séquence de la protéine présentée dans la partie précédente. Cette portion de la séquence est difficile à hydrater, elle aura tendance à « se cacher » de l'eau ; dans la structure native, elle est enfouie à l'intérieur de la protéine. D'autres stratégies peuvent être adoptées pour isoler des parties hydrophobes de l'eau. Les lipides par exemple, sont des molécules constituées d'une tête hydrophile et d'une queue hydrophobe. Pour isoler leur partie hydrophobe, certains lipides se rassemblent sous forme de micelles sphériques en mettant en contact leur queue hydrophobe et en présentant leur tête hydrophile au solvant. D'autres lipides peuvent aussi former des structures plus grandes comme des bicouches lipidiques, les agencements étant fonction de la géométrie de la molécule lipidique envisagée.

A pH 2, l' $\alpha$ -lactalbumine est dans sa forme *molten globule* et ne possède pas de structure tertiaire stable, ce qui l'empêche de séquestrer efficacement ses régions hydrophobes. Il est donc raisonnable de penser que la protéine adopte une stratégie semblable à celle des lipides pour répondre aux contraintes hydrophobes<sup>168</sup>. Les protéines dans l'état *molten globule* mettraient alors leurs parties hydrophobes en contact les unes des autres pour les isoler du solvant.

#### Forces prises en compte pour l'interprétation des résultats

L'interaction entre les protéines dépend de la concentration en sel. L'interaction électrostatique entre les protéines doit donc être prise en compte.

Les interactions de van der Waals et hydrophobes peuvent, toutes deux, être utilisées. Les interactions de van der Waals sont difficilement accessibles et n'apportent pas d'informations directes. Les interactions hydrophobes sont importantes car dans les solutions concentrées, les protéines sont proches les unes des autres. De plus, ces interactions permettent d'expliquer le fait que la protéine ne peut former des fibres qu'à partir de son *molten globule*, lorsque sa portion hydrophobe est exposée.

Les deux forces utilisées pour interpréter les résultats sont donc la répulsion électrostatique et la « force » hydrophobe.

### **IV.2.3. Diffusion de rayons X aux petits angles**

#### Introduction

Les rayons X sont, depuis des années, utilisés pour déterminer les structures des macromolécules du vivant<sup>9,10</sup>. C'est un rayonnement électromagnétique à haute fréquence, donc haute énergie et faible longueur d'onde. Ces caractéristiques lui permettent d'interagir avec les électrons. Du fait de leur faible masse, les électrons (de la protéine dans notre cas) peuvent diffuser élastiquement, c'est à dire sans échange d'énergie, les photons qui composent le rayon X.

Au cours d'une expérience de diffusion de rayons X aux petits angles, une solution de protéine est placée dans le faisceau de rayons X. Derrière l'échantillon, à une distance variable, se trouve des capteurs CCD qui permettent de compter les photons. Ce dispositif permet d'obtenir des figures de diffusion.

L'échantillon est isotrope, la protéine ne privilégie pas de direction, d'orientation particulière. Les photons mesurés par les capteurs sont donc caractérisés par l'angle  $\theta$  avec lequel ils sont déviés par la solution de protéine. Sur les courbes de diffusion, l'intensité du faisceau diffusé est représenté en fonction du vecteur de diffusion  $s=2\sin\theta/\lambda$ . (la forme normalisée  $Q=4\pi\sin\theta/\lambda=2\pi s$  est aussi utilisée).

Les courbes obtenues donnent des informations sur la forme et les interactions entre protéines. La courbe de diffusion peut être décomposée en un produit de deux fonctions :

$$I(c, s) = I(0, s) \cdot S(c, s)$$

Avec  $I(0, s)$ , appelé facteur de forme, qui correspond à la courbe de diffusion de la protéine en solution idéale, c'est-à-dire sans interaction, et  $S(c, s)$ , appelée facteur de structure, dû aux interactions entre protéines.

Les facteurs de forme des protéines sont obtenus à partir des profils de diffusion enregistrés à faible concentration en protéine. Dans ce cas, la distance moyenne entre les molécules est

grande et les interactions peuvent être négligées ; le facteur de structure est considéré comme égal à 1 et les courbes de diffusion enregistrées correspondent au facteur de forme. Il peut ensuite être analysé pour déterminer le rayon de giration de la protéine.

La zone de Guinier correspond aux vecteurs de diffusion  $Q$  pour lesquels le produit du rayon de giration avec  $Q$  est inférieur à 1,3. Dans cette zone, l'intensité diffusée par la particule est égale à  $I_0 \cdot \exp(-1/3 \cdot R_g^2 \cdot Q^2)$  avec  $I_0$  l'intensité diffusée à l'origine. Ainsi, en se plaçant dans la zone de Guinier et en utilisant l'équation précédente, il est possible de calculer le rayon de giration ( $R_g$ ) des protéines en suspension.

Le facteur de structure renseigne sur les interactions entre particules : lorsque la concentration en protéine est élevée, la figure de diffusion obtenue tient compte de la forme de ces macromolécules, mais aussi des interactions qui existent entre elles. Le facteur de structure est calculé en divisant les courbes de diffusion d'une solution concentrée en protéine (donc qui tient compte des interactions entre molécules) par le facteur de forme (qui dépend uniquement de leur forme). Cette opération permet de faire apparaître les effets des interactions entre les protéines.

Le facteur de structure  $S$  ainsi obtenu renseigne sur les interactions entre protéines et peut notamment être utilisé pour calculer le second coefficient du Viriel ( $A_2$ ). Ce paramètre permet de quantifier les interactions entre les particules. Il est relié à la pression osmotique ( $\Pi$ ) par la relation suivante :

$$\frac{\Pi}{cRT} = \frac{1}{M} + A_2c$$

Avec  $M$  la masse molaire de la protéine et  $c$  sa concentration.

Un coefficient du Viriel négatif correspond à une interaction attractive, un coefficient du Viriel positif correspond à une interaction répulsive.

La valeur du facteur de structure en 0 est reliée à la pression osmotique :

$$S(c,0) = (RT/M)(\partial\Pi/\partial c)$$

Le coefficient du Viriel est donc déduit de la valeur du facteur de structure en 0 :

$$\frac{1}{S(c,0)} = 1 + 2MA_2c$$

### Conditions expérimentales

Les expériences de diffusion ont été réalisées sur la ligne ID02 de l'ESRF, et sur des solutions de protéine dont la concentration varie entre 1 et 40 mg/mL. Les solutions sont ajustées à pH 2 et quatre concentrations en NaCl sont utilisées (0, 30, 60 et 100 mM).

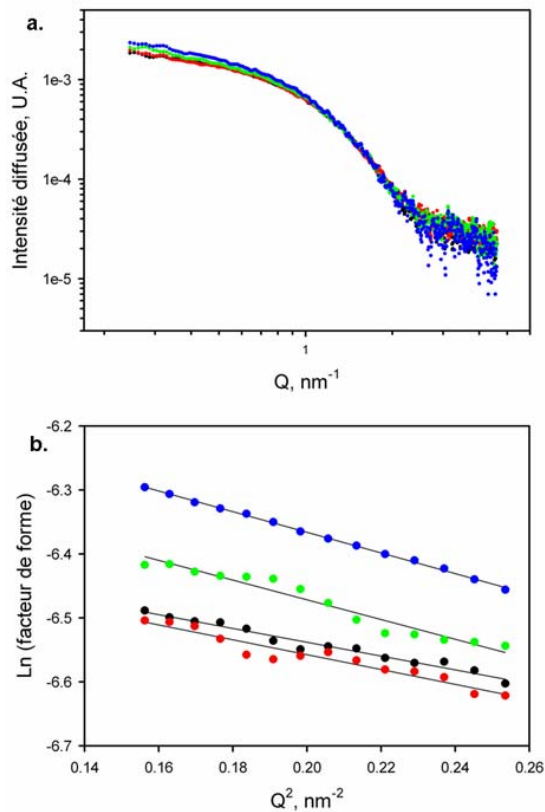
Chaque courbe est ensuite traitée en soustrayant la diffusion due au tampon et en normalisant à la concentration en protéines.

### Résultats

Les courbes de diffusion enregistrées à basse concentration en protéine (2,5 mg/mL) sont représentées sur la figure IV.7a. L'intensité diffusée est représentée en fonction du vecteur de diffusion Q. Ces courbes sont typiques de courbes de diffusion de rayons X par des protéines : une forte intensité aux faibles Q qui diminue lorsque Q augmente. Les quatre courbes de la figure correspondent aux résultats obtenus en présence de différentes concentrations de sel. La



force ionique modifie peu la figure de diffusion obtenue. Pour les faibles concentrations (0 mM NaCl en noir et 30 mM en rouge), les courbes sont quasiment superposées. Lorsqu'elle augmente (60 mM en vert puis 100 mM en bleu), l'intensité diffusée semble augmenter, en particulier pour les petits angles.



**Fig IV.7 : Facteur de forme de l' $\alpha$ -lactalbumine en présence de différentes concentrations en sel.**

a. Facteur de forme  
b.  $\ln(I)$  en fonction de  $Q^2$ , estimation du rayon de giration par regression linéaire.

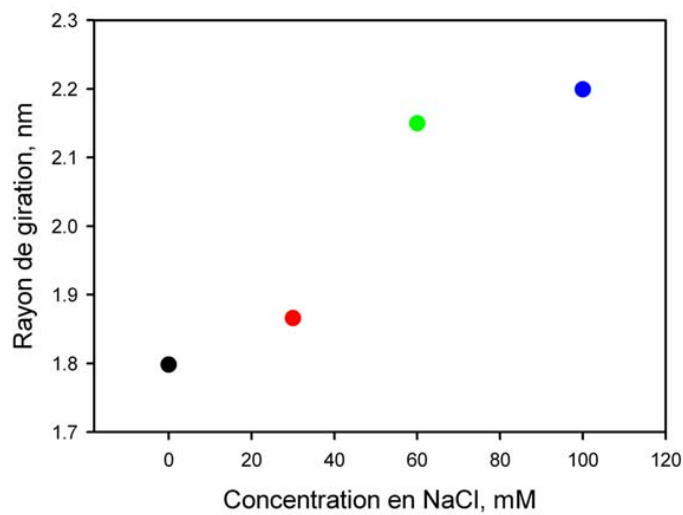
Les concentrations en sel utilisées sont: (●) 0mM, (●) 30mM, (●) 60mM, (●) 100mM

Ces courbes sont enregistrées à basse concentration en protéines et les interactions sont négligées, pour pouvoir assimiler ces courbes au facteur de forme de la protéine.

Les données issues de ces expériences sont utilisées pour calculer les rayons de giration. Le logarithme de l'intensité diffusée est représenté en fonction de  $Q^2$  dans la figure IV.7b. Les données sont représentées uniquement pour les Q qui se situent dans la zone de Guinier. Les quatre séries de points correspondent aux quatre concentrations en NaCl utilisées. Les données sont un peu bruitées, mais la relation entre les  $\ln(I)$  et  $Q^2$  est clairement linéaire. Comme expliqué ci-dessus, l'intensité diffusée pour les faible valeurs de Q (dans la zone de Guinier) est égale à  $I_0 \cdot \exp(-1/3 \cdot R_g^2 \cdot Q^2)$  pour le facteur de forme. Les pentes des droites

formées par les points de la figure IV.7b sont donc égales à  $-1/3$  de  $R_g^2$ , avec  $R_g$  le rayon de giration de la protéine pour les différentes concentrations de sel.

Les rayons de giration ainsi obtenus sont reportés dans la figure IV.8. Ces rayons sont compris entre 18 et 22 nm et ils semblent augmenter avec la concentration en sel. Ce point sera discuté dans la suite.



**Fig IV.8 :** Rayons de giration de l' $\alpha$ -lactalbumine dans l'état molten globule.

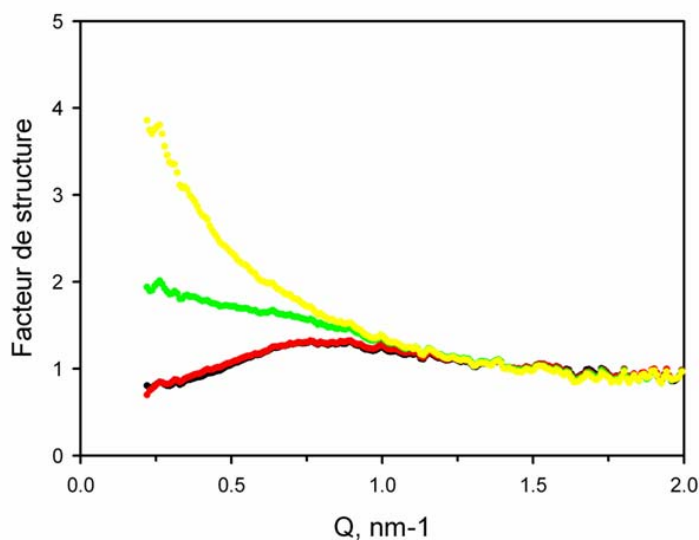
Les rayons de giration sont déterminés pour différentes concentrations en sel à partir des expériences de diffusion de SAXS.

La courbe de diffusion enregistrée à 2,5 mg/mL est assimilée au facteur de forme de l' $\alpha$ -lactalbumine. Ceci permet de déterminer les facteurs de structure en divisant les courbes enregistrées à plus hautes concentrations en protéine par ce facteur de forme. Les courbes de diffusion enregistrées à 5, 10, 20 et 40 mg/mL sont donc divisées par la courbe de diffusion enregistrée à 2,5 mg/mL, et ce pour toutes les concentrations en sel.

A concentration en sel fixe, tous les facteurs de structures calculés à partir des différentes courbes de diffusion ont la même forme globale. Cependant, leurs amplitudes dépendent de la

concentration en protéine. En effet, plus cette dernière est élevée, plus les interactions entre protéines sont importantes et plus le facteur de structure aura une amplitude forte.

Lorsque la concentration en sel varie, le facteur de structure change de forme. Cet effet est représenté dans la figure IV.9. Les facteurs de structures présentés dans cette figure ont été calculés à partir des solutions à 40 mg/mL car c'est ceux pour lesquels l'effet du sel est le plus visible.



**Fig IV.9 :** Facteurs de structures déterminés par diffusion de rayons X aux petits angles

Les facteurs de structures représentés sont calculés à partir des courbes de diffusion des solutions à 40 mg/mL.

Plusieurs concentrations en sel sont utilisées: (●) 0mM, (●) 30mM, (●) 60mM, (●) 100mM

Les facteurs de structure pour les faibles concentrations en sel (noir : 0 mM, rouge : 30 mM) sont inférieures à 1 pour les faibles valeurs de Q. Ils augmentent jusqu'à  $Q = 0,8$  pour atteindre une valeur proche de 1,3 puis ils décroissent lentement pour revenir à 1. Pour les hautes concentrations en sel (vert : 60 mM, jaune : 100mM), la valeur du facteur de structure en 0 est supérieure à 1 (environ 2 pour 60 mM et 5 pour 100 mM de NaCl). Cette valeur décroît et tend vers 1 pour les grands Q.

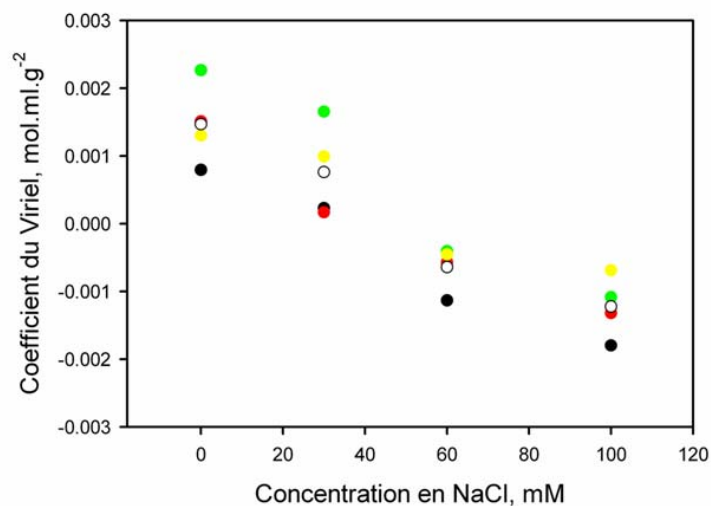
Les facteurs de structures  $S$  obtenus permettent de calculer les coefficients du Viriel  $A_2$  grâce à la relation suivante :

$$\frac{1}{S(c,0)} = 1 + 2MA_2c$$

Avec  $M$  la masse molaire de la protéine et  $c$  sa concentration.

Lorsque la valeur du facteur de structure en 0 est supérieure à 1, les interactions entre protéines sont de nature attractive et le coefficient du Viriel est négatif. Lorsque la valeur du facteur de structure en 0 est inférieure à 1, les interactions sont répulsives et le coefficient du Viriel est positif.

Ces coefficients sont calculés à partir des facteurs de structures obtenues pour toutes les concentrations en sel et en protéine. Ils sont représentés sur la figure IV.10.



**Fig IV.10 : Second coefficient du Viriel déterminés à partir des facteurs de structure.**

Les coefficients du Viriel sont déterminés à partir des expériences de SAXS effectuées sur les solutions à (●) 5 mg/mL, (●) 10 mg/mL, (●) 20 mg/mL et (●) 40 mg/mL. La moyenne des coefficients obtenus pour ces différentes concentrations est représentée par le symbole ○.

Les différentes couleurs représentent les différentes concentrations en protéine : noir = 5 mg/mL, rouge = 10 mg/mL, vert = 20 mg/mL et jaune 40 mg/mL. La moyenne des coefficients du Viriel obtenus à partir des solutions à différentes concentrations est représentée par des cercles vides. Les données sont relativement dispersées, mais on peut tout de même observer une diminution du coefficient du Viriel avec la concentration ionique. Lorsque la concentration en sel augmente, le coefficient du Viriel diminue et les interactions entre les protéines deviennent de plus en plus attractives.

## Discussion

### *Rayon de giration*

Kuwajima et ses collaborateurs ont calculé un rayon de giration de 17,1 Å pour le *molten globule* de l' $\alpha$ -lactalbumine. A faible concentration en sel, les valeurs calculées ici sont très légèrement supérieures (18 Å).

De plus, lorsque la concentration en sel augmente, le rayon de giration calculé est de plus en plus grand. Deux explications peuvent être avancées. L'augmentation de sel peut déstabiliser le *molten globule* et augmenter son rayon de giration. Cette augmentation du rayon de giration peut aussi être expliquée en considérant que les interactions entre particules ne sont pas négligeables pour une concentration en protéine égale à 2,5 mg/mL.

La première explication peut être écartée grâce aux dénaturations thermiques suivies par dichroïsme circulaire. Ces expériences ont montré que la concentration en sel n'influence pas la stabilité de la protéine. Le rayon de giration de la protéine ne devrait donc pas varier avec la concentration en sel.

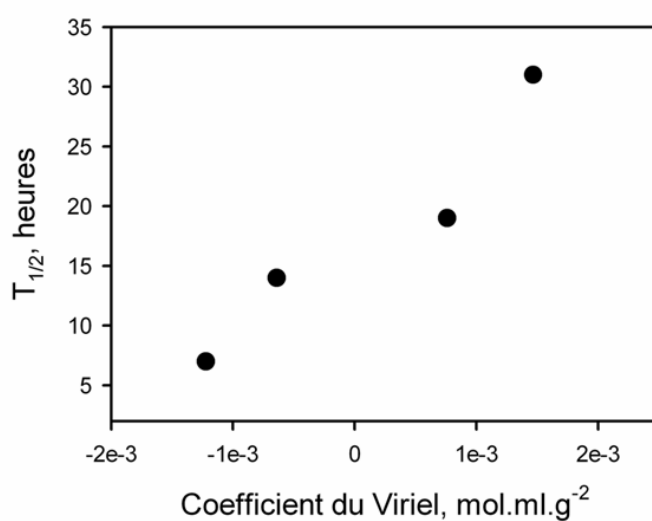
Pour calculer ce rayon, la solution de protéine à 2,5 mg/mL est utilisée. La courbe de diffusion obtenue à partir de cette solution est assimilée au facteur de forme, c'est-à-dire que les interactions entre les protéines sont considérées comme nulles. Le rayon de giration calculé par cette méthode n'étant pas constant, les interactions ne sont vraisemblablement pas négligeables.

La figure IV.9 représente les facteurs de structure pour différentes concentrations en sel. Le facteur de structure est calculé à partir du rapport de la courbe de diffusion sur le facteur de forme. Lorsque la concentration en sel augmente, la pente du facteur de structure à l'origine augmente. Comme le facteur de forme est constant, cela indique que la pente de la courbe de diffusion à l'origine augmente avec la concentration en sel. Cette augmentation est due à des changements dans les interactions entre protéines. Ceci induit une augmentation du rayon de giration apparent. Ainsi, si les interactions entre particules ne sont pas négligeables pour des concentrations en protéine égales à 2,5 mg/mL, le rayon de giration calculé sur ces courbes est légèrement surévalué et a tendance à augmenter avec la concentration en sel.

#### *Facteur de structure, coefficient du Viriel*

Le coefficient du Viriel calculé ici diminue clairement avec la concentration en sel, les protéines ont donc tendance à s'attirer davantage lorsque la concentration en sel augmente. Ce phénomène peut être expliqué en revenant sur les forces qui existent entre les protéines. Les deux principaux types d'interactions entre les protéines en solution sont une répulsion électrostatique due à la charge des protéines et une attraction due à leurs parties hydrophobes. Lorsque la concentration en sel augmente, la force ionique de la solution augmente, et les charges électrostatiques sont écrantées. La répulsion électrostatique diminue, il ne reste plus que l'attraction due aux forces hydrophobes.

Sur la figure IV.11, les temps de demi-réaction mesurés dans les expériences de cinétiques suivies par fluorescence de la thioflavine T sont représentés en fonction du coefficient du Viriel déterminé par les expériences de diffusion de rayons X aux petits angles. La figure montre bien une relation entre ces deux paramètres, soulignant ainsi l'importance des interactions entre protéines pour la cinétique de formation de fibres amyloïdes.



**Fig IV.11 :** Temps de demi-réaction de la formation de fibres en fonction du second coefficient du Viriel calculé par diffusion de rayons X.

### **IV.3. Rhéologie**

A haute concentration en protéine, la formation de fibres s'accompagne d'une augmentation de viscosité de la solution. Lorsque les fibres sont formées, elles peuvent même interagir entre elles pour former un gel. Cette augmentation de viscosité est suivie par rhéométrie.

#### **Conditions expérimentales**

Le rhéomètre décrit dans la partie précédente a été utilisé pour suivre la formation de fibres d' $\alpha$ -lactalbumine. Les expériences sont réalisées sur une solution de protéine à 40 mg/mL à pH 2. La température est de 40°C.

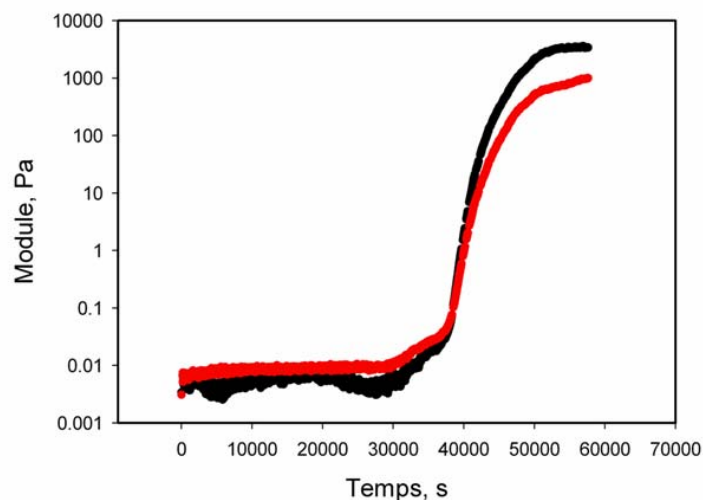
Dans ces expériences, il est difficile d'empêcher l'évaporation de la solution. Pour pouvoir détecter l'augmentation de viscosité due à la formation de fibres avant que la solution ne s'évapore, la réaction doit être rapide, c'est pourquoi une concentration de 150 mM de NaCl est utilisée. La cinétique est suivie pendant 16 h.

Pour les mesures réalisées ici, le rhéomètre est utilisé en mode « oscillation ». La partie mobile oscille autour d'un point d'équilibre. La fréquence d'oscillation est constante et égale à 1 Hz ( $2\pi$  rad/s). La contrainte (le taux de cisaillement) appliquée sur le fluide est elle aussi constante et égale à  $0,03 \text{ s}^{-1}$ . Elle est assez faible pour ne pas trop perturber le système. La réponse de l'échantillon à ces sollicitations permet de déterminer son module élastique et son module visqueux. Le module élastique traduit la capacité de l'échantillon à restituer l'énergie mécanique qui le déforme. Le module visqueux traduit la capacité de l'échantillon à dissiper irréversiblement (sous forme de chaleur) cette énergie. Ces modules sont très faibles pour un liquide, mais ils augmentent sensiblement lorsque l'échantillon est un gel.



## Résultats

La figure IV.12 montre l'évolution des modules élastique et visqueux au cours du temps. Ces modules sont faibles et constants pendant une dizaine d'heures. Ils augmentent ensuite fortement pendant environ 5 h, pour atteindre un plateau en fin de réaction.

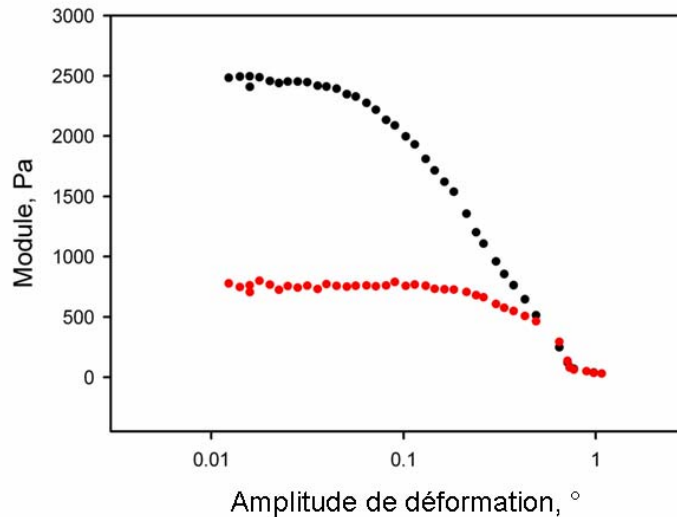


**Fig IV.12 :** Modules élastique et visqueux mesurés par rhéométrie pendant la formation de fibres amyloïdes.

- : Module visqueux ( $G'$ )
- : Module élastique ( $G''$ )

La cinétique enregistrée est similaire aux cinétiques de formation de fibres. En début de réaction, pendant la phase de nucléation, la viscosité de la solution ne varie pas et les modules élastique et visqueux restent constants. Pendant la phase de croissance, la solution devient de plus en plus visqueuse et prend progressivement la consistance d'un gel. Les modules augmentent. Lorsque toutes les fibres sont formées, il n'y a plus de monomères libres, la réaction est terminée et les modules mécaniques varient peu et sont caractéristiques du gel formé.

Pour caractériser ce gel, les modules visqueux et élastique sont mesurés en appliquant différentes contraintes sur l'échantillon. Les résultats obtenus sont représentés sur la figure IV.13.



**Fig IV.13 :** Modules élastique et visqueux du gel de protéine en fonction du taux de cisaillement.

- : Module visqueux (G')
- : Module élastique (G'')

Les modules sont constants jusqu'à une contrainte de cisaillement d'environ  $0,1 \text{ s}^{-1}$ . L'intensité des modules décroît pour des constantes supérieures. Ceci est dû à la fluidification du gel. Le gel a donc un comportement rhéofluidifiant, il devient moins visqueux lorsqu'il est soumis à de fortes contraintes.

### Discussion

Au premier ordre, la viscosité d'une suspension de particules est égale à :

$$\eta = \eta_0(1 + \nu_1 \cdot \Phi + \dots)$$

Avec  $\eta$  la viscosité de la solution,  $\eta_0$  la viscosité du solvant,  $\nu_l$  le facteur de Simha dépendant de la forme des particules en suspension et  $\Phi$  le volume spécifique occupé par ces particules.

Dans le cas de la formation de fibres,  $\Phi$  varie peu. La taille des particules en solution augmente au cours de la formation de fibres, mais leur nombre diminue. Le volume partiel occupé par les particules varie donc peu. En revanche, le facteur de Simha dépend de la forme de la particule. Plus la particule est allongée, plus son facteur de Simha est important. Cette augmentation explique en partie l'augmentation de la viscosité.

Un autre phénomène a davantage d'effet sur les propriétés viscoélastiques de la solution et doit être considéré pour expliquer l'augmentation de viscosité. Aux concentrations importantes, la solution prend la forme d'un gel, et la formule de viscosité applicable aux suspensions de particules ne peut raisonnablement pas être utilisée. Les fibres formées interagissent entre elles, ce qui explique la consistance des échantillons, leur forte viscosité. La présence de ces interactions est visible lorsque le gel est caractérisé en utilisant différents taux de cisaillements. Pour des faibles taux de cisaillements, les modules élastiques et visqueux sont importants. Pour des forts taux de cisaillements, ces modules décroissent. Cependant, les fibres ne sont pas détruites : la solution a toujours un aspect trouble et gélatineux. Lorsqu'on laisse le système se relaxer, il retrouve rapidement ses caractéristiques viscoélastiques. Les modules élastique et visqueux dépendent des interactions entre les fibres. Lorsque le cisaillement est faible, ces interactions entre fibres sont peu perturbées et le gel est solide. Lorsque le cisaillement augmente, l'énergie mécanique donnée à l'échantillon par le rhéomètre perturbe suffisamment ces interactions pour fluidifier le gel.

#### **IV.4. Résumé - discussion**

Les fibres d' $\alpha$ -lactalbumine sont formées à partir de l'état *molten globule* à pH 2, à 40°C et en présence de différentes concentrations de sel.

Les expériences de microscopie confirment la formation de fibres amyloïdes. Des filaments aux bords sombres, d'un diamètre de 5 à 10 nm et quelques  $\mu\text{m}$  de longueur sont visibles sur les clichés. Ces caractéristiques sont typiques de fibres amyloïdes. Les spectres infrarouges confirment la présence d'une grande quantité de feuillets  $\beta$ .

Les cinétiques de formation de fibres présentent plusieurs phases : tout d'abord une phase pendant laquelle le signal n'évolue pas, la phase de latence qui correspond à la formation de « noyaux » de fibres amyloïdes ; dans un deuxième temps, l'intensité de fluorescence augmente, cette phase correspond à l'élongation des fibres par adjonction de nouveaux monomères sur les fibres et les noyaux déjà formés, elle est appelée phase de croissance ; en fin de réaction, le signal n'évolue plus et reste à une haute intensité de fluorescence, la solution est composée de fibres qui ne grandissent plus.

Les cinétiques de formation et la durée des phases varient avec la concentration en sel. Plus la concentration en sel est élevée, plus la réaction est rapide.

Pour déterminer l'origine de ces différences, des expériences ont été menées sur l'état initial de la réaction, à savoir l'état *molten globule*.

Des expériences de dénaturation thermique sont effectuées en présence de différentes concentrations de sel. Les résultats obtenus sont semblables quelle que soit la concentration utilisée. La stabilité ne dépend donc pas de la concentration en sel. Les différences observées sur les cinétiques ne peuvent pas s'expliquer par une déstabilisation du *molten globule* et les explications sont à chercher du côté des interactions entre protéines.

Pour mesurer ces interactions, des expériences de diffusion de rayons X aux petits angles sont réalisées. Ces expériences permettent de mesurer les rayons de giration des protéines, et aussi d'estimer les coefficients du Viriel qui caractérisent les interactions entre protéines. Ces coefficients varient beaucoup avec la concentration en sel et indiquent que les interactions sont de plus en plus attractives lorsque la concentration en sel augmente. Cette variation du coefficient du Viriel sous l'influence de sel est bien liée à la vitesse de formation de fibres. C'est clairement visible sur la figure IV.11 qui représente ces coefficients en fonction du temps de demi-réaction de formation de fibres.

Comment expliquer les modifications dans les interactions entre protéines, et donc les changements dans les cinétiques avec la variation de la concentration en sel ? Les deux principales interactions entre protéines sont l'interaction attractive hydrophobe et l'interaction électrostatique qui est répulsive car les protéines présentes en solution possèdent toutes la même charge globale. Lorsque la concentration en sel augmente, la force ionique de la solution augmente et les charges des protéines sont écrantées. Les interactions électrostatiques sont donc moins importantes. L'équilibre entre les attractions hydrophobes et les répulsions électrostatiques est perturbé, et les protéines se rencontrent plus facilement.

Enfin, les fibres à hautes concentrations interagissent entre elles pour former un gel. La cinétique de formation de fibres peut donc être suivie par rhéométrie. C'est ce qui a été réalisé pour des fortes concentrations en sel. Le gel formé est lui aussi caractérisé par rhéométrie ce qui met en évidence le caractère rhéofluidifiant de l'échantillon.

# **V. Conclusion, perspectives**

## $\alpha$ -lactalbumine et métaux

### Discussion

La première partie est consacrée à l'étude de l'influence des métaux sur la constante de vitesse de repliement et la stabilité de l' $\alpha$ -lactalbumine. Les résultats obtenus sont interprétés à l'aide du paysage d'énergie de la protéine.

Pour expliquer l'évolution de la constante de vitesse de repliement en fonction de la concentration en calcium, il est intéressant de revenir sur la définition des états natifs et *molten globule*. Ces macro états sont accessibles expérimentalement et sont définis par rapport à leurs caractéristiques structurales. L'état *molten globule* possède la structure secondaire de la protéine native mais pas de structure tertiaire stable. L'état natif possède de la structure secondaire et de la structure tertiaire. A l'échelle atomique, ces macroétats sont constitués d'une multitude de microétats. Le nombre de ces microétats correspond à l'entropie du macroétat.

Lorsque du calcium est introduit en solution, le paysage d'énergie est modifié. Les expériences de cinétiques indiquent que le repliement est accéléré car le calcium se lie aux états *molten globule* et de transition. L'effet de cette liaison est principalement de nature entropique. Parmi les microétats qui constituent le *molten globule* en absence d'ion métallique, certains ont des structures qui leur permettent de lier un calcium. La liaison de cet ion sur ces microétats diminuera leurs énergies internes et ils seront stabilisés, en particulier par rapport aux microétats du *molten globule* apo qui ne peuvent pas lier d'ion. Ainsi, en présence de calcium, seul les microétats susceptibles de lier le calcium pourront être peuplés et constitueront l'état *molten globule*. Les macroétats *molten globule* de la forme holo et apo ne sont donc pas constitués par les mêmes microétats. Ils sont moins nombreux pour la forme

holo que pour la forme apo, ce qui explique la diminution d'entropie du *molten globule* en présence de calcium et l'accélération de la réaction de repliement.

Les expériences de dénaturation thermique peuvent également être expliquées à l'aide de cette description. Notamment pour l'augmentation de température de dénaturation avec la concentration en calcium. Cette augmentation est attribuée à une augmentation d'entropie de l'état natif car l'enthalpie de la réaction est supposée constante. Or les valeurs de l'enthalpie calorimétrique et de l'enthalpie de Van't Hoff diffèrent. Avant de s'intéresser à la modification du paysage d'énergie, il est utile de revenir sur les différences entre les enthalpies mesurées par les deux méthodes.

Les expériences de calorimétrie donnent directement accès à la valeur « réelle » de l'enthalpie en mesurant des flux de chaleur entre l'échantillon et l'extérieur. En revanche, pour les dénaturations thermiques suivies par dichroïsme circulaire, l'enthalpie est calculée à partir d'équilibre entre la protéine repliée et sa forme dénaturée. Il s'agit dans ce cas de l'enthalpie de Van't Hoff. Lorsque la dénaturation implique uniquement deux états, les deux valeurs d'enthalpie doivent être égales. Dans le cas de l' $\alpha$ -lactalbumine, on considère en plus de l'état natif et de l'état déplié, l'état *molten globule*. La différence entre les deux enthalpies peut donc provenir de ce troisième état. Cependant, cette explication n'est pas pleinement satisfaisante. En effet, en présence de calcium, seul les états natifs et dépliés sont peuplés, le repliement est coopératif et on se retrouve dans une réaction à deux états où les enthalpie devrait être égales. Pour expliquer les différences, il faudrait déterminer en quoi l'état *molten globule* qui n'est pas peuplé au cours de la dénaturation pourrait avoir une influence sur cette réaction.

Une autre piste peut aussi être évoquée : les équations utilisées pour le calcul de l'enthalpie de Van't Hoff considèrent implicitement que le  $\Delta C_p$  de la protéine est nul. Dans ce cas :



$$\Delta G = \Delta H - T\Delta S$$

A la température de dénaturation,

$$\Delta G = 0$$

Donc

$$\Delta H = T_m \Delta S$$

Ainsi,

$$\Delta G = \Delta H - \frac{T}{T_m} \Delta H$$

Comme

$$\Delta G = -RT \ln K_{eq}$$

$$K_{eq} = \exp\left(-\frac{\Delta H}{R}\left(\frac{1}{T} - \frac{1}{T_m}\right)\right)$$

$$K_{eq} = \frac{[U]}{[N]}$$

Les courbes de dénaturation sont donc paramétrées en utilisant l'équation suivante :

$$\frac{[U]}{[N]+[U]} = \frac{K_{eq}}{1+K_{eq}} = \frac{\exp\left(-\frac{\Delta H}{R}\left(\frac{1}{T}-\frac{1}{T_m}\right)\right)}{1+\exp\left(-\frac{\Delta H}{R}\left(\frac{1}{T}-\frac{1}{T_m}\right)\right)}$$

Dans le cas où les  $\Delta C_p$  ne sont pas nuls, la première équation devient :

$$\Delta G = \Delta H - T\Delta S + \Delta C_p(T - T_m - T \ln(T/T_m))$$

Ce qui donne :

$$\frac{[U]}{[N]+[U]} = \frac{K_{eq}}{1+K_{eq}} = \frac{\exp\left(-\frac{\Delta H}{R}\left(\frac{1}{T}-\frac{1}{T_m}\right) - \frac{\Delta C_p}{R}\left(1 - \frac{T_m}{T} - \ln\left(\frac{T}{T_m}\right)\right)\right)}{1+\exp\left(-\frac{\Delta H}{R}\left(\frac{1}{T}-\frac{1}{T_m}\right) - \frac{\Delta C_p}{R}\left(1 - \frac{T_m}{T} - \ln\left(\frac{T}{T_m}\right)\right)\right)}$$

Les  $\Delta C_p$  calculés par calorimétrie ne sont pas précis mais ils ne semblent pas nuls. Cela peut expliquer la différence observée dans les enthalpies.

La calorimétrie donne un accès direct à l'enthalpie et seul ces valeurs seront considérées.

L'enthalpie est donc constante et cela permet d'attribuer l'augmentation de température de dénaturation à une variation d'entropie d'après l'équation :

$$\Delta H = T_m \Delta S$$

Dans cette équation, le  $\Delta S$  correspond à la différence d'entropie entre l'état natif et l'état déplié. La concentration en calcium n'a pas d'influence spécifique sur l'état déplié et la variation d'entropie mesurée est donc due à une augmentation de l'entropie de l'état natif. Ce phénomène est expliqué dans la partie suivante à l'aide du paysage d'énergie.

La forme apo est moins stable que la forme holo. Cela s'explique structuralement par la présence de charges négatives à proximité du site de liaison. Elles peuvent être compensées par les charges positives du calcium lorsque la protéine se trouve sous sa forme holo. Certains microétats appartenant à l'état de transition ou l'état *molten globule* en absence de calcium peuvent appartenir à l'état natif lorsque la protéine lie un calcium et que les charges déstabilisantes à proximité du site de liaison sont compensées par les charges de l'ion.

Ce constat permet d'expliquer les dénaturations thermiques. A faible concentration en calcium, seuls les microétats les mieux adaptés peuvent lier un ion calcium et être stabilisés. L'effet est visible en comparant les expériences réalisées en présence de calcium à celle réalisées à faible concentration de calcium. La stabilisation de l'état natif explique le caractère coopératif de la dénaturation et le fait que le *molten globule* ne soit plus peuplé au cours de la dénaturation. Lorsque la concentration en calcium augmente, l'ion pourra se lier sur des structures moins définies, moins adaptées. Ces micro états seront alors stabilisés. Certains d'entre eux qui étaient considérés comme instables et appartenant à l'état *molten globule* et de transition seront stabilisés, par compensation des charges négatives par exemple, de telle sorte qu'ils pourront alors être considérés comme appartenant à l'état natif. Le nombre de microétats qui constitue l'état natif, et donc l'entropie de cet état augmentera donc avec l'ajout de calcium ce qui augmentera par là même sa stabilité et la température de dénaturation.

## Perspectives

Actuellement, les expériences de dénaturation thermiques suivies par dichroïsme circulaire et par calorimétrie différentielle à balayage sont répétées et effectuées en présence d'autres métaux. Elles doivent permettre de décrire un peu mieux la modification du paysage d'énergie induite par les métaux.

L'origine de la différence entre les  $\Delta H$  calorimétriques et les  $\Delta H$  de Van't Hoff reste obscur et un approfondissement est nécessaire sur ce point.

D'autres techniques sont aussi utilisées pour préciser le paysage d'énergie de la protéine, notamment des expériences de repliement et de dépliement déclenchées par saut de température.

## **Repliement en absence de métaux**

### Feuillets intermoléculaires, multimères, fibres.

Lorsque la concentration en protéine augmente, la constante de vitesse de repliement diminue et des feuillets  $\beta$  non natif sont formés. La diminution de la constante de vitesse de repliement est attribuée à des interactions entre protéines. La quantité de feuillets non natifs augmente avec la concentration ce qui indique que ces feuillets sont vraisemblablement intermoléculaires. Cette hypothèse est corroborée par l'étude de la séquence de l' $\alpha$ -latalbumine. L'état qui possède ces feuillets non-natifs est donc multimérique. Pourtant, aucun oligomères n'est détecté par diffusion quasi-élastique de lumière au cours de la réaction. Ces résultats sont interprétés en considérant que les multimères formés sont de petite

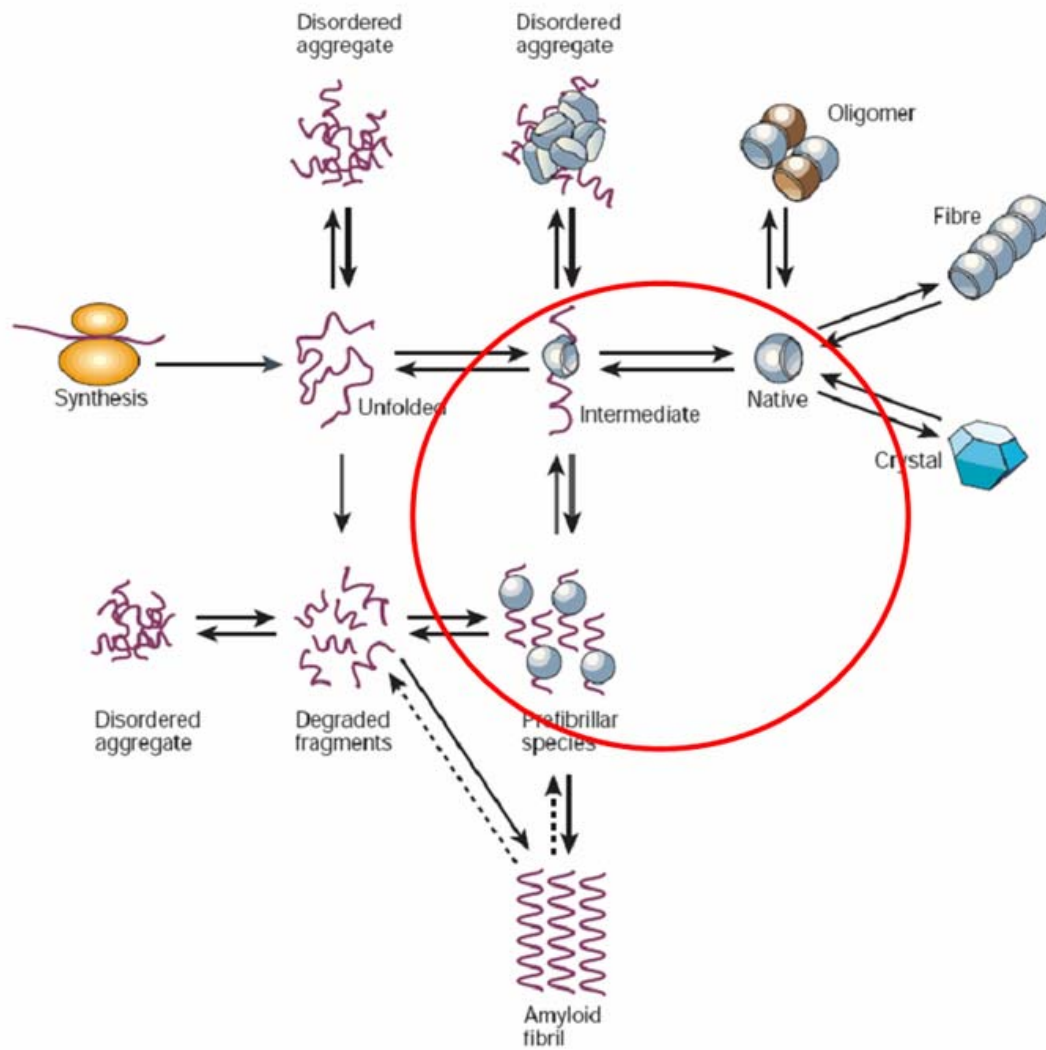
taille et que la résolution de l'appareil ne permet pas de les distinguer des monomères. Pour étayer ce raisonnement, estimons le rayon de petits oligomères constitués de quelques entités monomériques. Pour ce faire les protéines et les multimères sont assimilés à des sphères et leur densité est supposée constante. Soit  $R$  le rayon d'une protéine et  $R_N$  le rayon d'un multimère composé de  $N$  monomères. Le volume de la protéine est de  $\frac{4}{3}\pi R^3$ . Le multimère de  $N$  monomères a un volume égal à  $N$  fois le volume de la protéine. On a donc  $\frac{4}{3}\pi R_N^3 = N(\frac{4}{3}\pi R^3)$ . Ainsi,  $R_N/R = (N)^{1/3}$ . Un dimère a un rayon  $2^{1/3} = 1,25$  fois plus grand que le rayon de la protéine, un trimère 1.44 fois le rayon de la protéine. Le rayon de l' $\alpha$ -lactalbumine est de l'ordre de 20 nm, soit 25 nm pour un dimère et 29 nm pour un trimère. Pour les rayons que l'ont veut mesurer, la résolution de l'appareil est de quelques nanomètres. Des pics correspondant à un monomère et à de petits multimères peuvent effectivement ne pas être distinguables.

Une approche alternative peut être proposée pour expliquer le fait que les multimères ne sont pas détectés. Lorsque la concentration en protéine est élevée, elles interagissent toutes entre les unes avec les autres, la solution peut donc être considérée comme un pseudo-gel où toutes les protéines forment un seul grand multimère. Les interactions entre protéines sont relativement faibles, ce qui laisse à la solution son aspect liquide et ce qui permet à la protéine de se replier. Cette description présente en outre l'avantage d'expliquer le caractère rhéofluidifiant de ces solutions : lorsque la solution subit une contrainte élevée, les interactions entre protéines sont cassées et la solution devient plus liquide.

Le calcul des distances entre protéine permet de mieux comprendre cette idée de pseudo-gel. Ce calcul a été réalisé dans la partie IV.2.2. La distance entre les centres des protéines est estimé à l'aide de l'équation suivante :  $(10^{-3}M/(c \cdot Na))^{1/3}$ . Pour mesurer la distance entre les bords des protéines, il faut soustraire le diamètre de la protéine à la valeur obtenue. Pour les

hautes concentrations, la distance entre les protéines est plus petite que la taille des protéines elle-même. A 60 mg/mL par exemple, cette distance est égale à 3,3 nm. Comme la protéine a un diamètre de 4nm, aux hautes concentration, la protéine est ne peu plus se déplacer librement dans la solution. Elle ne peut plus passer entre ses voisines et est confinée spatialement.

L'état multimérique riche en feuillettes  $\beta$  non natifs présente des similitudes avec les fibres amyloïdes qui sont composées elles aussi de plusieurs protéines reliées entre elles par des feuillettes intermoléculaires. De plus, l' $\alpha$ -lactalbumine peut former ce type de fibres à partir de son *molten globule*. L'état riche en feuillettes non natif est donc considéré comme un état préfibrillaire. En solution la réaction de repliement est mise en compétition avec la formation de fibres. Le repliement l'emporte mais des états préfibrillaires sont tout de même détectés. Cette compétition a déjà été proposée (figure V.1) mais aucune preuve expérimentale directe n'a, à nos connaissances, été publiée jusqu'à présent.



**Fig V.1 : États accessibles aux chaînes polypeptidiques. (D'après Dobson 2003)**

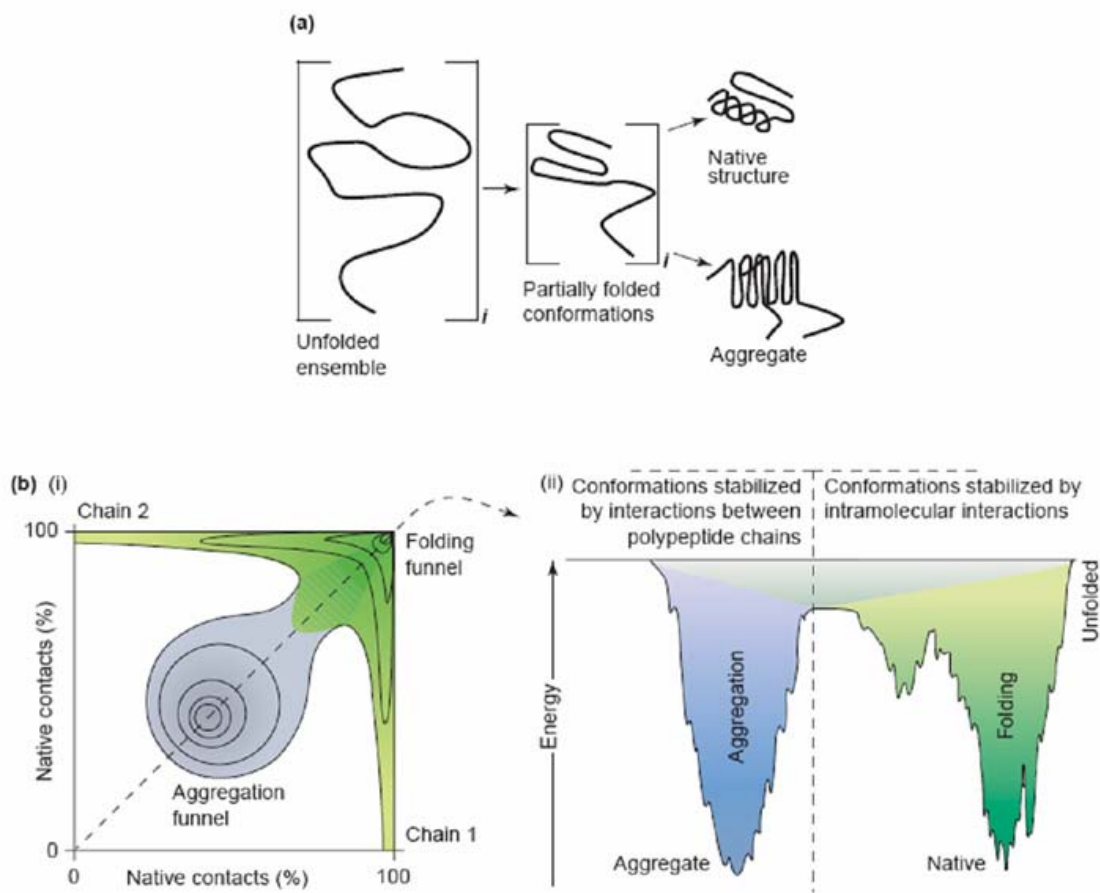
Les états explorés et les réactions étudiées par les expériences présentées précédemment sont entourés en rouge.

### Utilisation d'un paysage d'énergie pour décrire le repliement en absence de métaux.

Les expériences en présence de métaux sont décrites à l'aide des paysages d'énergie. Elles ont permis de montrer qu'en absence de calcium, l'état *molten globule* est de grande entropie. C'est ce qui permet à la protéine d'explorer de nombreuses configurations avant de se replier et de peupler l'état riche en feuillets  $\beta$ . En présence de calcium, aucun feuillet non natif est détecté. La description en terme de paysage d'énergie est plus difficile dans le cas de la forme apo à cause des interactions intermoléculaires entre les protéines dans l'état *molten globule*.

Des paysages d'énergies conçus pour tenir compte des interactions ont été proposés par Patricia Clark<sup>169</sup>. Pour les tracer, l'auteur considère un « super paysage » d'énergie qui décrit non plus l'énergie d'une protéine en fonction de sa configuration mais celle de plusieurs protéines. Pour réduire ce super paysage à un paysage d'énergie classique, seuls les cas où les protéines sont toutes dans la même configuration sont retenus. La figure V.2 est extraite de l'article de Clark. Sur cette figure, les configurations des protéines sont représentées par leur pourcentage de contacts natifs. La figure b(i) présente le super paysage d'énergie de deux protéines. En regardant ce qui se passe sur la diagonale, lorsque la configuration des deux chaînes est identique, un paysage d'énergie dépendant de la configuration d'une seule protéine peut être tracé (figure b(ii)). Ce paysage d'énergie présente deux minima : un pour la forme native de la protéine et un pour la forme agrégée.

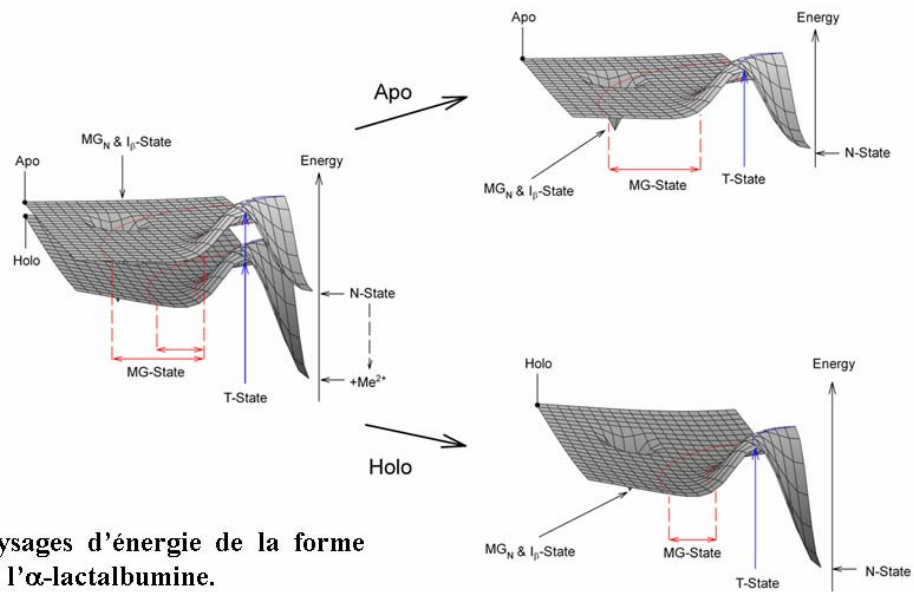




**Fig V.2 : Paysages d'énergie modifiés pour tenir compte des interactions entre protéine. D'après Clark 2004)**

- a) Différents états de la protéine
- b) Paysages d'énergie

De tels paysages d'énergie peuvent être imaginés dans notre cas. Ils sont représentés sur la figure V.3. Ils présentent l'avantage d'expliquer facilement pourquoi les feuillets  $\beta$  non-natifs ne sont pas détectés lorsque la protéine se replie en présence de métaux. En effet, les métaux ont la particularité de diminuer l'entropie de l'état *molten globule*. Les protéines dans cet état ne peuvent pas explorer autant de configuration qu'en absence de calcium et ils n'ont pas accès au minimum local qui correspond à la formation de l'état  $I_\beta$ .



**Fig V.3 : Paysages d'énergie de la forme apo et holo de l' $\alpha$ -lactalbumine.**

Les paysages présentés dans la première partie pour expliquer l'évolution des vitesses avec la concentration en protéine ont été modifiés pour prendre en compte les interactions entre protéines et les états  $MG_N$  et  $I_\beta$ .

## Perspectives

Des méthodes doivent être imaginées pour réaliser des expériences destinées à confirmer ou infirmer l'hypothèse de la formation d'un pseudogel lorsque la concentration en protéines est importante.

La portion de la protéine impliquée dans la formation des feuillets  $\beta$  non natifs a été mise en évidence en utilisant des méthodes de prédiction. La portion déterminée par prédiction est bien cohérente avec tous les résultats expérimentaux, elle explique bien que les feuillets  $\beta$  intermoléculaire sont formés depuis l'état *molten globule*. Cependant, une vérification expérimentale directe peut être utile. Elle peut être effectuée par des expériences de mutagenèse dirigée. Des acides aminés de la séquence peuvent être mutés pour favoriser la

structuration en hélice  $\alpha$ . Une diminution ou une disparition de la phase lente confirmerait alors l'implication de l'acide aminé muté dans le feuillet non natif. Ce type d'analyse permettrait ainsi de déterminer avec précision la portion de la protéine qui forme des états non-natifs.

## **Formation de fibres amyloïdes**

Dans la dernière partie, l'effet du sel sur les cinétiques de formation de fibres amyloïdes est exploré. Les cinétiques suivies par fluorescence de la thioflavine montre clairement cet effet. Les expériences de dichroïsme circulaire et de diffusion de rayons X aux petits angles indiquent que le sel n'agit pas directement sur les protéines mais sur leurs interactions. L'augmentation de la concentration en sel diminue la répulsion électrostatique en écrantant les charges des protéines.

Ces études permettent de mettre en évidence l'importance des interactions et notamment de l'équilibre entre les attractions hydrophobes et les répulsions électrostatiques<sup>164</sup>.

L'effet du sel sur les cinétique est observé mais la conséquence sur la structure de fibres formés, leur dimension et leur stabilité n'est pas exploré ici.

L'étude du gel formé lorsque la concentration en protéine est importante à été effleuré seulement et demande d'être approfondie. Les liaisons qui se créent entre les fibres au sein du gel doivent être caractérisées ainsi que l'influence du sel sur les caractéristiques mécaniques du gel.

1. **Mulder G. (1839)** On the composition of some animal substances. *Journal für praktische Chemie* **16**, 129.
2. **Benhabilés N., Thomas A., Brasseur R. (2000)**. Les mécanismes de repliement des protéines solubles. *Biotechnol. Agron. Soc. Environ* **4** (2): 71-81.
3. **Grimaux E. (1882)** Sur des colloïdes azotes. *Bull. soc. chim.*, **38**, 64.
4. **Hofmeister F. (1902)** Über Bau und Gruppierung der Eiweisskörper. *Ergebnisse der Physiologie*, **1**, 759.
5. **Fischer E. (1902)** Die Hydrolyse der Proteinstoffe. *Zeitschrift für Lebensmitteluntersuchung und -Forschung A.* **24**, 1431-4630.
6. **Svedberg T., Tiselius A. (1926)** A new method for the determination of the molecular weight of the protein. *JACS*, **48**, 30.
7. **Wrinch D. M. (1937)** The Cyclol Hypothesis and the "Globular" Proteins. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical sciences*, **907**, 505-524
8. **Sanger F. (1952)** The arrangement of amino acids in proteins. *Adv Protein Chem* **7**: 1-67.
9. **Perutz M.F., Rossmann M.G., Cullis A.F., Muirhead G., Will G., North A.T. (1960)** Structure of haemoglobin: A three-dimensional Fourier synthesis at 5.5 Å resolution, obtained by X-ray analysis *Nature*, **185**, 416-422.
10. **Kendrew J.C., Dickerson R.E., Strandberg B.E., Hart R.J., Davies, D.R., Philips, D.C. (1960)** Structure of myoglobin: A three-dimensional Fourier synthesis at 2 Å resolution. *Nature*, **185**, 422-427.
11. **Astbury W.T., Street A. (1931)** X-ray studies of the structure of hair, wool and related fibres. I. General. *Trans. R. Soc. Lond.*, **A320**, 75-100
12. **Pauling L., Corey R.B. (1951)** The pleated sheet, a new layer configuration of polypeptide chains. *Proc Natl Acad Sci USA* **37(5)**, 251-256.
13. **Pauling L., Corey R.B., Branson H.R. (1951)** The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain. *Proc Natl Acad Sci USA* **37(4)**, 205-211.
14. **Rost B. (2001)** Protein secondary structure continues to rise. *Journal of Struct. Biol.* **134**, 204-218
15. **Rost B, Sander C. (2000)** Third generation prediction of secondary structure in Webster, D. (Ed.), Protein structure prediction: methods and protocols, p71-95, Humana press, Clifton, N.J.
16. **Wüthrich K. (1986)** NMR of Proteins and Nucleic Acids, Wiley, New York.

17. **Pande V.S., Baker I., Chapman J., Elmer S.P., Khaliq S., Larson S.M., Rhee Y.M., Shirts M.R., Snow C.D., Sorin E.J., Zagrovic B. (2003)** Atomistic protein folding simulations on the submillisecond timescale using worldwide distributed Computing. *Biopolymers*, **68(1)**, 91-109
18. **Marti-Renom M.A., Stuart A.C., Fiser A., Sanchez R., Melo F., Sali A. (2000).** Comparative protein structure modeling of genes and genomes. *Annu Rev Biophys Biomol Struct* **29**, 291-325.
19. **Bowie J.U., Lüthy R., Eisenberg D. (1991)** A method to identify protein sequences that fold into a known three-dimensional structure. *Science*. **253**, 164-170.
20. **Baldwin R.L. (1999)** Protein folding from 1961 to 1982. *Nature Struct. Biol.* **6(9)**, 814-817
21. **Anfinsen C.B., Haber E., Sela M., White, Jr., F.W. (1961)** The Kinetics of Formation of Native Ribonuclease during Oxidation of the Reduced Polypeptide Chain. *Proc. Natl. Acad. Sci. USA* **47**, 1309–1314.
22. **Anfinsen C.B. (1973)** Principles that govern the folding of protein chains. *Science* **181(96)**, 223-230.
23. **Wu, H (1931)** Studies on Denaturation of Proteins. XIII. A Theory of Denaturation. *Chinese Journal of Physiology* **5**, 321–344.
24. **Kauzmann W.<sup>I II III IV V</sup>, Simpson R. B.<sup>I II III</sup>, Schellman J.<sup>II</sup>, Frensdorff H. K.<sup>IV V</sup>, Watson M. T.<sup>IV V</sup> (1953)** The Kinetics of Protein Denaturation. I, II, III, IV, V. *JACS* **75(21)**, 5139-5172.
25. **Brandts J.F. (1964)** The Thermodynamics of Protein Denaturation. I, II. *JACS* **86(20)**, 4291-4314.
26. **Tanford C. (1968)** Protein denaturation. *Adv. Prot. Chem.* **23**, 122–282.
27. **Tanford C. (1970)** Protein denaturation. C. Theoretical models for the mechanism of denaturation. *Adv. Prot. Chem.* **24**, 1–95.
28. **Privalov P.L., Khechinashvili N.N., Atansov B.P. (1971)** Thermodynamic analysis of thermal transitions in globular proteins. I. Calorimetric study of chymotrypsinogen, ribonuclease and myoglobin. *Biopolymers* **10**, 1865–1890 (1971).
29. **Goldberg M.E.(1969)** Tertiary structure of Escherichia coli  $\beta$ -D-galactosidase. *J. Mol. Biol.* **46**, 441–446.
30. **Taniuchi H.(1970)** Formation of Randomly Paired Disulfide Bonds in Des-(121–124)-ribonuclease after Reduction and Reoxidation. *J. Biol. Chem.* **245**, 5459–5468.
31. **Wetlaufer D.B.(1973)** Nucleation, Rapid Folding, and Globular Intrachain Regions in Proteins. *Proc. Natl. Acad. Sci. USA* **70**, 697–701.

32. **Levinthal C.(1968)** Are there pathways for protein folding? *J. Chim. Phys.* **65**, 44–45.
33. **Ikai A., Tanford C. (1971)** Kinetic Evidence for Incorrectly Folded Intermediate States in the Refolding of Denatured Proteins. *Nature* **230**, 100–102.
34. **Tsong T.Y., Baldwin R.L., Elson E.L.(1971)** The Sequential Unfolding of Ribonuclease A: Detection of a Fast Initial Phase in the Kinetics of Unfolding. *Proc. Natl. Acad. Sci. USA* **68**, 2712–2715.
35. **Creighton T.E.(1974)** Renaturation of the reduced bovine pancreatic trypsin inhibitor. *J. Mol. Biol.* **87**, 579–602.
36. **Cook K.H., Schmid F.X., Baldwin R.L.(1979)** Role of proline isomerization in folding of ribonuclease A at low temperatures. *Proc. Natl. Acad. Sci. USA* **76**, 6157–6161.
37. **Schmid F.X., Baldwin R.L.(1979)** The rate of interconversion between the two unfolded forms of ribonuclease A does not depend on guanidinium chloride concentration. *J. Mol. Biol.* **135**, 199–215.
38. **Evans P.A., Radford S.E.(1994)** Probing the structure of folding intermediates. *Curr. Opin. Struct. Biol.* **4**, 100-106.
39. **Troullier A., Reinstädler D., Dupont Y., Naumann D., Forge V.(2000)** Transient non-native secondary structures during the refolding of  $\alpha$ -lactalbumin detected by infrared spectroscopy. *Nat. Struct. Biol.* **7**, 78-86
40. **Pollack L., Tate M.W., Darnton N.C., Knight J.B., Gruner S.M., Eaton W.A., Austin R.H. (1999)** Compactness of the denatured state of a fast-folding protein measured by submillisecond small-angle x-ray scattering. *Proc. Natl. Acad. Sci. USA* **96**, 10115-10117.
41. **Eliezer D., Jennings P.A., Wright P.E., Doniach S., Hodgson S.K., Tsuruta H.(1995)** The radius of gyration of an apomyoglobin folding intermediate. *Science* **270**, 487-488
42. **Balbach J. Forge V., van Nuland N.A.J., Winder S.L., Hore P.J., Dobson C.M. (1995)** Following protein folding in real-time using NMR spectroscopy. *Nat. Struct. Biol.* **2**, 865-870.
43. **Balbach J., Forge V., Lau W.S., van Nuland N.A.J., Brew K., Dobson C.M. (1996)** Protein folding monitored at individual residues during a two dimensional NMR experiment. *Science*, 274(5290), 1161-1163.
44. **Englander S.W. (2000)** Protein folding intermediates and pathways studies by hydrogen exchange. *Ann. Rev. Biophys. Biomol. Struct.* **29**, 213-238.
45. **Miranker A., Robinson C.V., Radford S.E., Dobson C.M. (1996)** Investigation of protein folding by mass spectrometry. *FASEB J.* **10**, 93-101.

46. **Fisher T.E., Oberhauser A.F., Carrion-Vazquez M., Marszalek P.E., Fernandez J.M. (1999)** The study of protein mechanics with the atomic force microscope. *Trends Biochem. Sci.* **24**, 379-384.
47. **Udgaonkar J.B., Baldwin R.L.(1995)** Nature of the early intermediates of ribonuclease A. *Biochemistry* **34**, 4088-4096.
48. **Itzhaki L.S., Evans P.A., Dobson C.M., Radford S.E. (1994)** Tertiary interactions in the folding pathways of hen lysozyme: Kinetic studies using fluorescent probes. *Biochemistry* **33**, 5212-5220.
49. **Eaton W.A., Munoz V., Hagen S.J., Jas G.S., Lapidus L.J., Henry E.R., Hofrichter J.(2000)** Fast kinetics and mechanics in protein folding. *Annu. Rev. Biophys. Biomol Struct.* **29**, 327-359.
50. **Roder H., Shastry M.C.R.(1999)** Methods for exploring early events in protein folding. *Curr. Opin. Struct. Biol.* **9**, 620-626
51. **Engelhard M., Evans P.A. (1996)a** Experimental investigation of side chain interactions in early protein folding intermediates. *Folding Des.* **1**, 31-37.
52. **Engelhard M., Evans P.A. (1996)b** Kinetics of interaction of partially folded proteins with a hydrophobic dye: Evidence that molten globule character is maximal in early folding intermediates. *Protein Sci.* **4**, 1553-1562.
53. **Rischek C., Poulsen F.M.(1995)** Modification of a specific tyrosine enables tracing of the end-to-tend distance during apomyoglobin folding. *FEBS Lett.* **374**, 105-109.
54. **Jones E., Beechem J.M., Matthews C.R. (1995)** Local and global dynamics during the folding of E.coli dihydrofolate reductase by time-resolved fluorescence spectroscopy. *Biochemistry* **34**, 1867-1877.
55. **Matouschek A., Kellis J.T., Serrano L., Fersht A.R.(1989)** Mapping the transition state and pathway of protein folding by protein engineering. *Nature* **340**, 122-126
56. **Brockwell D.J., Smith D.A., Radford S.E.(2000)** Protein folding mechanisms: new methods and emerging ideas. *Curr. Opin. Struct Biol.* **10**, 16-25.
57. **Li A., Daggett V. (1994)** Characterization of the transition state of protein unfolding by use of molecular dynamics: Chymotrypsin inhibitor 2. *Proc. Natl. Acad. Sci. USA*, **91**, 10430-10434.
58. **Lazaridis T., Karplus M. (1997)** "New view" of protein folding reconciled with the old through multiple unfolding simulations. *Science* **278(5345)**, 1928-1931.
59. **Hermans J.J., Lohr D., Ferro D. (1969)** Unfolding and hydrogen exchange of proteins: The three-dimensional Ising lattice as a model. *Nature* **224**, 175-177.

60. **Go N. (1983)** Theoretical studies of protein folding. *Ann. Rev. Biophys. Bioeng.* **12**, 183-210.
61. **Wolynes P.G., Onuchic J.N., Thirumalai D. (1995)** Navigating the folding route. *Science* **267**, 1619-1620.
62. **Dill K.A., Bromberg S., Yue K.Z., Fiebig K.M., Yee D.P., Thomas P.D., Chan H.S. (1995)** Principles of protein folding: A perspective from simple exact models. *Protein Sc.* **4**, 561-602.
63. **Skolnick J., Kolinski A. (1990)** Simulation of the folding of globular protein. *Science* **250**, 1121-1125.
64. **Miyazawa S., Jernigan R.L. (1985)** Estimation of effective interresidue contact energies from protein crystal structures: quasi chemical approximation. *Macromolecules* **18**, 534-552.
65. **Onuchic J.N., Nymeyer H, Garcia A.E., Chahine J., Socci N.D. (2000)** The energy landscape theory of protein folding: insights into folding mechanisms and scenarios. *Adv. Prot. Chem.* **53**, 87-152
66. **Dill K.A., Chan H.S. (1997)** From Levinthal to pathways to funnels. *Nat. Struct. Biol.* **4**(1) 10-19.
67. **Bryngelson J.D., Onuchic J.N., Socci N.D., Wolynes P.G. (1995)** Funnels, pathways, and the energy landscape of protein folding: a synthesis. *Proteins* **21**(3), 167-195.
68. **Wolynes P.G., Onuchic J.N., Thirumalai D. (1995)** Navigating the folding route. *Science* **267**, 1619-1620.
69. **Eyring H., Polanyi M. (1931)** On simple gas reaction. *Zeitschrift für physikalische Chemie B.* **12**, 279-311.
70. **Pelzer H., Wigner E., (1932)** Über die Geschwindigkeitkonstante von Austauschreaktionen. *Zeitschrift für physikalische Chemie B.* **15**, 445-463.
71. **Fersht A., (1999)** Structure and Mechanism in Protein Science. *New York: Freeman.*
72. **Bertini I., Gary H.B., Lippard S.J., Valentine, J.S. (1994)** Bioinorganic chemistry. *University science books, Mill Valley, CA.*
73. **Lippard S.J., Berg J.M. (1994)** Principles of bioinorganic chemistry. *University science books, Mill Valley, CA.*
74. **Fenton D.E. (1995)** Biocoordination chemistry. *Oxford UK: Oxford university press.*
75. **Gray H.B. (2003)** Biological inorganic biochemistry at the beginning of the 21<sup>st</sup> century. *Proc. Natl. Acad. Sc. USA* **100**, 3563-3568.



76. **Fenton D.E. (1995)** Biocoordination chemistry. *Oxford UK: Oxford university press.*
77. **Bertini I., Cowan J.A., Luchinat C., Natarajan K, Piccioli M. (1997)** Characterisation of a partially unfolded high potential iron protein. *Biochemistry* **36**, 9332-9339.
78. **Robinson C.R., Liu Y., Thomson J.A., Sturtevant J.M., Seigar S.G.. (1997)** Energetics of heme binding to native and denatured states of cytochrome b562. *Biochemistry* **36**, 16141-16146.
79. **Wittung-Stafshede P., Lee J.C., Winkler J.R. Gray H.B. (1999)** Cytochrome b562 folding triggered by electron transfer: approaching the speed limit for formation of four-helix-bundle protein. *Proc. Natl. Acad. Sc. USA.* **96**, 6587-6590.
80. **Alzheimer A.,(1907)** Über eine eigenartige Erkrankung der Hirnrinde. *Allg. Z. Psychiat. Psych.-Gerichtl. Med.* **64**, 146–148.
81. **Gallez C. (2005)** Rapport sur la maladie d'Alzheimer et les maladies apparentées. *Rapports de l'office parlementaire d'évaluation des politiques de santé n°2454.*
82. **Ross C.A., Poirier M.A. (2004)** Protein aggregation and neurodegenerative disease. *Nat. Med.* **10**, S10-S17.
83. **Creutzfeldt H. (1920)** Über eine eigenartige herdförmige Erkrankung des Zentralnervensystems. *Z. Ges. Neurol. Psychiat.* **57**, 1-20.
84. **Jakob A. (1921)** Über eigenartige Erkrankungen des Zentralnervensystems mit bemerkenswerten anatomischen Befunden. *Z. Ges. Neurol. Psychiatr.* **64**, 147-228.
85. **Huntington G. (1872)** On chorea. *The medical and surgical reporter* **26(15)**, 317-321.
86. **Manyam B.V., Sánchez-Ramos J.R. (1999).** Traditional and complementary therapies in Parkinson's disease. *Advances in neurology* **80**: 565-74.
87. **Parkinson J. (1817)** An essay on the shaking palsy.(Reproduced). *J Neuropsychiatry Clin Neurosci* **14** (2): 223-36 (2002).
88. **Dobson C.M. (2002)** Getting out of shape. *Nature* **418**, 729-730.
89. **Wickner R.B., Edskes H.K., Shewmaker F., Nakayashiki T. (2007)** Prions of fungi: inherited structures and biological roles. *Nature reviews microbiology* **5**,611-618.
90. **Wickner R.B. (1997)** A new controls fungal cell fusion incompatibility. *Proc. Natl. Acad. Sc. USA* **94**, 10012-10014.
91. **Si K., Lindquist S, Kandel E.R. (2003)** A neuronal isoform of the aplysia CPEB has prion-like properties. **115(7)**, 879-891.
92. **Shorter J., Lindquist S. (2005)** Prions as adaptive conduits of memory and inheritance. *Nat Rev Genet* **6** (6), 435-50.

93. **Fowler D.M., Koulov A.V., Alory-Jost C., Marks M.S., Balch W.E., Kelly J.W. (2006)** Functional amyloid formation within mammalian tissue. *PLOS Biology* **4(1)**, 100-107.
94. **MacPhee C.E., Dobson C.M. (2000)** Formation of mixed fibrils demonstrates the generic nature and potential utility of amyloid nanostructures. *JACS* **122(51)**, 12707-12713.
95. **Guijarro J.I., Sunde M., Jones J.A., Campbell I.D., Dobson C.M. (1998)** Amyloid fibril formation by an SH3 domain. *Biochemistry* **95(8)**, 4224-4228.
96. **Dobson C.M. (2003)** Protein folding and misfolding. *Nature* **426**, 884-890.
97. **Astbury W.T., Dickinson S. (1935)** The X-ray interpretation of denaturation and the structure of the seed globulins. *Biochem. J.* **29(10)**, 2351-2360.
98. **Makin O.S., Serpell L.C., (2005)** Structures for  $\mu$ amyloid fibrils. *Febs J.* **272**, 5950-5961
99. **Nelson R., Eisenberg D. (2006)** Structural models of amyloid-like fibrils. *Adv. Protein Chem.* **73**, 235-282.
100. **Lomakin A., Teplow A.B., Kirschner D.A., Benedek G.B. (1997)**. Kinetic theory of amyloid- $\beta$  protein. *Proc. Natl. Acad. Sc. USA.* **94**, 7942-7947.
101. **Wetzel R. (2006)** Kinetics and thermodynamics of amyloid fibril assembly. *Acc. Chem. Res.* **39**, 671-679.
102. **Chothia C., Janin J. (1975)** Principles of protein-protein recognition. *Nature* **256**, 705-708.
103. **Hortschansky P., Schroeckh V., Christopeit T., Zandomenighi G., Fändrich M. (2005)** The aggregation kinetics of Alzheimer's  $\beta$ -amyloid peptide is controlled by stochastic nucleation. *Protein Science* **14**, 1753-1759.
104. **Nielsen L., Khurana R., Coats A., Frokjaer S., Brange J., Vyas S., Uversky N.U., Fink A.L. (2001)** Effect of environmental factors on the kinetics of insulin fibril formation Elucidation of the molecular mechanism. *Biochemistry* **40**, 6036-6046.
105. **Sabaté R., Gallardo M., Estelrich J. (2005)** Temperature dependence of the nucleation constant rate in  $\beta$  amyloid fibrillogenesis. *International Journal of Biological Macromolecules* **35**, 9-13.
106. **Kusumoto Y., Lomakin A., Teplow D.B., Benedek G.B. (1998)** Temperature dependence of amyloid  $\beta$ -protein fibrillization. *Proc. Natl. Acad. Sc. USA.* **95(21)**, 12277-12282.
107. **Uversky V.N., Fink A.L. (2004)** Conformational constraints for amyloid fibrillation: the importance of being unfolded. *Biochim. Biophys. Acta* **1698(2)**, 131-153.

108. **Fändrich M., Forge V., Buder K., Kittler M., Dobson C.M., Diekmann S. (2003)** Myoglobin forms amyloid fibrils by association of unfolded polypeptide segments. *Proc. Natl. Acad. Sc. USA.* **100(26)** 15463-15468.
109. **Schmittschmitt J.P., Scholtz J.M. (2003)** The role of protein stability, solubility, and net charge in amyloid fibril formation. *Protein Science* **12**, 2374-2378.
110. **Booth D.R., Sunde M., Bellotti V., Robinson C.V., Hutchinson W.L., Fraser P.E., Hawkins P.N., Dobson C.M., Radford S.E., Blake C.C.F., Pepys M.B. (1997).** Instability, unfolding and aggregation of human lysozyme variants underlying amyloid fibrillogenesis. *Nature* **385**, 787-793.
111. **Kallberg Y., Gustafsson M., Persson B., Thyberg J., Johansson J. (2001)** Prediction of amyloid fibril-forming proteins. *J. Biol. Chem.* **276(16)**, 12945-12950.
112. **Hill R.L. and Brew K. (1975).** Lactose synthetase *Adv. Enzymol.Rel Areas Mol. Biol* **43**, 411-489.
113. **O'Keeffe E.T., Hill R.L. and Bell J.E. (1980).** Bovine galactosyltransferase: interaction with alpha-lactalbumin and the role of alpha-lactalbumin in lactose synthase. *Biochemistry* **19**, 4954-4962.
114. **Hakansson A., Zhivotovsky B., Orrenius S., Sabharwal H., Svanborg C. (1995).** Apoptosis induced by a human milk protein *Proc Natl Acad Sci USA* **92**, 8064-8068.
115. **Hakansson A., Svensson M., Mossberg A.K., Sabharwal H., Linse S., Lazou I., Lonnerdal B., Svanborg C. (2000).** A folding variant of alpha-lactalbumin with bactericidal activity against *Streptococcus pneumoniae*. *Mol Microbiol.* **35(3)**, 589-600
116. **Kohler C., Gogvadze V., Hakansson A., Svanborg C., Orrenius S., Zhivotovsky B. (2001).** A folding variant of human alpha-lactalbumin induces mitochondrial permeability transition in isolated mitochondria. *Eur J Biochem* **268(1)**, 186-91.
117. **Svensson M., Hakansson A., Mossberg A.K., Linse S., Svanborg C. (2000).** Conversion of alpha-lactalbumin to a protein inducing apoptosis. *Proc Natl Acad Sci USA.* **97(8)**, 4221-4226.
118. **Pellegrini A., Thomas U., Bramaz N., Hunziker P., and von Fellenberg R. (1999).** Isolation and identification of three bactericidal domains in the bovine alpha-lactalbumin molecule. *Biochim. Biophys. Acta* **1426**, 439- 448.
119. **Permyakov E.A, Berliner L.J. (2000).** Alpha-Lactalbumin: structure and function. *FEBS Lett* **473(3)**, 269-74. Review.
120. **Anderson. P.J., Brooks C.L., and Berlinger L.J. (1997).** Functional Identification of Calcium Binding residues in Bovine alpha-lactalbumin. *Biochemistry.* **36**, 11648-11654.

121. **Permyakov E. A., Kalinichenko L.P., Morozova L.A., Yarmolenko V.V. and Burstein E.A. (1981).** Alpha-lactalbumin binds magnesium ions: study by means of intrinsic fluorescence technique. *Biochem Biophys Res Commun* **102(1)**, 1-7.
122. **Permyakov E.A, Morozova L.A, Burstein E.A. (1985).** Cation binding effects on the pH, thermal and urea denaturation transitions in alpha-lactalbumin. *Biophys Chem* **21(1)**, 21-31.
123. **Hendrix T., Griko Y.V., Privalov P.L. (2000).** A calorimetric study of the influence of calcium on the stability of bovine alpha-lactalbumin. *Biophys Chem* **84(1)**, 27-34.
124. **Chrysina E.D., Brew K. and Acharaya K.R. (2000).** Crystal Structures of Apo- and Holo- bovine alpha-lactalbumin at 2.2-Å Resolution Reveal an Effect of Calcium on Inter-lobe Interactions. *J Biol Chem* **275(47)**, 37021-37029.
125. **Wijesinha-Bettoni R., Dobson C.M. and Redfield C. (2001).** Comparison of the Structural and dynamical properties of Holo and Apo alpha-lactalbumin by NMR Spectroscopy. *J. Mol. Biol* **307**, 885-898.
126. **Kuwajima K., Harushima Y., Sugai S. (1986)** Influence of Ca<sup>2+</sup> binding on the structure and stability of bovine α-lactalbumin studied by circular dichroism and nuclear magnetic resonance spectra. *Int. J. Pept. Protein Res.* **27(1)**, 18-27.
127. **Ishikawa N., Chiba T., Chen L.T., Shimizu A., Ikeguchi M. and Sugai S.(1998).** Remarkable destabilization of recombinant alpha-lactalbumin by an extraneous N-terminal methionyl residue. *Protein Eng* **11**, 333-335.
128. **Murakami K., Andree P.J, Berliner L.J. (1982).** Metal ion binding to alpha-lactalbumin species. *Biochemistry* **21(22)**, 5488-94.
129. **Aramini J.M., Kiraoki T., Grace M.R., Swaddle T.W., Chiancone E. and Vogel H.J. (1996).** NMR and stopped-flow studies of metal ion binding to alpha-lactalbumins. *BBA* **1293**, 72-82.
130. **Desmet J., Haezebrouck P., Van Cauwelaert F. (1991).** Thermodynamic data on the binding of six M2(+)-ions to bovine, goat, and human alpha-lactalbumin. *J Inorg Biochem* **42(2)**, 139-45.
131. **Desmet J., Van Dael H., Van Cauwelaert F., Nitta K., Sugai S. (1989).** Comparison of the binding of Ca<sup>2+</sup> and Mn<sup>2+</sup> to bovine alpha-lactalbumin and equine lysozyme. *J Inorg Biochem* **37(3)**, 185-91.
132. **Desmet J. and Van Cauwelaert F. (1988).** Calorimetric experiments of Mn<sup>2+</sup> - binding to alpha-lactalbumin. *Biochim Biophys Acta* **957(3)**, 411-9.
133. **Sandier A., Amiel C., Sebille B., Rouchaud J.C., Fedoroff M. (1999).** A study of strontium binding to albumins, by a chromatographic method involving atomic emission spectrometric detection. *Int J Biol Macromol* **24(1)**, 43-48.

134. **Kronman M.J. (1989).** Metal-ion binding and the molecular conformational properties of alpha-lactalbumin. *Crit Rev Biochem Mol Biol* **(6)**, 565-667.
135. **Chandra N., Brew K., Acharya K.R. (1998).** Structural evidence for the presence of a secondary calcium binding site in human alpha-lactalbumin. *Biochemistry* **37(14)**, 4767-72.
136. **Murakami K. and Berliner L.J. (1983).** A distinct Zinc binding site in the  $\alpha$ -lactalbumins regulates calcium binding. Is there a physiological role for this control? *Biochemistry* **22**, 3370-3374.
137. **Musci G., Berliner L.J. (1985).** Physiological roles of zinc and calcium binding to alpha-lactalbumin in lactose biosynthesis. *Biochemistry* **24(24)**, 6945-6948.
138. **Permyakov E.A., Morozova L.A., Kalinichenko L.P., Derezhkov Vyu. (1988).** Interaction of alpha-lactalbumin with  $\text{Cu}^{2+}$ . *Biophys Chem* **32(1)**, 37-42.
139. **Dolgikh D. A., Gilmanshin R. I. , Brazhnikov E.V. , V.E. Bychkova V.E., G. V. Semisotnov G. V., Venyaminov S. Yu., Ptitsyn O. B. (1981).**  $\alpha$ - Lactalbumin: compact state with fluctuating tertiary structure? *FEBS Letters* **136**, 311-315.
140. **Alexandrescu A.T., Evans P.A., Pitkethly M., Baum J. and Dobson C.M. (1993).** Structure and dynamics of the acid-denatured molten globule state of alpha-lactalbumin: a two-dimensional NMR study. *Biochemistry* **32(7)**, 1707-1718.
141. **Kuwajima K. (1996).** The molten globule state of alpha-lactalbumin. *FASEB J.* **10**, 102-109.
142. **Permyakov E.A., S.E. Permyakov, and Berliner L.J. (2003).**  $\alpha$ -Lactalbumin:  $\text{Ca}^{2+}$  binding protein with multiple functions. *Protein structures: Kaleidoscope of Structural properties and Functions.* 475-497.
143. **Forge V., Wijesinha R.T., Balbach J., Brew K., Robinson C. V., Redfield C. and Dobson C.M. (1999).** Rapid collapse and slow structural reorganisation during the refolding of bovine alpha-lactalbumin. *J. Mol. Biol.* **288**, 673-688.
144. **Balbach J., Forge V., Lau W.S., Jones J. A., van Nuland N.A., Dobson C.M. (1997).** Detection of residue contacts in a protein folding intermediate. *Proc Natl Acad Sci USA* **94(14)**, 7182-5.
145. **Gast K., Muller-Frohne M. and Damaschun G. (1998).** Compactness of the kinetic molten globule of bovine alpha-lactalbumin: a dynamic light scattering study. *Protein Sci* **7(9)**, 2004-2011.
146. **Kataoka M., Kuwajima K., Tokunaga F., and Goto Y. (1997)** Structural characterization of the molten globule of alpha-lactalbumin by solution X-ray scattering. *Protein. Sci* **6**, 422-430.

147. **Kuwajima K., Mitani M. and Sugai S. (1989).** Characterization of the Critical State in Protein Folding Effects of Guanidine Hydrochloride and Specific Ca<sup>2+</sup> Binding on the folding Kinetics of alpha-lactalbumin. *J. Mol. Biol.* **206**, 547-561.
148. **Arai M., Ito K., Inobe T., Nakao M., Maki K., Kamagata K., Kihara H., Amemiya Y., Kuwajima K. (2002).** Fast compaction of alpha-lactalbumin during folding studied by stopped-flow X-ray scattering. *J Mol Biol* **321(1)**, 121-32.
149. **Balbach J., Forge V., Lau W.S., van Nuland N.A., Brew K., Dobson C.M. (1996).** Protein folding monitored at individual residues during a two-dimensional NMR experiment. *Science* **274(5290)**, 1161-1163.
150. **Arai M., Kuwajima K. (1996).** Rapid formation of a molten globule intermediate in refolding of alpha-lactalbumin. *Fold Des.* **1(4)**, 275-87.
151. **Schellman J.A. (1975)** Macromolecular binding. *Biopolymers* **14**, 999-1018.
152. **Kuznetsova I.M., Khaitlina S.Y., Konditerov S.N., Surin A.M., Turoverov K.K. (1988)** Changes of structure and intermolecular mobility in the course of actin denaturation. *Biophys. Chem.* **32**, 73-78.
153. **Jelesarov I., Bosshard H.R. (1999)** Isothermal titration calorimetry and differential scanning calorimetry as complementary tools to investigate the energetics of biomolecular recognition. *Journal of molecular recognition* **12**, 3-18.
154. **Robertson A.D., Murphy K.P. (1997)** Protein structure and the energetics of protein stability. *Chem. Rev.* **97(5)**, 1251-1268.
155. **Simha R. (1940)** The influence of Brownian movement on the viscosity of solutions. *J. Phys. Chem.* **44(1)**, 25-34.
156. **Noda I. (1993)** Generalized two-dimensional correlation method applicable to infrared, Raman, and other types of spectroscopy. *Applied Spectroscopy.* **47(9)**, 1329-1336.
157. **Noda I. (1990)** Two-dimensional infrared (2D IR) spectroscopy: theory and applications. *Applied Spectroscopy.* **44(4)**, 550-561.
158. **Alexandrescu A.T., Broadhurst R.W., Wormald C., Chyan C.L., Baum J. and Dobson C.M. (1992).** <sup>1</sup>H-NMR assignments and local environments of aromatic residues in bovine, human and guinea pig variants of α-lactalbumin. *Eur. J. Biochem.* **210**, 699-709.
159. **Linding R., Schymkowitz I., Rousseau F., Diella F., Serrano L. (2004)** A comparative study of the relationship between protein structure and beta-aggregation in globular and intrinsically disordered proteins. *J. Mol. Biol.* **342(1)**, 345-353.

160. **Goers J, Permyakov S.E., Permyakov E.A., Uversky V.N., Fink A.L. (2002)** Conformational prerequisites for  $\alpha$ -lactalbumin fibrillation. *Biochemistry* **42**, 12546-12551.
161. **Naiki H., Higuchi K., Hosokawa M., taekeda T. (1989)** Fluorometric determination of amyloid fibrils in vitro using the fluorescent dye, thioflavine T. *Anal. Biochem.* **177(2)**, 244-249.
162. **Kelényi G. (1967)** Thioflavin S fluorescent and Congo red anisotropic stainings in the histologic demonstration of amyloid. *Acta Neuropathol.* **7(4)**, 336-348.
163. **Krebs M.R.H., Bromley E.H.C., Donald A.M. (2005)** The binding of thioflavin-T to amyloid fibrils: localisation and implications. *J. Struct. Biol.* **149**, 30-37.
164. **Raman B., Chatani E., Kihara M., Ban T., Sakai M., Hasegawa K., Naiki H., Rao C.M., Goto Y. (2005)** Critical balance of electrostatic and hydrophobic interactions is required for  $\beta_2$ -microglobulin amyloid fibril growth and stability. *Biochemistry*, **44(4)**, 1288-1299.
165. **Derjaguin B.V., Landau L. (1941)** Theory of the stability of strongly charged lyophobic sols and the adhesion of strongly charged particles in solutions of electrolytes. *Acta Physicochim. USSR*, **14**, 633-662.
166. **Verwey E.J.W., Overbeek J.T.G. (1948)** Theory of stability of lyophobic colloids. *Amsterdam: Elsevier*
167. **Gunton J.D., Shiryayev A., Pagan D.L. (2007)** Protein condensation, *Cambridge university press*
168. **Yong W., Lomakin A., Kirkitadze M.D., Teplow D.B., Chen S.H., Benedeck G.B. (2002)** Structure determination of micelle-like intermediates in amyloid b-protein fibril assembly by using small angle neutron scattering. *Proc. Natl. Acad. Sc. USA.* **99(1)**, 150-154.
169. **Clark P.L. (2004)** Protein folding in the cell: reshaping the folding funnel. *T.I.B.S.*, **29(30)**, 527-534.

## Résumé

Le repliement des protéines est un des problèmes centraux de la biologie. Il s'agit de comprendre comment la chaîne polypeptidique d'une protéine se replie pour acquérir sa structure tridimensionnelle biologiquement active. Il a été démontré dans les années 60 que la forme repliée de la protéine est le plus stable d'un point de vue thermodynamique et qu'il est défini par la structure primaire. La réaction de repliement correspond ainsi à la dernière étape de l'utilisation de l'information contenue dans l'ADN. Cependant, Il est possible que les protéines se replient mal et interagissent entre elles pour former des fibres amyloïdes. Ce sont des agrégats structurés impliqués dans plusieurs maladies comme la maladie d'Alzheimer, de Parkinson....

Ces phénomènes sont étudiés ici dans le cas de l' $\alpha$ -lactalbumine, une protéine du lait qui possède un site de liaison pour le calcium. Le repliement est tout d'abord étudié en présence de métaux se liant au site du calcium. Ces expériences sont couplées à des expériences de dénaturation thermiques pour caractériser le rôle de la fixation des métaux sur les différents états de la protéine et son influence sur la cinétique de repliement.

La réaction est ensuite caractérisée en absence d'ion métallique. Elle est alors beaucoup plus lente et complexe. Différentes techniques spectroscopiques sont utilisées. Les résultats obtenus permettent de proposer un schéma réactionnel selon lequel un état précurseur de fibres amyloïdes est transitoirement peuplé. Enfin, pour compléter cette étude, les effets des interactions entre protéines sur la formation de fibres amyloïdes ont été étudiés pour différentes concentrations en sel.

Mots clés : Repliement des protéines, fibres amyloïdes,  $\alpha$ -lactalbumine, cofacteur métallique, spectroscopie, paysage d'énergie

## Abstract

Protein folding is one of the central questions of Biology. How does polypeptidic chain fold to its three-dimensional, biologically active, structure? Experiments done in the 60's have shown that the folded form is the most stable one on a thermodynamic point of view. This form being defined by the primary structure of the protein, the folding reaction corresponds to the last step of the use of the information encoded in DNA. Proteins can sometimes misfold and form amyloid fibrils through intermolecular interactions. These structured aggregates are involved in several diseases such as Alzheimer's disease, Parkinson diseases....

These phenomenon's are studied here in the case of  $\alpha$ -lactalbumin, a milk protein which possesses a calcium binding site. At first, the folding is monitored in presence of various metal ions that bind to the calcium site. These experiments are coupled with thermal unfolding experiments. They permit to precise the effect of metal ion binding on the different states of the protein and on the folding kinetics.

The reaction is also monitored in the absence of metal ions with several biophysical methods. Then, the folding is much slower and intricate. A reaction scheme is drawn from the results. This scheme indicates that a state precursor of amyloid fibril is transiently populated during the reaction. Finally, the effect of intermolecular interactions on the formation of amyloid fibrils is characterized for different salt concentrations.

Keywords: Protein folding, amyloid fibrils,  $\alpha$ -lactalbumin, metallic cofactor, spectroscopy, energy landscape