



Caractérisations biochimiques et structurales de la γ E cristalline de rat, hydrogénée et perdeutériée en vue d'une étude de diffraction des neutrons.

Jean-Baptiste Artero

► To cite this version:

Jean-Baptiste Artero. Caractérisations biochimiques et structurales de la γ E cristalline de rat, hydrogénée et perdeutériée en vue d'une étude de diffraction des neutrons.. Sciences du Vivant [q-bio]. Université Joseph-Fourier - Grenoble I, 2005. Français. <tel-00207936>

HAL Id: tel-00207936

<https://tel.archives-ouvertes.fr/tel-00207936>

Submitted on 18 Jan 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.

Thèse

Présentée par

Jean Baptiste Artero

Pour obtenir le titre de

Docteur de l'université Joseph Fourier - Grenoble I

Spécialité Biologie Structurale et Fonctionnelle

(Arrêtés ministériels du 5 juillet 1984 et du 30 Mars 1992)

**Caractérisations biochimiques et structurales
de la γ E cristalline de rat, hydrogénée et perdeutériée
en vue d'une étude de diffraction des neutrons.**

Thèse soutenue publiquement le 23 mars 2005

devant le jury composé de :

- Mr Peter Timmins	Directeur de thèse
- Mr Sean Mc Sweeney	Directeur de thèse
- Mr Giuseppe Zaccai	Président du jury
- Mr Peter Lindley	Rapporteur
- Mr Yves Mechullam	Rapporteur
- Mr Rob Ruigrok	Examineur

Thèse préparée au sein de l'Institut Laue Langevin (ILL) et de l'installation européenne de rayonnement synchrotron (ESRF).

Abstract

All vertebrate eye lenses are transparent and have a refractive power depending on a smooth refractive index gradient for visible light. This is achieved by the regular arrangement of the fibre cell and by the differential expression of lens specific proteins, the crystallins. The increasing refractive index from the cortex to the nuclear region is associated with increasing concentration of crystallins relative to water. The nuclear region, where the refractive index is the highest, is enriched with γ crystallins, a family of monomeric polypeptides synthesised mainly during the early development, that play a major role in forming the closely packed medium. However, the high concentration is near a critical point whereby very soluble proteins undergo a low energy phase separation driven by the competing forces of protein-water, water-water and protein-protein interactions. This phase separation is crucial in the opacification of lens fibre cells found in certain forms of mammalian cataract.

A high resolution neutron diffraction study can supplement the existing X-ray data and allows us to make detailed and thorough analysis of the solvent organisation around the crystalline molecule. This will provide a structural base for studying the effect on protein-water, water-water and protein-protein interactions. This study will be greatly facilitated by the use of fully deuterated protein. Substituting deuterium for hydrogen will reduce the incoherent scattering background, strengthen the high resolution diffraction data and provide an order of magnitude gain in signal to noise. The exchange of the hydrogen by deuterium can be achieved by soaking protein crystals in deuterated liquor but these exchangeable hydrogens represent only 25% of total hydrogens of the protein. The exchange of all hydrogen atoms can only be achieved by an *in vivo* biosynthesis using a bacterial culture in a deuterated medium (deuterated glycerol and D₂O).

γ E crystalline from rat has been chosen as a model protein system to locate hydrogen atoms of water molecules surrounding the protein and hydrogens involved in different stabilising interactions. Large quantities of the hydrogenated and perdeuterated protein were expressed in *Escherichia coli* grown in minimal medium and then purified. The level of the isotope substitution on non-exchangeable sites of the protein was found to be 98% by electrospray ionisation mass spectrometry.

In the absence of known biochemical activity, the hydrogenated and deuterated γ E crystallins were characterised by non denaturing gel electrophoresis, isoelectric point determination and limited proteolysis. There are no major biochemical differences between these two forms of protein. Further characterisations by Fourier transform infrared (FTIR) and circular dichroism (CD) spectroscopies did not show any significant differences.

The hydrogenated and deuterated proteins were crystallised in H₂O and D₂O buffers. Crystallisation conditions, space groups and cell parameters were found to be the same for all forms of the protein. Comparison of these four forms of γ E crystallin revealed no significant structural difference between them at the atomic resolution around 1.4Å. However, the temperature factor variation of the structures at identical resolution (H γ Eh₂o H γ Ed₂o and D γ Eh₂o) depends on the isotope labelling. The structures of the deuterated protein in H₂O or the hydrogenated one with D₂O solvent have a lower temperature factor. The molecular model of γ E crystalline has been obtained at 1.36Å and some new relevant details on the structure and water network surrounding the protein are described in this report.

Neutron diffraction data collection cannot be carried out before crystal size improvement, but most importantly, we have shown that perdeuteration itself does not alter the structural features of the protein.

Key words

Protein perdeuteration
 γ E crystallin
Neutron and X-ray diffraction
Physicochemical properties
X-ray structure

Résumé

Tous les cristallins des vertébrés sont transparents et ont un pouvoir de réfraction. Ceci est dû à l'arrangement régulier de cellules et à l'expression différentielle de protéines spécifiques, les cristallines. L'indice de réfraction croissant du cortex vers la région nucléaire est associé à l'augmentation de la concentration en cristallines. La région nucléaire, où l'indice de réfraction est le plus haut, est enrichie en γ -cristallines, une famille de polypeptides monomériques synthétisés principalement pendant le développement précoce, qui jouent un rôle important en formant un ensemble de protéines très dense. Cependant, cette concentration élevée est près d'un point critique où les protéines très solubles subissent une séparation de phase, conduite par une concurrence de forces d'interactions protéine-eau, eau-eau et protéine-protéine. Cette séparation de phase est cruciale dans l'opacification des cellules du cristallin trouvée dans certaines formes de cataracte mammifère.

Une étude de diffraction de neutrons à haute résolution peut compléter les données existantes aux rayons X et peut nous permettre de faire une analyse détaillée et complète de l'organisation du solvant autour de la γ -cristalline. Ceci fournira une base structurale pour étudier les interactions protéine-solvant, solvant-solvant et protéine-protéine. Cette étude peut être considérablement facilitée par l'utilisation de la protéine entièrement deutériée. La substitution du deutérium à l'hydrogène réduit le bruit de fond incohérent causé par les hydrogènes, renforçant alors le rapport signal sur bruit. L'échange de l'hydrogène par le deutérium peut être fait en trempant le cristal de protéine dans une solution deutériée mais seulement 25% de hydrogènes totaux de la protéine sont alors échangés. Un échange total peut être uniquement effectué par une biosynthèse protéique *in vivo*, en utilisant une culture bactérienne dans un milieu deutérié.

La γ E cristalline de rat a été choisie comme un système modèle pour localiser les atomes d'hydrogène des molécules d'eau entourant la protéine et ceux impliqués dans des interactions stabilisantes. De grandes quantités de protéines hydrogénées et perdeutériées ont été exprimées dans *Escherichia coli*, qui a poussé dans un milieu minimal, et elles ont ensuite été purifiées. Par spectrométrie de masse, le niveau de substitution isotopique des hydrogènes non échangeables dans la protéine s'est avéré être de 98%.

En l'absence d'activité biochimique connue, les γ E cristallines hydrogénées et deutériées ont été caractérisées par gel natif, par détermination du point isoélectrique et par des protéolyses limitées. Il n'y a aucune différence biochimique évidente entre ces formes de la protéine. D'autres caractérisations, par spectroscopie infrarouge et en dichroïsme circulaire, ont été employées pour étudier des différences significatives, mais aucune n'a été décelée.

Les protéines hydrogénées et deutériées ont été cristallisées dans des tampons hydrogénés et deutériés. Les conditions de cristallisation, les groupes de l'espace et les paramètres de maille se sont avérés être les mêmes pour toutes les formes de la protéine. La comparaison de ces quatre formes de γ E cristalline n'a indiqué aucune différence structurale évidente entre elles à une résolution atomique aux alentours de 1,4Å. Cependant, la variation de facteur d'agitation thermique des structures à une résolution identique ($H\gamma E_{H_2O}$, $H\gamma E_{D_2O}$ et $D\gamma E_{H_2O}$) dépend du marquage d'isotopique. Les structures de la protéine deutériée en H_2O ou celle hydrogénée avec un tampon D_2O ont un facteur d'agitation thermique plus bas. Par la suite, le modèle moléculaire de la γ E cristalline a été obtenu à 1,36Å et quelques nouveaux détails étonnants de la protéine et du réseau de molécule d'eau entourant la protéine sont décrits dans ce rapport.

Jusqu'à l'amélioration de la taille des cristaux, toute collecte de données par diffraction de neutrons ne peut pas être envisagée, mais d'une manière globale, nous avons prouvé que la perdeutériation n'induisait pas de changement structural de la protéine.

Mots clés

Perdeutériation de protéine
 γ E cristalline
Diffraction de neutron et de rayon X
Propriétés physico-chimiques
Structure de rayon X.

Remerciements

Mon passage à l'ILL et à l'ESRF a été une expérience très enrichissante et beaucoup de personnes ont contribué que ce soit de près ou de loin, aux travaux décrits ici, ainsi qu'à la composition de cette thèse. Il m'est impossible de toutes les mentionner personnellement sans en oublier ; c'est pourquoi j'espère que ces personnes ne m'en voudront pas pour cet oubli involontaire.

Tout d'abord, je souhaite remercier mes deux superviseurs, Peter Timmins et Sean Mc Sweeney, qui m'ont donné ma chance avec ce projet. Je suis extrêmement reconnaissant pour tout le temps qu'ils ont passé à essayer de m'apprendre la cristallographie que ce soit des rayons X ou des neutrons, pour leur enthousiasme constant sur ce sujet, pour les réponses à mes questions (peut être pas assez nombreuses), et surtout pour leur patience.

Ensuite je tiens à remercier Michael Haertlein qui a directement participé à ce projet et surtout pour toutes nos nombreuses discussions qui m'ont éclairé au cours de cette thèse.

Je remercie les membres du jury, Peter Lindley, Yves Mechullam, Giuseppe Zaccai et Rob Ruigrok d'avoir accepté de faire partie du jury, et tout particulièrement Peter Lindley et Yves Mechullam d'avoir accepté d'être rapporteurs de ce travail.

Pendant ces trois années de thèse (voire un peu plus), j'ai eu l'occasion de collaborer avec David Lemaire, avec qui, c'était toujours un véritable plaisir d'effectuer les nombreux tests de spectrométrie de masse, Vincent Forge qui m'a initié à la spectroscopie FTIR, Monika Budayova Spano pour ces nombreux conseils en cristallogenèse, et surtout à Susana Teixeira, pour avoir accepté d'intégrer le monde des γ cristallines. A vous tous, je vous remercie pour votre disponibilité et votre patience, en espérant que ces collaborations continuent.

Je souhaite remercier, aussi, les nombreuses personnes des différents laboratoires que j'ai arpentés pour leur patience, pour leur humour, et pour leurs discussions que cela soit de nature techniques ou tout autre. Je souhaite en particulier dire merci à Marie-Thérèse Dauvergne pour son attention constante qu'elle soit personnelle ou professionnelle, et qui m'a convertie aux cristallines, à Flora Meilleur, pour toutes ses ouvertures d'esprits sur les neutrons, à Ingrid Parrot pour sa bonne humeur et son humour toujours présents, à Martine

Moulin qui m'a écouté et m'a souvent conseillé, et à Benoit Gallet, qui m'a laissé tomber pour rejoindre l'IBS.

Ingar Leiros et Gordon Léonard m'ont apporté beaucoup de conseils lors de l'enregistrement des données expérimentales. A l'époque, mes très incomplètes connaissances de la diffraction aux rayons X ne m'auraient pas permis de faire seul les bons choix.

Je remercie sincèrement les personnes de l'EMBL et de l'ESRF, qui m'ont apporté leurs conseils en cristallographie : Nicolas Tarbouriech, Cyril Dian, Yvain Nicolet, Carlo Petosa, Win Burmeister, Carlos Fernandez Tornero...

Un grand merci, à Karine Sultan, Rémi Pink, Benjamin Girardier, Mary Jane Villot, Virginie Bertholet, Annie Simon, Sandrine Mario, Monique Romero, Denis Pognant, Jose Barranco, et François Dauvergne pour leur dévouement dans leurs travaux respectifs, sans lesquels la science serait bien peu de chose.

Je tiens à ne pas oublier tous les thésards et amis de l'ILL, de l'EMBL, de l'IAB, et de l'ESRF, pour leur conseils, scientifiques ou non, pour leur soutien, leur réconfort et pour toute leur aide : Marlyse, Thibault, Nicolas, Lucie, Aurélie, Jeanne, Franck, Philippe, Cédric, Céline, Annabelle, Cyril(s), Carlos, Momo, Barry, Alex, Titi, Benoit(s), Anne Laure, Gilles, Mike, Gwladys...

Enfin, ceux qui comptent le plus à mes yeux, je tiens à remercier Armelle et Valentine, ainsi que mes parents, pour leur soutien moral et orthographique.

“Advances are made by answering questions.
Discoveries are made by questioning answers.”
- Bernhard Haisch, astrophysicist -

Sommaire

Chapitre 1 : Présentation du sujet

	Page
Abstract	1
Résumé	3

Chapitre 2 : Œil, cristallin, γ cristallines et neutrons

Abstract	6
2.1 Le système protéique : les γ cristallines	7
2.1.1 De l'œil à la vision	7
2.1.1.1 L'œil	7
2.1.1.2 Le trajet optique	8
2.1.1.3 le cristallin : un organe exceptionnel	9
2.1.1.3.1 L'organisation du cristallin	9
2.1.1.3.2 Les fibres du cristallin	11
2.1.1.3.3 Réfraction et transparence	12
2.1.2 Les cristallines	14
2.1.2.1 Les cristallines : des protéines	14
2.1.2.2 Les cristallines : les protéines essentielles du cristallin	16
2.1.3 Les α cristallines	16
2.1.3.1 Les α cristallines : « small Heat Shock Protein » et autres fonctions	17
2.1.3.2 Les α cristallines : des oligomères hétérogènes de grande taille	17
2.1.3.3 Les α cristallines : cataractes et maladies	17
2.1.4 La super famille des β/γ cristallines	18
2.1.4.1 La clef grecque : dénominateur commun des β/γ cristallines	18
2.1.4.2 Les β cristallines	19
2.1.4.3 La protéine γ S cristalline	21
2.1.4.4 Les γ cristallines	21
2.1.4.4.1 Gènes de γ cristallines	22
2.1.4.4.2 Maladies	22
2.1.4.4.3 Transition de phase des γ cristallines	23
2.1.4.4.4 Structures des γ cristallines cryo et non cryo-précipitables	24
2.1.4.5 Evolution de la super famille β/γ cristalline	29
2.1.5 L'eau et les γ cristallines	34
2.1.5.1 L'eau et la biologie	33
2.1.5.2 L'hydratation des γ cristallines	35
2.2 Pourquoi envisager la cristallographie aux neutrons	36
2.2.1 La visualisation des hydrogènes par la cristallographie aux neutrons	38
2.2.2 Diffusion incohérente des atomes d'hydrogène	40
2.2.3 La réduction du bruit de fond en cristallographie des neutrons	41
2.2.4 Faible flux de neutrons	42
2.2.5 La diffraction polychromatique dite diffraction de Laue	43
2.3 Bilans	45
Bibliographie	46

Chapitre 3 : Production de γ cristallines totalement deutériées

Abstract	49
3.1 Expression de γ cristallines en bio-réacteur	50
3.1.1 Notions essentielles pour les cultures cellulaires	51
3.1.1.1 Apports nutritifs et conditions de culture	51
3.1.1.1.1 La source de carbone	52
3.1.1.1.2 Cas de l'oxygène	52
3.1.1.1.3 Effet de la température	53
3.1.1.1.4 Le pH	54
3.1.1.2 Paramètres de la croissance	54
3.1.1.3 Les phases de croissance : cas d'une culture en « batch »	56
3.1.1.4 La culture en « fed batch »	57
3.1.1.5 La production de protéine recombinante	58
3.1.2 La culture en milieu deutérié	59
3.2 Matériels et méthodes	60
3.2.1 Clones de γ cristallines disponibles	60
3.2.2 Adaptation au milieu minimum deutérié	61
3.2.3 Expression en fermenteur	61
3.3 Résultats et discussions	63
3.3.1 Choix du plasmide	63
3.3.2 Adaptation des bactéries à la croissance en D ₂ O	64
3.3.3 Culture en mode « batch fed batch »	65
3.3.3.1 La phase de « batch »	67
3.3.3.2 La phase de « fed batch »	67
3.3.3.3 La phase d'induction	68
3.3.4 Comparaisons cultures hydrogénées et deutériées	69
3.4 Bilans	71
Bibliographie	73

Chapitre 4 : Caractérisations biochimiques et spectrales

Abstract	74
4.1 Caractérisations Biochimiques	75
4.2 Matériels et méthodes	76
4.2.1 Préparations des échantillons	76
4.2.1.1 La protéine γ cristalline exprimée avec queue poly-histidines	76
4.2.1.2 La protéine γ cristalline exprimée sans queue poly-histidines	76
4.2.1.3 Préparation de la protéine en milieu deutérié (D ₂ O)	77
4.2.2 Spectrométrie de masse	77
4.2.3 Clonage de la Méthionine Amino Peptidase (MAP) dans le pLysS, expression et purification de la γ E cristalline	78
4.2.4 Gel natif	79
4.2.5 Détermination du point isoélectrique	79
4.2.6 Protéolyse des γ cristallines	80
4.2.7 Spectroscopie infrarouge par transformée de Fourier	80

4.2.8 Dichroïsme circulaire	81
4.3 Résultats et discussions	83
4.3.1 Lyse cellulaire	83
4.3.2 Purification de la protéine avec une queue poly-histidines	85
4.3.3 Purification de γ cristalline sans queue poly-histidines	85
4.3.4 Evaluation de l'incorporation isotopique par spectrométrie de masse	86
4.3.5 Gel natif	89
4.3.6 Point isoélectrique	90
4.3.7 Protéolyse limitée	91
4.3.8 Spectroscopie infrarouge par transformée de Fourier	94
4.3.8.1 Attribution des structures secondaires par FTIR	95
4.3.8.2 Comparaison des données infrarouge avec d'autres analyses hydrogénées	98
4.3.9 Dichroïsme circulaire	100
4.3.9.1 Dénaturation chimique	101
4.3.9.2 Dénaturation thermique	102
4.3.10 La stabilité des protéines	104
4.3.11 Effets isotopiques du solvant sur la stabilité des protéines	107
4.3.12 Effet de la perdeutériation sur la stabilité	109
4.4 Bilans	111
Bibliographie	113

Chapitre 5 : La cristallographie aux rayons X : rappels, collectes et traitements des données

Abstract	115
5.1 La cristallographie aux rayons X	116
5.2 Rappels sur la cristallographie aux rayons X	117
5.2.1 Principe de la diffusion ou de la diffraction	117
5.2.2 Principe de diffraction des rayons X par un cristal	118
5.2.2.1 Un cristal	118
5.2.2.2 Principe de la diffraction par un cristal	119
5.2.2.3 La loi de Bragg	120
5.2.2.4 Notion d'espace réciproque et de réseau réciproque	120
5.2.2.5 La sphère d'Ewald	122
5.2.3 Traitements des données	123
5.2.3.1 Des images de diffraction aux données de diffraction	123
5.2.3.2 Résolution du problème de phase	125
5.2.3.3 Reconstruction et affinement	126
5.2.3.3.1 La reconstruction	127
5.2.3.3.2 Les méthodes numériques d'affinement	127
5.2.3.3.3 Critères de validité d'un affinement	128
5.3 Matériels et méthodes	129
5.3.1 Cristallisation des γ cristallines	129
5.3.2 Optimisation des solutions cryogéniques	130
5.3.3 Collectes des données cristallographiques	131
5.3.3.1 Le synchrotron : source de rayons X	131
5.3.3.2 Diffraction des rayons X	132

5.3.4 Traitements des données	133
5.3.4.1 Traitements des images	133
5.3.4.2 Résolution du problème de phase	133
5.3.4.3 Reconstruction et affinement	134
5.3.4.4 Validation des structures	134
5.4 Résultats et discussions	135
5.4.1 Condition de cristallisation	135
5.4.2 Collectes des données	137
5.4.3 Traitements des données	139
5.4.4 Remplacement moléculaire	139
5.4.5 Reconstructions et affinements	141
5.4.6 Qualité des modèles	145
5.5 Bilans	145
Bibliographie	146

Chapitre 6 : Structure de la γE cristalline à 1,32 Å et comparaison des formes hydrogénées et deutériées en H₂O et en D₂O

Abstract	147
6.1 Introduction	148
6.2 Comparaison des différentes structures résolues	149
6.2.1 Surfaces et volumes	149
6.2.2 Déviation des chaînes	150
6.2.3 Facteur thermique	153
6.3 Structure de la γE cristalline	158
6.3.1 Quatre motifs en clef grecque organisés symétriquement	158
6.3.2 Stabilisation de la structure par interactions entre les deux domaines	160
6.3.3 Contacts cristallins	161
6.3.4 Boucle entre les motifs <i>cd</i> dans le motif en clef grecque 3	164
6.3.5 Le tri peptide et la partie C terminale	169
6.3.6 Autres détails sur la structure	172
6.3.6.1 Analyse des résidus cystéines	172
6.3.6.2 Empilement des résidus aromatiques	173
6.4 Analyse de la structure du solvant	174
6.4.1 Comparaison du solvant entre les différentes structures	175
6.4.2 Les liaisons hydrogène avec la protéine et la solvatation	177
6.4.3 L'eau autour des résidus chargés	179
6.4.4 Distribution des molécules d'eau	181
6.4.5 Réseaux de molécules d'eau	183
6.5 Bilans	184
Bibliographie	187

Chapitre 7 : Cristallisation de la γ E cristalline, conclusions et perspectives

Abstract	189
7.1 Des cristaux de grande taille pour les neutrons	190
7.2 Cristallisation de la γ E cristalline	190
7.2.1 Méthodes de diffusion de vapeur et de dialyse	190
7.2.2 Obtention de cristaux uniques	192
7.2.2.1 Ensemencement	192
7.2.2.2 Cristallisation en capillaire par la technique d'acuponcture	192
7.2.3 Des cristaux en trois dimensions	193
7.3 Conclusions et perspectives	195
Bibliographie	199

Annexes

Annexe A : Intégration du glycérol dans la glycolyse	200
Annexe B : Constructions plasmidiques et analyse de séquence	201
Annexe C : Milieux de culture	204
Annexe D : Tampons de purification	207
Annexe E : Structure des γ E cristallines	209
Annexe F : Concentration de la γ E cristalline dans un cristal	220

Listes des figures

Page

Chapitre 2 : Œil, cristallin, γ cristallines et neutrons

2.1	Représentation antique du système oculaire	7
2.2	Schéma d'une coupe transversale d'un œil humain	8
2.3	Schéma d'une coupe transversale d'un cristallin humain	10
2.4	Représentation tridimensionnelle d'une coupe d'un cristallin	11
2.5	Couches superficielles du cristallin	12
2.6	Empilement hexagonale et régulier des fibres du cristallin	13
2.7A	Poterie grecque ayant le motif de la clef grecque	18
2.7B	Super structure secondaire de la clef grecque	18
2.8	Structure de la β B2 cristalline	20
2.9	Exemples de cataractes aculéiformes et de Coppock	23
2.10	Courbes de température de transition	24
2.11	Alignement des séquences primaires des γ cristallines	25
2.12	Comparaison tri dimensionnelles des γ cristallines	27
2.13	Comparaison de la boucle entre les brins c et d du motif 3	28
2.14	Représentations de la surface moléculaire de γ F cristalline	28
2.15	Structure des différents constituants des β/γ cristallines	30
2.16	Arrangements des séquences codantes des protéines assimilées à la famille des β/γ cristallines	31
2.17	Résumé des évènements de l'évolution de la famille des β/γ cristallines	33
2.18	L'eau et la γ B cristalline	34
2.19	Facteurs de formes en cristallographie aux rayons X et aux neutrons	39
2.20	Visualisation des hydrogènes	40
2.21	Cliché de diffraction quasi-Laue	43
2.22	Construction d'Ewald pour la diffraction de Laue	44

Chapitre 3 : Production de γ cristallines totalement deutériées

3.1	Courbe de croissance d'une culture en « batch »	56
3.2	Courbe de croissance d'une culture en « fed batch »	57
3.3	Représentation schématique de l'adaptation au D ₂ O d' <i>E. coli</i>	61
3.4	Photographie d'une unité de fermentation	62
3.5	Profil de fermentation d' <i>E. coli</i>	66
3.6	Expression de la protéine recombinante γ E cristalline	69
3.7	Comparatif des courbes de croissance hydrogénée et deutérée	70

Chapitre 4 : Caractérisations biochimiques et spectrales

4.1	Première étape de purification de la γ E cristalline avec queue poly-histidines	83
4.2	Digestion de la queue histidine	84
4.3	Chromatographie d'exclusion de γ E cristalline avec queue poly-histidines	84

4.4	Première étape de purification de la γ E cristalline sans queue poly-histidines	85
4.5	Chromatographie d'exclusion de γ E cristalline sans queue poly-histidines	85
4.6	Spectres de masse de la γ E cristalline hydrogénée et deutérée	87
4.7	Co-expression de la γ E cristalline et de la MAP	88
4.8	Purifications de la γ E cristalline sans queue poly histidines exprimée avec la MAP	88
4.9	Spectre de masse de la γ E cristalline purifiée et exprimée avec la MAP	89
4.10	Gel natif de la γ E cristalline	90
4.11	Analyse en IEF	91
4.12	Digestion protéolytique de la γ E cristalline hydrogénée et deutérée par la subtilysine	91
4.13	Digestion protéolytique de la γ E cristalline hydrogénée et deutérée par l'élastase, la trypsine et la subtilysine	92
4.14	Digestion protéolytique de la γ E cristalline hydrogénée et deutérée par la pepsine en condition semi dénaturante	93
4.15	Spectres d'absorption infrarouge des γ E cristalline hydrogénées et deutérées	95
4.16	Spectres infrarouge et leurs dérivés secondes de l'amide I	96
4.17	Spectres de dichroïsme circulaire de la γ E cristalline	101
4.18	Courbes de dénaturation chimique	102
4.19	Courbes de dénaturation thermique	103
4.20	Représentation schématique d'une dénaturation réversible de protéine	104
4.21	Contributions relatives à l'énergie libre pour le repliement d'une protéine	105
4.22	Diagramme représentant les changements d'énergie libre lors de repliement et de dénaturation d'une protéine	107
4.23	Effets de la deutération sur la dénaturation de quelques protéines	110

Chapitre 5 : La cristallographie aux rayons X : rappels, collectes et traitements des données

5.1	Interaction d'une onde électromagnétique avec un objet	117
5.2	Représentation graphique de la loi de Bragg	120
5.3	Représentation de l'espace réciproque et des relations avec l'espace réel	121
5.4	Construction d'Ewald pour les rayons X	122
5.5	Détermination de la résolution par la surface du détecteur	123
5.6	L'E.S.R.F.	132
5.7	Représentation schématique d'une expérience de diffraction de rayons X par un cristal	133
5.8	Morphologie des différentes formes cristallines obtenues pour la γ E cristalline	135
5.9	Cristaux en 3D de la protéine γ E cristalline	136
5.10	Clichés de diffraction	138
5.11	Densité électronique bien définie autour du résidu W42	143
5.12	Qualité de la densité électronique $2 F_o - F_c $	144

Chapitre 6 : Structure de la γ E cristalline à 1,32 Å et comparaison des formes hydrogénées et deutériées en H₂O et en D₂O

6.1	RMS déviation de la chaîne principale des γ E cristalline de rat étudiées	151
6.2	RMS déviation des chaînes latérales des γ E cristalline de rat étudiées	152
6.3	Représentation de la chaîne principale suivant le facteur <i>B</i>	154
6.4	Comparatif des facteurs <i>B</i> de la chaîne principale, des chaînes latérales et des résidus	154
6.5	Représentation schématique de l'existence des doubles conformations d'un point de vue énergétique	158
6.6	Structure générale des γ E cristalline résolues et leurs symétries internes	159
6.7	Représentation du motif 1 en clef grecque	159
6.8	Surfaces d'interaction entre les deux domaines N et C terminaux	160
6.9	Empilement cristallin de la protéine recombinante deutérée	161
6.10	Représentation des zones de contact cristallin	162
6.11	Exemple de contact cristallin entre deux équivalents cristallins	162
6.12	Empilement cristallin de façon « tête bêche »	164
6.13	Représentation de la double conformation des résidus Phe115 et His116	165
6.14	Représentation de la double conformation de la boucle entre les brins <i>c</i> et <i>d</i> du motif 3	165
6.15	Représentation des boucles <i>cd</i> du motif 3 des γ cristallines	167
6.16	Déviation des C α des résidus 90 à 140	168
6.17	Représentation du tripeptide RGD et des acides aminés proches	170
6.18	Empilement des tyrosines 16 et 28 et de la proline 27	173
6.19	Comparaison des facteurs <i>B</i> et de la déviation moyenne des 98 molécules d'eau identiques dans tous les modèles	176
6.20	Organisation des molécules d'eau autour des résidus chargés	181
6.21	Représentation des molécules d'eau dispersées autour de la γ E cristalline	182
6.22	Représentation de l'extension du motif en clef grecque	182
6.23	Représentation d'une des deux molécules d'eau les plus enfouies dans la protéine	183
6.24	Schéma des 5 plus grands réseaux de molécules d'eau de la γ E cristalline	184

Chapitre 7 : Cristallisation de la γ E cristalline, conclusions et perspectives

7.1	Cristallisation par diffusion de précipitant ou de vapeur	191
7.2	Diagramme de solubilité	191
7.3	Différentes étapes de cristallisation de la γ E cristalline	193
7.4	Système de cristallisation en capillaire	193
7.5	Induction de la croissance cristalline de la γ E cristalline par variation de température	194

Sommaire des tableaux

Page

Chapitre 2 : Œil, cristallin, γ cristallines et neutrons

2.1	Répartition de l'indice de réfraction	12
2.2	Structures connues des γ cristallines	26
2.3	Caractéristiques des principaux atomes constitutifs des protéines	38

Chapitre 3 : Production de γ cristallines totalement deutériées

3.1	Comparatif des prix des produits deutérés pouvant être utilisés en fermentation	50
3.2	Temps de doublement moyen de culture de différentes souches	59
3.3	Comparaison des résultats de fermentations hydrogénées et deutériées	70

Chapitre 4 : Caractérisations biochimiques et spectrales

4.1	Spécificité, tampons et conditions de protéolyses ménagées	80
4.2	Attribution des bandes infrarouge de l'amide I	97
4.3	Estimation en pourcentage des structures secondaires de différentes γ cristallines	99
4.4	Différentes interactions non covalentes	105
4.5	Caractéristiques physico-chimiques de l'eau légère et de l'eau lourde	108

Chapitre 5 : La cristallographie aux rayons X : rappels, collectes et traitements des données

5.1	Les 7 systèmes cristallins	119
5.2	Conditions de cristallisation ayant donné des cristaux	130
5.3	Paramètres de différentes collectes de données	139
5.4	Solutions proposées pour le remplacement moléculaire	140
5.5	Facteurs R et R_{libre} après le premier affinement	141
5.6	Statistiques des collectes de données et des affinements finaux	142
5.7	Statistiques issues des diagrammes de Ramachandran	145

Chapitre 6 : Structure de la γ E cristalline à 1,32 Å et comparaison des formes hydrogénées et deutériées en H₂O et en D₂O

6.1	Comparatif des surfaces et des volumes des structures de la γ E cristalline	149
6.2	Nombres d'acides aminés ayant une double conformation dans les différentes structures	156
6.3	Occupation des différentes conformations des résidus ayant 2 conformations dans les différents modèles	157
6.4	Appariements des résidus de contacts dans l'empilement cristallin	163

6.5	Alignements des acides 117,120 et 124 des γ cristallines caractérisées pour les hautes et basses températures de transition	166
6.6	Nombre de molécules de solvant identifiées	175
6.7	Distribution des molécules d'eau en fonction de leur nombre de coordination	177
6.8	Distribution des molécules d'eau en fonction de leur nombre de liaisons à la protéine	178
6.9	Répartition des liaisons hydrogène en fonction de leur interaction avec les atomes de la protéine et du solvant	179
6.10	Contribution des différents résidus polaires ou chargés dans l'hydratation de la γ E cristalline	180

Abréviations et symboles

Å : Ångström (1Å = 0,1 nm)

α : alpha

β : bêta

γ : gamma

λ : Longueur d'onde d'une radiation monochromatique

Γ : Temps de doublement d'une culture

μg : Microgramme

μl : Microlitre

μM : Micromolaire

Ac : Acétate

ADN : Acide désoxyribonucléique

ADNc : Acide désoxyribonucléique complémentaire

Amp : Ampicilline

ARN : Acide ribonucléique

CCP4 : Collaborative Computational Project numéro 4

CD ou DC : « circular dichroïsm » dichroïsme circulaire

CEA : Commissariat à l'énergie atomique

CO₂ Dioxyde de carbone

D₂O : Eau lourde

Da : Dalton

DTT : Dithiothréitol

DγEd₂o : γE cristalline deutériée en eau lourde

DγEh₂o : γE cristalline deutériée en eau légère

E.coli : *Escherichia coli*

EDTA : Disodium dihydrogéoéthylènediamine tétraacétate

EMBL : Laboratoire européen de biologie moléculaire

ESRF : European Synchrotron Radiation Facility

$F(hkl)$: Facteur de structure pour les réflexions hkl représenté comme un vecteur

Facteur B : facteur thermique

Facteur R : Facteur indiquant le niveau de désaccord entre les structures observées en diffraction et les calculs effectués à partir du modèle moléculaire construit

Facteur R_{libre} : Facteur indiquant le critère de désaccord « objectif » sur un jeu de données

F_c : Amplitude du facteur de structure calculé

F_o : Amplitude du facteur de structure observé

FTIR : « Fourier Transform InfraRed » spectroscopie infrarouge par transformée de Fourier

g/L : Gramme par litre

GK : « Greek Key » motif en clé grecque

H₂O : Eau légère

HCDC : « High Cell Density Culture » culture de haute densité cellulaire

HEPES : Acide sulfonique n-2-hydroxyéthylpipérazine-n-2-éthane

HSP : « Heat shock Protein » protéine du choc thermique

H γ Ed₂o : γ E cristalline hydrogénée en eau lourde

H γ Eh₂o : γ E cristalline hydrogénée en eau légère

IBS : Institut de Biologie Structurale

IEF : « IsoElectric Focusing »

ILL : Institut Laue Langevin

IPTG : Isopropyl- β -D-thiogalactoside

Kana : kanamycine

kDa : kiloDalton

LADI : LAue Diffractomètre

LB: Milieu Luria Bertani

MAP : Méthionine Amino Peptidase

MeOH : méthanol

mg : Milligramme

MgSO₄ sulfate de magnésium

ml : Millilitre

NaCl : Chlorure de sodium

O₂ : Dioxygène

°C : Degré celcius

OD : « Optical density » densité optique

P. pastoris : *Pichia pastoris*

pb : Paire de base d'ADN

pD : «power of deuterium-power of the concentration of deuterium ion» mesure de l'acidité ou de l'alcalinité d'une solution deutériée

PEG : Polyéthylène glycol

pH : «Power of hydrogen-power of the concentration of hydrogen ion» mesure de l'acidité ou de l'alcalinité d'une solution hydrogénée

r.m.s.d.: « Root mean square deviation », déviation des moindres carrés

RGD : Arginine Glycine Aspartate

RMN : Résonance magnétique nucléaire

SDS : Sodium dodécyl sulfate

SDS-PAGE : SDS- «page polyacrylamide gel electrophoresis» gel d'électrophorèse de polyacrylamide

TFA : Acide trifluoroacétique

Tris: 2-amino-2-(hydroxyméthyl)propane-1,3-diol

UV : Ultra violet

Chapitre 1

Présentation du sujet

Abstract

All vertebrate eye lenses are transparent and have a refractive power depending on a smooth refractive index gradient for visible light. This is achieved by a regular arrangement of the fibre cells and by a differential expression of lens specific proteins, the crystallins. The nuclear region of the lens is enriched at high levels with γ crystallins. This high concentration is near a critical point whereby very soluble proteins undergo a phase separation driven by the competing forces of protein-water, water-water and protein-protein interactions. This phase separation is crucial in the opacification of lens fibre cells found in certain forms of mammalian cataract.

X-ray crystallography and NMR spectroscopy are powerful methods for determining the 3D structure of proteins; however, it is difficult for such techniques to determine the positions of all the hydrogen atoms accurately. Because of their low scattering power for X-rays, hydrogens are not seen in protein structures except at very high resolution. The scattering of neutrons by hydrogen or deuterium is comparable to that of other atoms in a protein, with the result that neutron diffraction is an ideal method to obtain hydrogen atom positions in a protein. However this technique is limited in several ways. The intensity of even the strongest neutron sources today is several orders of magnitude weaker than a typical synchrotron X-ray source. Consequently, one needs large crystals, with a volume greater than 1 cubic mm, to produce reflections that are strong enough to be detected. Another technical problem, intrinsic to the non labelled samples, is the significant background produced by the large incoherent scattering of hydrogen atoms. As hydrogens can make up half of a protein, the incoherent background scattering from them can be dominating in a weakly diffracting system, making the measurement difficult or impossible. In previous studies, this was partially addressed by using crystals soaked in mother liquor containing D₂O. This method

leads in replacement of the exchangeable hydrogens with deuterium, and some specific hydrogens can be directly located. But exchangeable hydrogens account for only about 25% of the total hydrogen atoms. The only way to replace the nonexchangeable hydrogens by deuterium is the synthesis of the fully deuterated protein. This technique is developed at the ILL deuteration laboratory on bacterial (*E. coli*) and yeast (*P. pastoris*) expression systems.

I report in this thesis, a study based on the deuterium labelling of a γ -crystallin. This work describes expression of the perdeuterated protein, its biochemical characterisation in comparison with its hydrogenated counterpart, and a structural study by X-ray crystallography. In chapter 2, the biological system, the γ -crystallins, their importance and why neutron crystallography will be an advantage will be presented. To obtain a fully deuterated crystallin, the technique using a bacterial system will be presented, in chapter 3, with the adaptation to growth in deuterated minimal media and the use of a bio-reactor to compensate for low yields due to the isotope effects. Purification of the γ E-crystallin, biochemical and biophysical comparative studies will be described in chapter 4. It is crucial that there are no modifications of the proteins induced by the perdeuteration. Some notions of X-ray crystallography will introduce chapter 5 where the data collection and processing will represent the main part of this chapter. In chapter 6, structures of hydrogenated and deuterated γ E-crystallin crystallised in H₂O and in D₂O solvents will show details, not visible in the structure previously resolved at 2.3 Å. A comparative crystallographic analysis of the different isotopic forms will show the influence of the deuterium in a crystallographic study. To obtain huge crystals necessary for neutron crystallography, different techniques of crystallisation were tested and are presented in chapter 7 with conclusions and perspectives to close my work.

Résumé

Le cristallin normal d'un œil de vertébré possède des capacités physiques de transparence et de réfraction. Ces propriétés dépendent d'un gradient précis de l'indice réfractif. Cet indice est établi par l'arrangement des fibres du cristallin mais aussi par l'expression différentielle de protéines spécifiques, appelées les cristallines, par leur organisation et leurs interactions. Ainsi, au centre du cristallin, les γ cristallines sont présentes à de très hautes concentrations, proches d'un point critique où ces protéines solubles subissent une séparation de phase. Cette transition de phase est dirigée par une compétition entre des interactions protéine - protéine, protéine - eau (solvant) et eau - eau. Il est fortement supposé que les interactions eau - γ cristalline sont impliquées dans les séparations de phase conduisant à l'opacification du cytoplasme des fibres du cristallin observée dans certaines formes de cataractes. De plus, il a été aussi démontré que le contenu en eau du cristallin diminuait au cours de la vie. Donc si une partie des molécules d'eau qui entourent les γ cristallines disparaît, il est normal que la protéine partiellement déshydratée va être forcée à changer sa structure tertiaire, incluant sa distribution de charges. Ainsi, à cause de la faible concentration locale en eau, la protéine risque de devenir insoluble et, par la suite de former des agrégats. De ce fait, nous sommes en droit de nous poser la question : à quel endroit et comment les molécules d'eau jouent-elles un rôle dans l'organisation et la stabilisation des γ cristallines ?

De nombreuses techniques telles que la cristallographie aux rayons X et la spectroscopie RMN sont très puissantes pour déterminer les structures tridimensionnelles des protéines ; cependant, il est encore difficile d'utiliser uniquement ces méthodes pour déterminer avec précision la position des atomes d'hydrogène de la protéine mais aussi du solvant. A cause de leur faible pouvoir de diffusion des rayons X, les atomes d'hydrogène ne sont généralement pas visibles dans les structures déterminées à hautes résolutions ($< 1 \text{ \AA}$). En revanche, lors d'une expérience de diffraction neutronique, la diffusion par les hydrogènes (ou des deutériums) est comparable à celle des autres atomes présents dans une protéine. Ainsi, la diffraction neutronique est utilisée pour visualiser les positions atomes d'hydrogène. Cependant des cristaux de très grande taille ($>1 \text{ mm}^3$) sont nécessaires pour compenser le faible flux utilisable de neutrons. De plus, la large diffusion incohérente des hydrogènes réduit considérablement le signal en comparaison du bruit. Ce problème peut être surmonté en deutériant partiellement, par trempage des cristaux dans une solution deutériée, ou en

produisant des échantillons protéiques entièrement deutériés, c'est la perdeutériation. Cette dernière technique permet d'augmenter le rapport signal/bruit d'un facteur 10 et donc de visualiser sur les cartes de diffraction, les atomes d'hydrogènes de la protéine ainsi que les molécules d'eau dans leur totalité.

La deutériation partielle ne permet de remplacer qu'une faible partie des hydrogènes responsables de la diffusion incohérente (~ 25%) alors que la perdeutériation remplace la totalité de ces atomes d'hydrogènes. C'est pourquoi au laboratoire de deutériation de l'ILL/EMBL, sont développées des techniques de production et la purification de protéines deutériées recombinantes à partir de systèmes d'expression bactériens (*E. coli*) et lévuriens (*P. pastoris*).

L'étude que j'ai entreprise est basée sur le développement de la technique de perdeutériation, mais aussi sur des comparaisons à la fois biochimiques, spectroscopiques et cristallographiques (rayons X) des γ cristallines hydrogénées et deutériées. Le choix de ces protéines pour une étude cristallographique est dirigé par les caractéristiques d'interactions, décrites dans le chapitre 2, de ces protéines mais aussi par le fait que le contenu en solvant (l'eau) est sensiblement identique dans le cristallin et dans un cristal de protéine.

Dans cette thèse, est présentée une étude sur un système protéique : les γ cristallines où les molécules d'eau environnantes à la protéine sont d'une importance capitale pour la transparence et la réfraction de la lumière au niveau du cristallin des mammifères. Leur localisation, leurs caractéristiques et leurs importances sont développées dans le chapitre 2, ainsi que les besoins et ce que va nous apporter une étude neutronique sur de telles protéines.

De façon, à substituer tous les atomes de la protéine, nous avons développé des techniques de sur-expression de protéines recombinantes en système bactérien, et qui seront discutées dans le chapitre 3. Dans ce chapitre, la fermentation de *Escherichia coli* en milieu de culture hydrogéné et deutérié, et l'adaptation à la croissance en milieu deutérié de ces bactéries, vont être dissertés.

Après avoir développé les techniques de purification de la γ E cristalline sous forme hydrogénée et deutériée, le chapitre 4 sera consacré à la caractérisation biochimique des protéines purifiées. Les techniques employées vont mettre en évidence les similarités et les différences qu'a engendrées la deutériation. Il est important que le remplacement isotopique n'altère pas la conformation de la protéine, ainsi que ses possibles propriétés biologiques.

Dans le chapitre 5, les rappels de la cristallographie aux rayons X, ainsi que les collectes de données et le traitement de ces dernières sont exposés. Dans le chapitre 6, l'étude des structures de la protéine γ E cristalline sous forme hydrogénée et deutériée, cristallisée dans des solvants hydrogénés et deutériés a révélé certains détails de la structure qui n'étaient pas visibles avec la structure de la protéine native à 2,3 Å (Norledge et al. 1997). Une étude comparative sur l'influence que peut avoir une substitution isotopique, soit sur les hydrogènes échangeables, soit sur les non échangeables, est décrite au niveau de la protéine mais aussi au niveau de la couche de solvant structuré qui entoure la protéine.

L'inconvénient majeur de la cristallographie aux neutrons est que le flux de neutrons provenant de réactions nucléaires, est beaucoup plus faible que celui des photons X d'une source synchrotron. Pour palier à ce problème, l'utilisation de cristaux de volume supérieur à 1 mm³ est nécessaire. Dans le chapitre 7, seront discutées les difficultés et les différentes techniques de cristallisation utilisées pour essayer d'obtenir des cristaux deutériés de volume important, en vue d'expérience de diffraction neutronique. Ce chapitre est aussi consacré aux conclusions et aux perspectives de ce travail.

Chapitre 2

Œil, cristallin, γ cristallines et neutrons

Abstract

The optical quality and refractive power of eye lenses are dependent on the form of the refractive index gradient, which in turn is determined by the regular arrangement of the fibre cells and by the protein concentration gradient. Although, lenses are made of many different crystallin proteins, the regions of high and low protein concentration are characterised by different proportions of component proteins, with the γ -crystallins being particularly enriched in the dehydrated nucleus where the refractive index is the highest. The bulk of the lens proteins belong to the α -crystallins, which function as small heat shock proteins, and the β - and γ -crystallin superfamily which are structurally related. β - and γ -crystallin polypeptides chains are folded into two similar β -sheet domains, each domain being composed of two Greek-key motifs. The closely related family of monomeric γ -crystallins comprises around 170 amino acids in which the two domains are linked covalently by a connecting peptide.

The X-ray structure of five γ -crystallins is known. Each of them has numerous ion pairs on the surface that lead to a high surface charge density but low overall charge because the negative and positive charge are well balanced, with arginines and aspartates frequently forming interacting chains. This may give extra stability to the structure and a reduced possibility of interaction with the solvent. It would be consistent with the low water content of lens cells and especially, the nucleus; the water content of the crystal is roughly that found at the centre of the lens.

A high resolution neutron diffraction study will supplement the existing X-ray data and allow us to make detailed and thorough analysis of the solvent organisation around the crystallin molecule. This will provide a structural base for studying the effect on protein-water, water-water and protein-protein interactions.

2.1 Le système protéique : les γ cristallines

2.1.1 De l'œil à la vision

Avant de discuter de ces cristallines, il est important de les localiser au niveau de l'œil et de la vision mais il est aussi indispensable de dissenter sur cet organe unique qu'est le cristallin, ainsi que sur ses propriétés étonnantes.

2.1.1.1 L'œil

Une sphère d'un diamètre d'environ 2,5 cm, d'un poids n'excédant pas 7 grammes, enchâssée dans l'orbite, protégée par les paupières, lubrifiée par des sécrétions lacrymales et animée par trois paires de muscles. Voilà l'œil humain, organe unique avec lequel l'homme jauge son environnement, ses couleurs, ses formes, et ses mouvements. Voici donc la vision, l'un des cinq sens « traditionnels », et peut être celui le plus étudié depuis l'antiquité (voir figure 2.1).

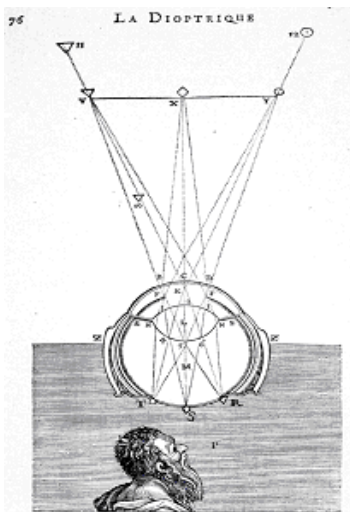
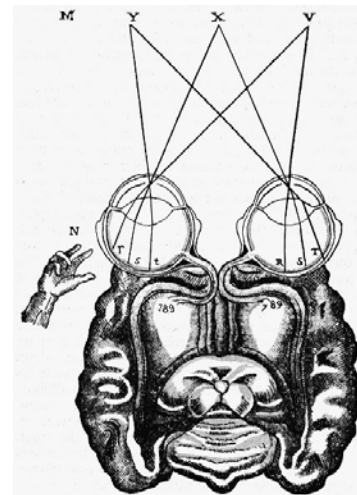


Figure 2.1: Représentation antique du système oculaire. Durant des siècles, la vision n'est qu'un jeu de rayons, une question de géométrie. A partir de Descartes, on s'interroge sur la façon dont cette image est transmise à l'âme (Descartes 1637).



Le plus étudié car peut être le plus sensible. La preuve : 70% des récepteurs sensoriels de l'organisme sont localisés dans les yeux. Plus exactement, 250 millions de photorécepteurs – cônes et bâtonnets des deux yeux – concentrés sur la rétine. En venant les frapper, les photons entraînent, *via* une cascade de réactions enzymatiques, la formation de signaux nerveux. Ceux-ci sont alors orientés vers quelques deux millions de cellules ganglionnaires, situées sur la couche

interne de la rétine. Ainsi élaboré, le signal visuel quitte l'œil par le nerf optique ; soit un million de fibres nerveuses qui l'acheminent vers les structures sous-corticales dévolues à la vision, à l'arrière du cerveau (~ 15% du cortex cérébral chez l'homme).

2.1.1.2 Le trajet optique

De nos jours, le trajet optique de la lumière à travers l'œil, est considéré de la façon suivante (voir figure 2.2 (photographie issue du site web <http://eyesite.cryst.bbk.ac.uk>)). Lorsque la lumière entre dans l'œil, elle traverse la cornée, puis un liquide appelé l'humeur aqueuse. Immersé dans l'humeur aqueuse, se trouve un muscle circulaire, l'iris, qui est percé d'un trou, la pupille. L'iris sert de diaphragme et règle l'ouverture de la pupille afin de contrôler la quantité de lumière qui entre dans l'œil (le diamètre de la pupille peut varier entre 2 mm en pleine lumière et 8 mm dans l'obscurité).

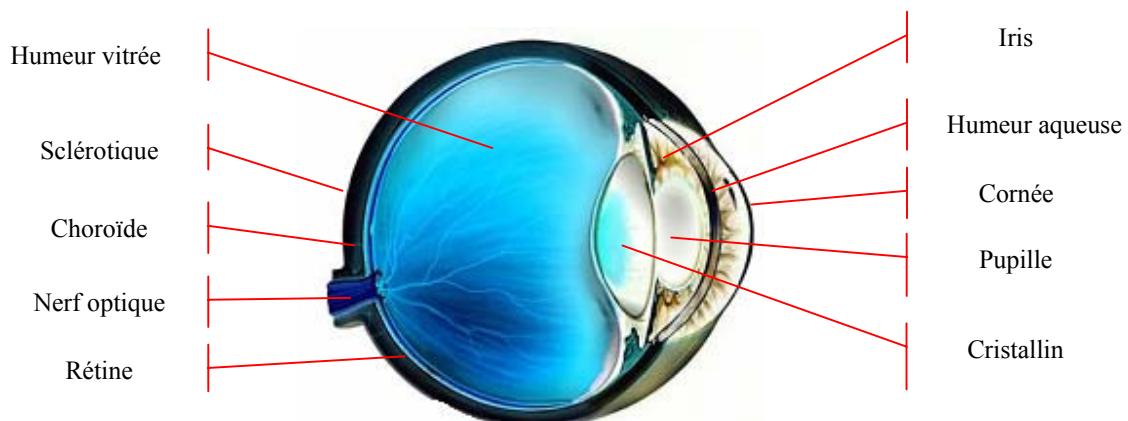


Figure 2.2 : Schéma d'une coupe transversale d'un œil humain.

Derrière l'iris se trouve **le cristallin** qui se comporte comme une lentille convergente biconvexe. Il a une structure feuilletée et comporte 22000 fines couches. Le cristallin, en se contractant, fait varier la distance focale de la lentille équivalente de 15.6 mm à 24.3 mm. Derrière le cristallin se trouve un autre liquide, l'humeur vitrée. Le fond de l'œil est constitué de trois membranes : à l'extérieur, la sclérotique, en avant la choroïde qui a pour fonction de ne pas laisser passer la lumière vers l'extérieur (elle se comporte comme une couche noire), et enfin la rétine. C'est sur la rétine que se forme l'image produite par le cristallin. La rétine est formée de cellules photo réceptrices, les cônes et les bâtonnets, qui ont des fonctions différentes : les

bâtonnets sont sensibles à très peu de lumière et ne distinguent pas les couleurs (ces cellules se comportent comme une pellicule très sensible en noir et blanc), les cônes ont besoin de beaucoup de lumière pour réagir mais distinguent les couleurs (chacune de ces cellules comprend trois pigments différents, l'un sensible préférentiellement au rouge, l'autre au vert et le troisième au bleu). Les informations enregistrées par ces cellules sont transmises au cerveau par l'intermédiaire du nerf optique. Notons qu'une portion de la rétine, celle qui est reliée au nerf optique, ne contient aucune cellule sensible à la lumière : il existe donc un point de l'œil que l'on appelle le point aveugle. Il existe également une petite partie de la rétine qui forme une dépression, ou fovéa : cette région ne contient aucun bâtonnet et renferme énormément de cônes, c'est une région importante pour la vision « en couleurs ».

Ainsi un certain nombre de structures importantes joue un rôle dans le fonctionnement de l'œil et de la vision, mais le cristallin y a une place essentielle de part sa conception et ses propriétés.

2.1.1.3 Le cristallin : un organe exceptionnel

Le cristallin est un tissu inhabituel. Comme dans différentes parties du corps humain, le cristallin présente un niveau étonnant d'interaction cellulaire, mais ce n'est pas ceci qui le différencie des autres organes. En premier lieu, le cristallin provient d'un seul type tissulaire, les cellules épithéliales de la surface ectodermique. Les fibres du cristallin sont dérivées de ce type de cellules. Il n'y a pas de remplacement (ou de mort programmée) des fibres cristallines, seulement une croissance continue de leur population. Malgré ce manque de turn-over cellulaire, le cristallin reste limpide pendant de nombreuses décades. Le cristallin ne possède pas de vascularisation sanguine, de connections nerveuses, ou tout autre contact avec les autres tissus. Les concentrations protéiques du cristallin sont très élevées (environ 35% poids/volume) alors que la pression colloïdale osmotique (capacité à garder l'eau à l'intérieur) est faible.

2.1.1.3.1 L'organisation du cristallin

Le cristallin est un organe à la fois dense, et flexible, situé entre la surface réfractante de la cornée et la rétine (voir figure 2.2). Comme la cornée, il possède deux principales propriétés optiques : la transparence et le pouvoir réfractif.

La forme du cristallin humain a pu être étudiée en partie par microscopie. Il est communément décrit comme une lentille bi convexe avec une surface antérieure et une postérieure (voir figure 2.3) en forme de parabole. Il possède un diamètre équatorial d'environ 10 mm et une épaisseur sagittale (le long de l'axe visuel) entre 3.5 et 5 mm dépendant de l'âge et du moment dans lequel se trouve l'œil. Cette lentille grossit en épaisseur sagittale tout au long de la vie et sur les deux surfaces par étapes après la jeune adolescence.

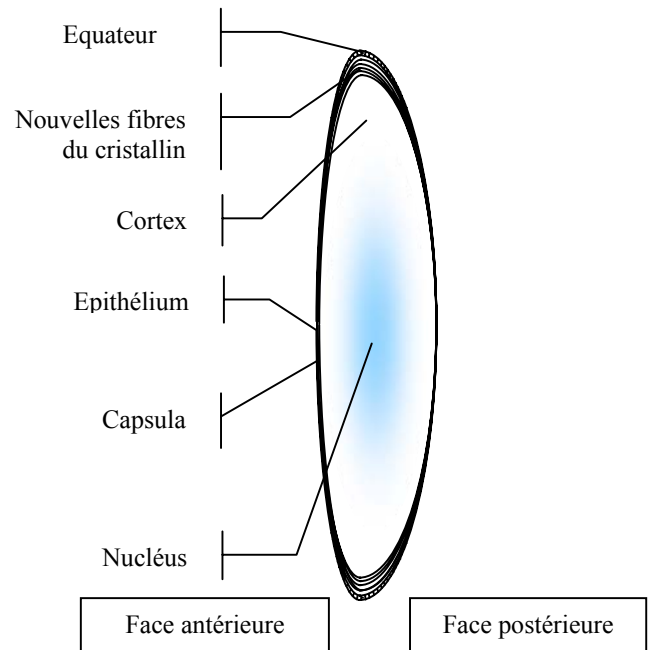


Figure 2.3 : Schéma d'une coupe transversale d'un cristallin humain.

Le cristallin est suspendu à partir du corps ciliaire qui l'encercle, par des fibres de la zonula attachées de part et d'autres à l'équateur (voir figure 2.4). Quand le muscle à l'intérieur de ce corps ciliaire se contracte, ce dernier grossit radialement et réduit l'anneau ciliaire qui, *via* les zonules, permet l'élasticité du cristallin pour prendre une forme préférentielle suivant l'état désiré. La tension est renvoyée au système de zonules par relaxation du muscle ciliaire, permettant à la lentille de retrouver sa forme aplatie. Par ce mécanisme le cristallin a augmenté son indice de réfraction par rapport aux milieux aqueux et vitreux qui l'entourent. Il s'agit du phénomène d'accommodation.

La capsule de la lentille englobe tous les éléments du cristallin. Cette capsule est une membrane épaisse, élastique produite par les cellules épithéliales de la lentille. Le gros du cristallin est constitué de fibres cellulaires. En fait, il existe un vaste réseau de cellules épithéliales qui ont perdu leur noyau ainsi que beaucoup de leurs organelles. A partir du moment où celles-ci ne grandissent plus et qu'elles ont atteint une taille idéale, elles sont appelées fibres cellulaires. Entre la capsule (située complètement à l'extérieur) et la dernière couche de fibres cellulaires se situent les cellules épithéliales. Ces « véritables » cellules continuent à se diviser et

donc à créer de nouvelles fibres. Ainsi les plus anciennes fibres se situent au centre du cristallin et forment le « nucléus ».

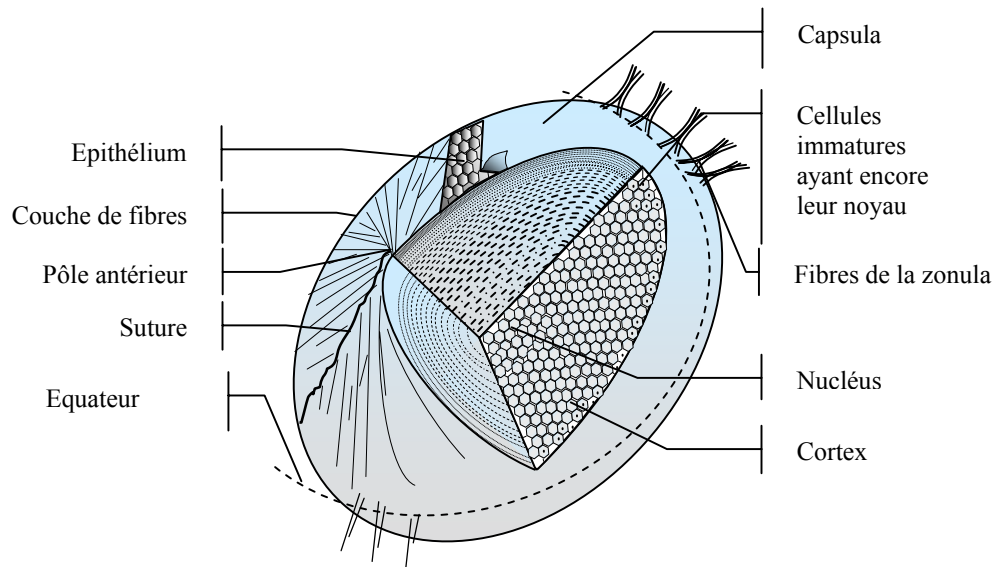


Figure 2.4 : Représentation tri-dimensionnelle d'une coupe d'un cristallin révélant les fibres disposées en couche.

Ces fibres forment donc une multitude de couches comme celles d'un oignon autour des premières fibres. Les fibres les plus récentes se retrouvent en périphérie du cortex et sont en continue formées par l'épithélium permettant au cristallin de croître en taille et en poids tout au long de la vie. Ces fibres oculaires semblent être extrêmement connectées par un réseau de jonctions et de modifications membranaires avec un petit espace intracellulaire.

2.1.1.3.2 Les fibres du cristallin

Au niveau de l'épithélium, les cellules se divisent, repoussant les cellules « filles » vers la position équatoriale. Ces dernières vont alors s'allonger et se différencier en fibres cellulaires du cristallin (voir figure 2.5).

Le côté apical des cellules (éloigné de la capsule) va pousser de l'avant, sous l'épithélium et en dessous des futures fibres, pour éventuellement terminer la suture antérieure. La partie basale (proche de la capsule) va, au contraire, reculer jusqu'à ce que les cellules rencontrent les sutures postérieures. Etant donné que les cellules s'allongent, les organelles

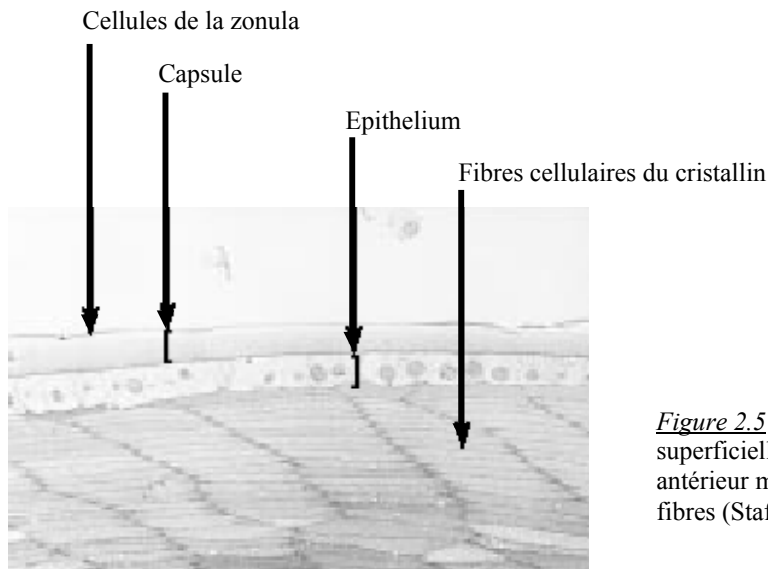


Figure 2.5 : Préparation optique des couches superficielles du cristallin, proche de l'équateur antérieur montrant la capsule, l'épithélium et les fibres (Stafford 2001).

cytoplasmiques, à l'exception des mitochondries, vont réduire en nombre jusqu'à disparaître. Le noyau cellulaire va, lui aussi, s'effacer quasiment en même temps que la partie basale de la cellule est terminée.

2.1.1.3.3 Réfraction et transparence

Le cristallin peut être considéré comme une lentille convergente, due à son pouvoir de réfraction élevé. Ce phénomène résulte d'une différence d'indice de réfraction entre deux milieux, ici l'eau et le cristallin. Lorsqu'un rayonnement lumineux arrive à la surface d'une lentille, il est dévié à cause d'une propagation moins rapide dans ce milieu. Ainsi le cristallin de l'œil est une lentille biconvexe et asymétrique (voir figure 2.3). Avec ses deux rayons de courbures différents et grands devant l'épaisseur du cristallin, on parle de lentille mince. Le gradient de l'indice de réfraction entre le cortex (indice de réfraction = 1.38) et le nucléus (indice de réfraction = 1.56) est essentiellement dû à l'expression différentielle de protéines dans ces zones. En effet, dans le cortex, sont produites majoritairement des α cristallines et, dans le nucléus, sont exprimées massivement les protéines de la super famille des β - γ cristallines (voir tableau 2.1).

<i>Localisation</i>	<i>Indice de réfraction</i>	<i>Protéines exprimées</i>	<i>Contenu en eau</i>
<i>Cortex</i>	1.38	α cristallines	Elevé
<i>Nucléus</i>	1.56	β γ cristallines	Faible

Tableau 2.1 : Répartition de l'indice de réfraction suivant les zones du cristallin. L'augmentation de cet indice du cortex vers le nucléus est associée à la concentration de γ cristallines

La transparence du cristallin n'est pas encore clairement comprise mais elle est due à son homogénéité optique. Comme les autres parties transparentes de l'œil, le cristallin n'est pas vascularisé. Seule, l'humeur aqueuse apporte aux fibres l'oxygène et les éléments nutritifs dont il a besoin. Ainsi le cristallin est essentiellement cellulaire. Or des cellules dites « normales » ne sont pas transparentes à la lumière à cause de l'absorption et de la diffusion de la plupart des constituants d'une cellule normale incluant les parois cellulaires, le cytosquelette, et les organelles. Il a été précisé auparavant que le cristallin perd la majorité de ses organelles au cours de son développement. Les cellules qui ne sont pas considérées comme matures sont confinées dans les régions équatoriales, cachées par l'iris.

En fait, la transparence du cristallin est essentiellement due à un empilement en couche des fibres, où l'espace extracellulaire est minimisé. Chaque couche est disposée de façon perpendiculaire à l'entrée de la lumière dans l'œil (voir figure 2.6 (photographies du site web <http://bca.cryst.bbk.ac.uk/BCA/CNews/1998/Dec98/eyeln.html>)).

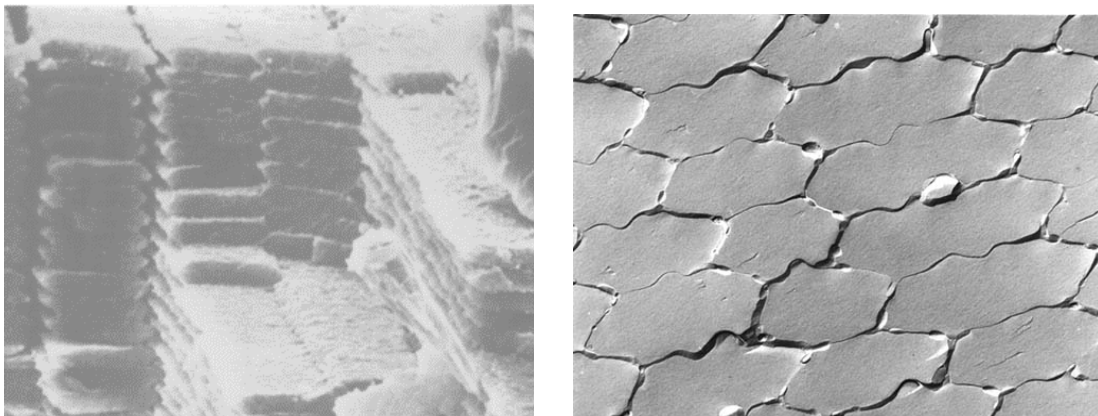


Figure 2.6: Images de microscopie électronique d'une coupe transversale d'un cristallin de bœuf montrant l'empilement hexagonal et régulier des fibres du cristallin

Ainsi la diffusion de la lumière par les membranes plasmiques, bien qu'elle soit encore présente, est diminuée. Mais il existe une légère diffusion de la lumière provenant du cytoplasme et des protéines qu'il contient. Il a été suggéré (Trokel 1962) que la transparence était la conséquence d'une distribution ordonnée (para-cristalline) de protéines à travers le cytoplasme ne créant ainsi pas de grosses fluctuations dans l'indice de réfraction au niveau de la cellule. Cette idée est depuis longtemps évoquée étant donné que l'origine grecque du mot cristallin signifie « *un empilement comme de la glace* ». Il a été aussi pensé (Benedek 1971) qu'une organisation dense des protéines à très haute concentration (identique à celle présente dans le cristallin), mais non cristalline est suffisante pour réduire l'intensité de la lumière diffusée. Ainsi la transparence est due essentiellement à l'homogénéité optique des fibres cellulaires qui dépend d'un arrangement bien ordonné à très haute concentration de petites protéines solubles : **les cristallines**.

Donc, plus la proportion de cristallines solubles est considérable par rapport au solvant (l'eau) dans les fibres du cytoplasme, plus l'indice de réfraction est élevé. Et le pourcentage de protéines solubles est important pour la transparence ; mais trop ou pas assez d'eau et le cytoplasme deviendrait trouble.

2.1.2 Les cristallines

Le cytoplasme cristallin est constitué d'une solution protéique extrêmement concentrée de protéines globulaires caractéristiques appelées **les cristallines**. Ces cristallines représentent 90 à 95 % des protéines solubles dont le poids diffère de 20 à plus de 800 kDa avec un nombre de sous unités variant de 1 à 40 environ.

2.1.2.1 Les cristallines : des protéines

Le terme « *protéine* » vient d'un mot grec signifiant « *de première importance* » ou « *qui tient la première place* ». En effet, les protéines sont des constituants extrêmement importants des cellules vivantes, tant d'un point de vue quantitatif (elles représentent en général plus de la moitié du poids sec des cellules), que d'un point de vue qualitatif, car certaines possèdent des implications biologiques capitales (en particulier les enzymes) ou encore des rôles beaucoup plus

structuraux. Toutes les protéines contiennent en majorité les quatre éléments : C, H, O et N ; beaucoup contiennent du soufre, certaines du phosphore ou encore des métaux... Les protéines, aussi appelées protides, sont de grandes molécules, des *macromolécules*, formées par la condensation de plus de 50 acides aminés, *via* des liaisons peptidiques. Les propriétés des acides aminés (charge, hydrophobicité...) gouvernent la structure de la protéine, globulaire ou fibrillaire, qui peut être décrite à différents niveaux :

- la structure primaire, qui est la séquence linéaire des acides aminés dans la protéine ;
- la structure secondaire, qui rend compte de l'organisation de groupes d'acides aminés en éléments structuraux simples : hélices α , feuillettes et brins β et boucles;
- la structure tertiaire, qui correspond au repliement de la protéine dans l'espace tridimensionnel. Cette structure rend compte de l'organisation entre eux des éléments de structure secondaire ;
- la structure quaternaire, qui définit l'association (multimérisation) entre des protéines de structures primaires identiques (homoassociation) ou distinctes (hétéroassociation).

La variété des protéines provient du nombre important de combinaisons possibles entre les 20 acides aminés principaux. Par exemple, avec une chaîne peptidique de 50 acides aminés, il est possible de construire 20^{50} protéines différentes par leur séquence. Mais, toutes ces combinaisons ne correspondent pas forcément à un polypeptide fonctionnel. En effet, le plus souvent, seront viables seules les séquences permettant un repliement tridimensionnel (le plus généralement relatif à une fonction).

Le repliement et la stabilité des protéines sont déterminés par des interactions intra et intermoléculaires (liaison disulfure, liaison ionique, liaison hydrogène, liaison hydrophobe, force de van der Waals), par diverses interactions avec l'environnement (solvant et autres protéines).

Dans ce travail, les macromolécules étudiées, les γ cristallines sont de petites protéines monomériques de forme globulaire. Ces protéines font partie d'une famille ayant des structures 3D identiques mais avec de légères variations de séquences essentiellement localisées en surface, permettant des interactions différentes avec leur environnement.

2.1.2.2 Les cristallines : les protéines essentielles du cristallin

Les cristallines sont synthétisées dans les fibres cellulaires du cristallin, en même temps que la perte du noyau ; il n'y a donc aucune chance d'avoir un renouvellement protéique. Les protéines, qui apparaissent pendant la pré-embryogénèse, doivent donc persister pendant toute la vie de l'organisme. Les cristallines sont donc pressenties comme des protéines extrêmement stables. De plus il existe une absence de système de dégradation des protéines au niveau du cristallin ; ceci assure la survie de toutes les protéines au sein des fibres oculaires. Il doit être aussi considéré que la stabilité des molécules dans l'œil peut être due à la présence d'une structure extrêmement ordonnée. Mais il est aussi possible que le facteur de viscosité, produit par de haute concentration protéique, peut contribuer à cette fabuleuse stabilité des protéines du cristallin (Jacob et al. 1997).

A travers toutes les cristallines, il est possible de discerner trois groupes majoritaires de protéines : les α -cristallines, les β -cristallines et les γ -cristallines. Les α -cristallines (αA et αB) ont environ 55 % de similarité de séquence et de réactivité immune. Les β -cristallines sont composées de polypeptides acides βA (65% similaire) et basiques βB (56 % similaire). Les γ cristallines sont, quant à elles, homologues à 70-80% suivant les organismes. Les β et γ cristallines ont une origine commune dans l'évolution (D'Alessio 2002), conduisant à des similarités de séquences et de structures et d'immunoréactivité. Ainsi, d'après ces conservations de séquences et de structures, seulement deux grands groupes de protéines se détachent : la famille des α -cristallines et la super famille des β/γ cristallines. Les α et les β cristallines sont des oligomères (association de chaînes polypeptidiques) polydisperses, alors que les γ cristallines forment des monomères compactes de 21 kDa environ.

2.1.3 Les α cristallines

Les α cristallines représentent jusqu'à 50% des protéines solubles du cristallin. Celles-ci sont considérées comme des protéines structurales du cristallin des vertébrés jouant un rôle important dans le maintien de la transparence du cristallin. A cause de leur polydispersité, elles n'ont pas été encore cristallisées mais quelques structures de sHSP ont déjà été résolues (Kim et al. 1998; van Montfort et al. 2001).

2.1.3.1 Les α cristallines : « small Heat Shock Proteins » et autres fonctions

Il existe deux classes d' α cristallines, les α A cristallines et α B cristallines. Les α A sont composées de 173 acides aminés et les α B de 175 (sauf chez les rongeurs où les α A font 196 acides aminés). Les α A sont légèrement plus acides que les α B mais leur point isoélectrique se situe aux alentours de 4,75. La similarité de séquence des α cristallines avec des protéines du choc thermique de la drosophile a conduit à l'établissement d'une nouvelle famille, les petites protéines du choc thermique (Ingolia et Craig 1982). Le domaine relatif aux HSP, environ 100 acides aminés, se situe sur la partie C-terminale. La conservation de ce domaine au cours de l'évolution démontre l'importance fonctionnelle de ce domaine (de Jong et al. 1998). Il a été démontré que les α cristallines pouvaient aussi avoir des fonctions d'auto-kinase (Kantorow et Piatigorsky 1994) et de protéines chaperonnes vis à vis des β et γ cristallines contre l'agrégation (Klemenz et al. 1991; Horwitz 1992).

2.1.3.2 Les α cristallines : des oligomères hétérogènes de grande taille

Les α cristallines natives sont présentes dans l'œil, sous forme d'un large complexe de 3 sous unités α A et d'une α B formant des poids moléculaires de 300 à 1000 kDa. De 15 jusqu'à 50 sous unités peuvent former ces complexes. Les α cristallines subissent aussi de nombreuses modifications post-traductionnelles telles que la phosphorylation, la racémisation aspartique, la déamidation ou encore des coupures.

2.1.3.3 Les α cristallines : cataractes et maladies

De nombreuses cataractes sont dues à des modifications des α cristallines. La plupart sont le résultat d'irradiation UV, de mécanisme d'oxydation et de stress osmotique. Généralement la non expression ou des mutations des α cristallines conduisent à des pertes de fonctions protectrices des α cristallines et donc, dans certains cas, à la formation de cataractes. Une accumulation de α B cristallines a été observée dans de nombreux cas de maladies neurologiques dégénératives (la maladie de Creutzfeld-Jacob, la maladie d'Alexander, *etc*).

2.1.4 La super famille des β/γ cristallines

Les protéines β - et γ - cristallines ont été reconnues comme étant des membres de la super famille des β/γ cristallines par différents alignements de leur séquence primaire mais aussi, plus récemment, par des structures cristallographiques relevant le même type de repliement (Lubsen et al. 1988; Wistow et Piatigorsky 1988). En effet, ces protéines se révèlent avoir comme motif structural de base : la clef grecque.

2.1.4.1 La clef grecque : dénominateur commun des β/γ cristallines

Les β/γ cristallines possèdent la même topologie du motif de la clef grecque dont le nom est basé sur la similarité de cette super structure secondaire avec le modèle trouvé sur les poteries grecques (voir figure 2.7).

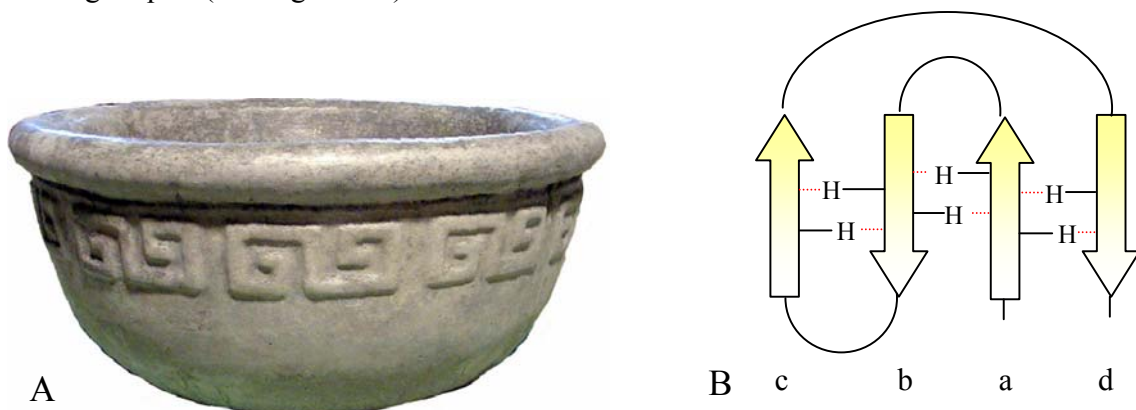


Figure 2.7 : A) Poterie grecque ayant le motif de la clef grecque. B) super structure secondaire nommée la clef grecque où 4 brins β antiparallèles sont stabilisés par des liaisons hydrogène.

Ce motif est composé de 4 éléments de structures secondaires formant un feuillet β (voir figure 2.7). Ce feuillet s'organise en 4 brins antiparallèles, liés par de courtes connections entre les brins a et b, b et c respectivement et une plus grande boucle entre les brins c et d, forçant le brin d à faire des liaisons hydrogène avec le brin a. Des similarités de structure avec des protéines de réponse à des stress (Superoxide Dismutase humaine, la protéine S de *Myxococcus*

xanthus produite pendant la sporulation) conduisent à une comparaison directe des fonctions de ces protéines ; d'un côté, la transparence du cristallin, et de l'autre, la réponse aux stress. Il est très difficile de trouver des points communs à ces deux phénotypes, mis à part le dénominateur commun de la « stabilité ». Ainsi, il est envisageable de dire que le motif de la clef grecque procure une haute stabilité à la protéine. La fonction de la clef grecque n'est pas encore très bien identifiée, cependant, de récentes études suggèrent que ce motif peut être responsable d'interactions protéine-protéine particulières que l'on retrouve chez les immunoglobulines (Crabbe et Goode 1995). Comme les immunoglobulines, les β/γ cristallines sont repliées en une structure β complète. Sans aucun doute, une accumulation des liaisons hydrogène dans cette structure ainsi que des interactions hydrophobes entre les domaines contribuent largement à la stabilité. Cette stabilité générale peut être comprise, en fait, comme une somme de contribution des repliements de petits domaines. Une corrélation spatiale des molécules à une haute concentration donnée dans le cristallin, pourrait ainsi conduire à un arrangement précis des cristallines entre elles. Cet ordre mènerait à une diminution de la diffusion de la lumière, et donc à un optimum de transparence du cristallin.

2.1.4.2 Les β cristallines

La famille des β cristallines peut être divisée en 2 sous-groupes : un acide : les βA cristallines, et l'autre basique : les βB cristallines. Chaque sous-groupe est codé par 3 gènes (*Cryba1*, -2, -3 ; *Crybb1*, -2,-3). Cependant *Cryba1* code pour deux protéines ($\beta A1$ et $\beta A3$ cristallines). Les protéines natives β cristallines sont présentes sous différentes formes d'oligomères (le poids moléculaire d'un monomère étant situé entre 22 et 28 kDa) avec un point isoélectrique allant de 5,7 à 7,0. Les dimères, les trimères et peut être les tétramères de β cristallines forment une sub-fraction en chromatographie d'exclusion, appelée la βL –cristalline (pour les faibles poids moléculaires : 40-100 kDa) ; les formes les plus grandes (βH –cristalline, jusqu'à 200kDa) sont composées essentiellement d'octamères. Les β cristallines sont acétylées sur leur partie N-terminale ; et elles subissent d'autres modifications post-traductionnelles telles que la glycosylation et la phosphorylation. Seules des mutations ponctuelles (entraînant la formation de codon STOP), des délétions ou modifications post-traductionnelles (formation de ponts disulfure) des β cristallines sont connues pour former des cataractes.

Contrairement aux α cristallines, la structure tridimensionnelle des β cristallines a été résolue (Bax et al. 1990; Wistow et al. 1981), et a confirmé l'association des β B2- cristallines en homodimère (voir figure 2.8). Cette caractéristique peut être attribuée au peptide connecteur (ou de liaison) entre deux domaines, chacun formé par des motifs en clef grecque.



Figure 2.8 : Structure de la β B2 cristalline. Le domaine N- terminal de la chaîne A (en bleu) interagit avec le domaine C terminal de la chaîne B (en orange) pour former un dimère. Chaque domaine est composé d'un motif en clef grecque.

De plus, la flexibilité de ces domaines autour de ce peptide permet la formation de dimères ou de tétramères qui ne peuvent s'inter-convertir que par dénaturation. En effet, les β cristallines peuvent former des dimères stables par contacts inter-domaines intermoléculaires. Pour les β B2- cristallines, le domaine N-terminal d'un monomère interagit avec le domaine C-terminal de l'autre domaine. Les γ cristallines possèdent aussi un peptide connecteur dont la conformation, beaucoup moins allongée, est différente du peptide des β cristallines. Des expériences de biologie moléculaire ont été entreprises afin de tester l'importance de ce peptide

de liaison. Le remplacement de ce dernier chez les γ cristallines par celui des β cristallines ne permet pas la formation de dimères, ce qui démontre que les interactions inter-domaines favorisent la dimérisation (Mayr et al. 1994). En revanche, le remplacement du peptide des β cristallines par celui des γ cristallines prouve que la dimérisation est le résultat de la linéarité du peptide de liaison (Trinkl et al. 1994).

2.1.4.3 La protéine γ S cristalline

La γ S cristalline est peut être l'un des ancêtres intermédiaires entre les β - et les γ -cristallines. La comparaison des séquences montre une plus proche relation avec les γ cristallines (53% d'identité) qu'avec les β cristallines (35%). γ S possède 4 résidus supplémentaires en N-terminal (comme les β cristallines) et en a perdu 2 en sa partie C-terminal. De plus, la γ S cristalline de bœuf est monomérique et n'a pas tendance à former des assemblages jusqu'à des concentrations de l'ordre du millimolaire. Le peptide connecteur des deux domaines a un acide aminé supplémentaire comparé à la β B2- cristalline et à la γ B cristalline ; cependant l'analyse de la surface inter domaine montre que les résidus chargés sont appariés expliquant le comportement monomérique des γ S cristallines. Une importante fonction des γ S cristallines au niveau du cristallin est de prévenir la séparation de phase du cytoplasme. Liu et ses collaborateurs (Liu et al. 1996) ont démontré que les γ S cristallines jouent un rôle prédominant dans la prévention de la formation de certaines cataractes. La présence de protéines γ S dans des solutions concentrées diminue les températures de transition de phase. Les γ S cristallines diminuent l'agrégation des solutions de γ cristallines, et donc aident au maintien de la transparence du cristallin.

2.1.4.4 Les γ cristallines

Sur des fractions totales du cristallin, les γ cristallines apparaissent en dernier, en chromatographie d'exclusion, à cause de leur faible masse de 20 kDa et de leur structure monomérique. Elles sont considérées comme des protéines basiques avec un point isoélectrique allant de 7,1 à 8,6. Ces dernières possèdent de nombreuses cystéines (4 à 7 par molécule) permettant la formation de ponts disulfure comme c'est le cas pour la γ B cristalline de veau

entre les cystéines 18 et 22. Les protéines qui ont été étudiées pendant cette thèse, sont la γ E et γ C cristallines de rat (*Rattus norvegicus*), γ D cristalline humaine (*Homo sapiens*) et la γ B de bœuf (*Bos taurus*). Dans ce rapport, je me suis concentré essentiellement à la γ E cristalline de rat.

2.1.4.4.1 Gènes de γ cristallines

La famille des γ cristallines est étroitement liée sur le chromosome 2q chez l'homme et sur le chromosome 9q chez le rat. Il y a 6 gènes dans les deux génomes récemment séquencés. En revanche chez l'homme, deux gènes γ cristallines (γ E et γ F) sont des pseudogènes. Comme il a déjà été spécifié, les γ cristallines sont de petites protéines (173-174 acides aminés) de deux domaines. Ces deux domaines sont codés par de larges exons : l'exon 2 étant de 243 paires de bases et l'exon 3 de 273 à 274 paires de bases. Les deux exons sont séparés par une séquence non codante de taille variable (0,9 et 3,8 kb). Les gènes de ces protéines comportent un exon de très petite taille (40-50 paires de bases) codant pour une région non transcrite et les 3 premiers acides aminés. La première séquence non codante est très courte. Tous les gènes codant pour les γ cristallines possèdent une région promotrice conservée à -50 : la séquence MARE (Maf Recognition Element, séquence reconnaissant le facteur de transcription Maf) et une séquence reconnaissant le facteur de transcription Sox 1. Il a été démontré que de nombreuses cataractes étaient induites par des mutations au niveau de ces facteurs de transcriptions et de ces régions promotrices sans pour autant en associer un phénotype précis.

2.1.4.4.2 Maladies

Les recherches génétiques sur le cristallin ont permis diverses mutations sur les gènes des γ cristallines. Par exemple, les souris mutantes *Elo* (Eye lens obsolescence), présentent une délétion d'un seul nucléotide sur le gène *Cryge* (gène codant pour la cristalline E). Cette mutation détruit le cadre de lecture du gène, et donc au niveau de la protéine, le quatrième et dernier motif de clef grecque de la γ E cristalline. L'instabilité de cette macromolécule sera ensuite observable sur l'opacification du cristallin de la souris. D'autres mutations ont été observées (voir figure 2.9). Le remplacement de l'arginine 58 par une histidine chez la γ D cristalline humaine, est identifiable dans les cataractes aculéiformes. Ces cataractes sont surprenantes par la formation de cristaux à l'intérieur même du cristallin. La mutation R58H

induit un changement de charge et donc une modification des interactions avec les résidus voisins (171-173) et les molécules d'eau environnantes (Basak et al. 2003).

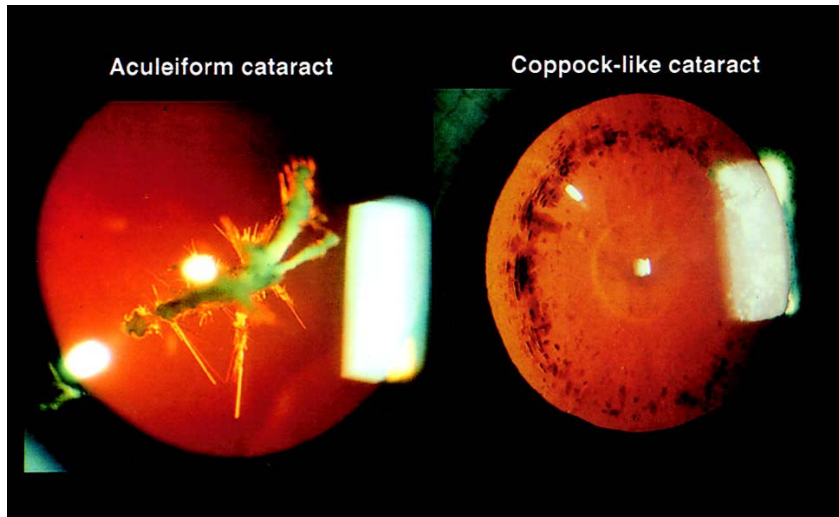


Figure 2.9: Exemples cliniques d'une cataracte aculéiforme (à gauche) et d'une cataracte de Coppock (à droite), montrés par photographie trans-illuminatoire (Heon et al. 1999).

Dans certaines cataractes de Coppock, il est possible de trouver le changement d'une adénine en cytosine dans l'exon 2 (position 225) des γ C cristallines entraînant la substitution d'une thréonine en une proline dans une région qui est normalement, hautement conservée chez les γ cristallines. Cette mutation va entraîner des changements conformationnels et une déstabilisation de la protéine (Castan et Enfors 2002) ; et ceci sera visible par la formation d'agrégats au niveau du cristallin.

2.1.4.4.3 Transition de phase des γ cristallines

In vitro, lorsqu'une solution de γ cristallines est emmenée en dessous d'une certaine température (température critique), cette solution va s'opacifier et présenter un comportement de séparation de phase liquide-liquide, que l'on peut observer dans la cataracte froide (Siezen et al. 1985; Thomson et al. 1987; Broide et al. 1991). Ce phénomène de cryo-précipitation est réversible par une augmentation de la température au dessus de la limite critique. Plusieurs études ont permis de discerner deux types de γ cristallines dont les caractéristiques physiques sont différentes : les non cryo-précipitables et les cryo-précipitables. Broide et ses collaborateurs ont

démontré que les non cryo-précipitables présentaient une température de transition de 5 °C alors que celle des cryo-précipitables était située entre 35 et 40 °C (voir figure 2.10).

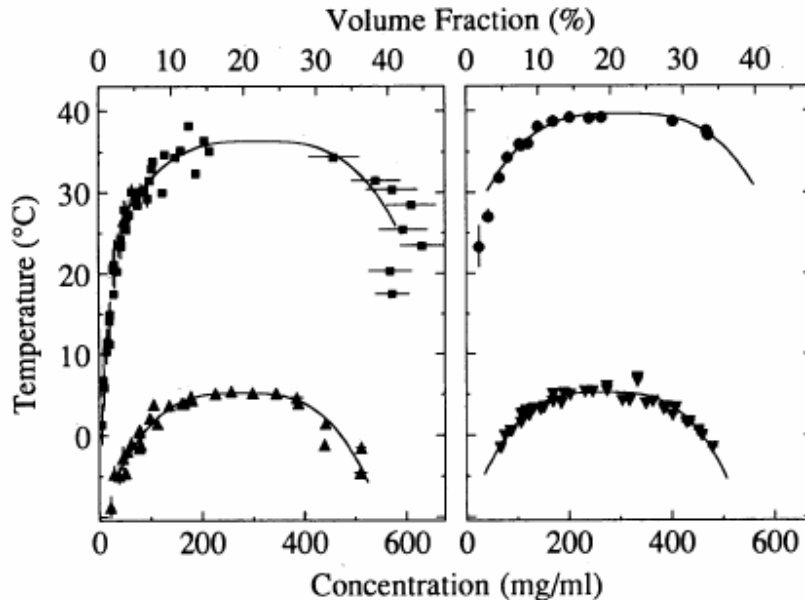


Figure 2.10 : Courbes de co-existence de solutions aqueuses de γ cristallines purifiées à partir de veau mesurées à pH7.1 dans un tampon phosphate 0.1M (force ionique 240 mM). Les cristallines cryo-précipitables : γ E (γ IIIa- ■) et γ F (γ IVa- ●) présentent une transition à partir de 35 °C ; et les cristallines non cryo-précipitables ont une température de transition à 5 °C : γ D (γ IIIb- ▲) et γ B (γ II- ▼) (Broide et al. 1991). En dessous de ces courbes, la solution de γ cristallines est transparente (région à une phase), alors qu'en dessus de cette courbe, la solution est opaque (région à deux phases) à cause de la séparation de phase.

Ce comportement de séparation de phase dépend de nombreux facteurs dont la température, le pH, la force ionique, la concentration en protéine. Il a été proposé que cette différence de température de transition serait liée aux cystéines contenues dans les protéines (Pande et al. 1993). Mais il est aussi possible que les γ F cristallines jouent un rôle dans le maintien de la transparence du cristallin en abaissant la température de transition de phase des γ A et des γ B cristallines (Liu et al. 1996).

2.1.4.4.4 Structures des γ cristallines cryo- et non cryo- précipitables

Les γ cristallines possèdent une structure primaire fortement conservée, que ce soit au sein même de l'organisme ou à travers les espèces (voir figure 2.11).

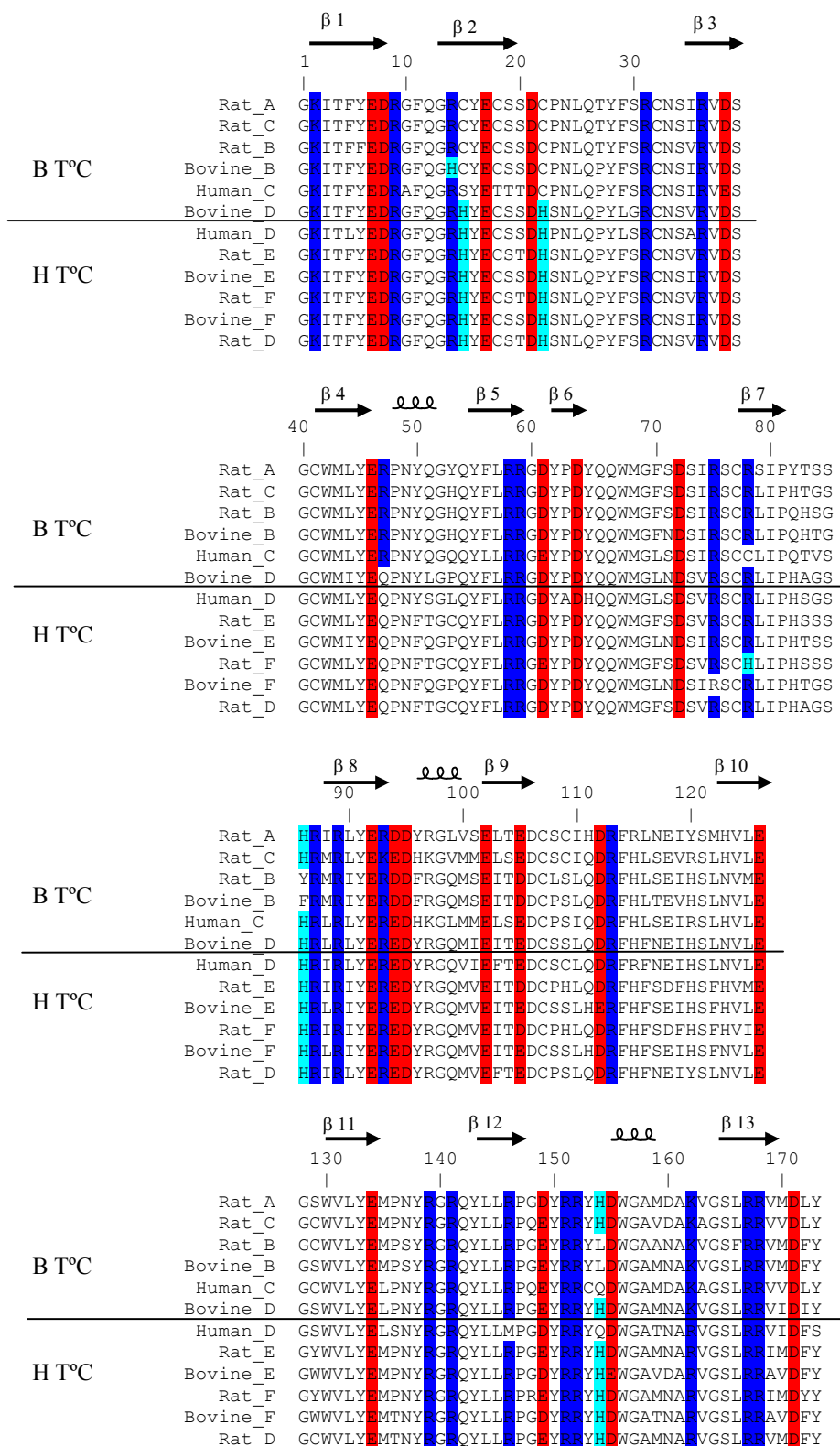


Figure 2.11 : Alignement des séquences primaires des γ cristallines dont le comportement de séparation de phase a été caractérisé. Les γ cristallines non cryo- précipitables (B T°C) sont représentées au dessus de la ligne noire et les non cryo- précipitables en dessous. Les résidus arginines et lysines qui sont convertis sont surlignés en bleu foncé et les histidines en bleu clair. Les résidus acides (aspartate et glutamate) les plus conservés sont surlignés en rouge.

En effet, les séquences en acides aminés démontrent que les γ cristallines (et les β cristallines) appartiennent à une même famille. Une analyse plus poussée des séquences permet aussi de se rendre compte que les motifs en clef grecque sont aussi répétés et conservés à l'intérieur même de la protéine. Sur la figure 2.11, les feuillets β se répètent quasiment de façon symétrique. Il ne serait donc pas étonnant que ces protéines présentent le même repliement tri-dimensionnel. A ce jour, plusieurs structures de γ cristallines de formes natives, recombinantes, ou encore des mutants ont déjà été résolues (voir tableau 2.2).

Protéine	γB	γD	γD	γE	γF	γS C-terminal
Organisme (recombinante)	<i>Bos taurus</i>	<i>Bos taurus</i>	<i>Homo sapiens</i> (<i>E. coli</i>)	<i>Rattus</i> <i>norvegicus</i>	<i>Bos taurus</i>	<i>Bos taurus</i> (<i>E. coli</i>)
ID de la PDB	1AMM	1ELP	1HK0	1A5D	1A45	1A7H
Résolution	1,20 Å	1,95 Å	1,25 Å	2,30 Å	2,30 Å	2,56 Å
Groupe d'espace	P 4 ₁ 2 ₁ 2	P 2 ₁ 2 ₁ 2 ₁	P 2 ₁ 2 ₁ 2 ₁	P 1 2 ₁ 1	C 2 2 21	P 6 ₁ 2 2
Basse ou haute T° de transition	Basse	Basse	Haute	Haute	Haute	-
Maille				$a=37,76, b=43,71, c=107,44,$		
Dim. en Å	$a=56,43, b=56,43, c=97,12.$	$a=57,81, b=70,03, c=117,25.$	$a=33,22, b=52,97, c=90,47.$	$\alpha=\gamma=90,00, \beta=107,44.$	$a=35,10, b=46,20, c=186,20.$	$a= b=62,25, c=170,00.$
Nbre de mol. par u. a.	1	2	1	2	1	2
Nbre de mol. d'H ₂ O. par mol.	392	109	282	128	250	12
Référence	(Kumaraswamy 1996)	(Chirgadze et al. 1996)	(Basak et al. 2003)	(Norledge et al. 1997)	(Norledge et al. 1997)	(Basak et al. 1998)

Tableau 2.2 : Quelques structures cristallines des γ cristallines. Sont citées les structures natives ou les mutants (γS C-terminal) dont la structure native n'est pas connue. D'autres structures mutantes sont connues telles que la partie C- terminale (1DLS) ou le mutant C18S (1I5I) de la γB bovine ou encore le mutant R58S (1H4A) de la γD cristalline humaine.

Ainsi les chaînes principales de la γ B cristalline de bœuf (Kumaraswamy 1996), de la γ D cristalline humaine (Basak et al. 2003), la γ E cristalline de rat et la γ F cristalline de bœuf (Norledge et al. 1997) se superposent (voir figure 2.12) comme pouvait le faire penser le haut degré d'identité de séquence. La combinaison de deux motifs formant un domaine est stabilisée par les brins a, b et d d'un motif et c d'un autre motif par des liaisons hydrogènes. La molécule entière peut être comparée à un "sandwich" de quatre feuillets β , les deux motifs au centre étant les motifs 2 et 4, alors que les motifs 3 et 1 exposent essentiellement les feuillets au solvant extérieur. Les boucles liant les feuillets sont très flexibles spécialement celle qui connecte les brins c et d du motif 3.

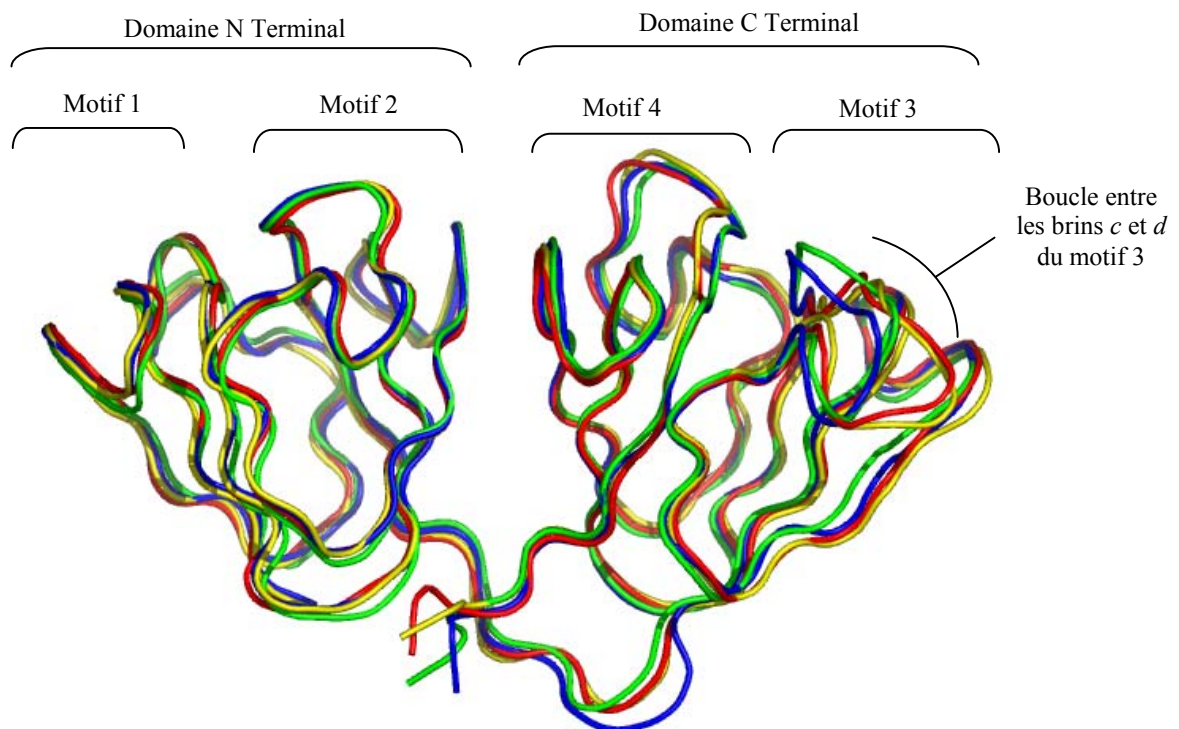


Figure 2.12 : Comparaison des structures tri- dimensionnelles des γ cristallines dont la structure est déjà connue : la γ B (bleu), la γ D (jaune), la γ E (rouge) et la γ F (vert).

Les γ cristallines présentent toutes une symétrie interne très forte. En effet, les deux domaines sont reliés par un axe de symétrie d'ordre 2. De plus, deux autres axes du même ordre connectent le motif 1 au motif 2, et le motif 3 au motif 4 respectivement.

Norledge et ses collaborateurs, en 1997, se sont intéressés aux différences existantes entre les cryo- et les non cryo-précipitables γ cristallines. La différence majeure, d'un point de vue structural, réside dans la boucle entre les brins c et d du motif 3, qui semble être due à la présence de gros résidus hydrophobes au cœur du motif chez les γ cristallines de haute température de transition (voir figure 2.13).

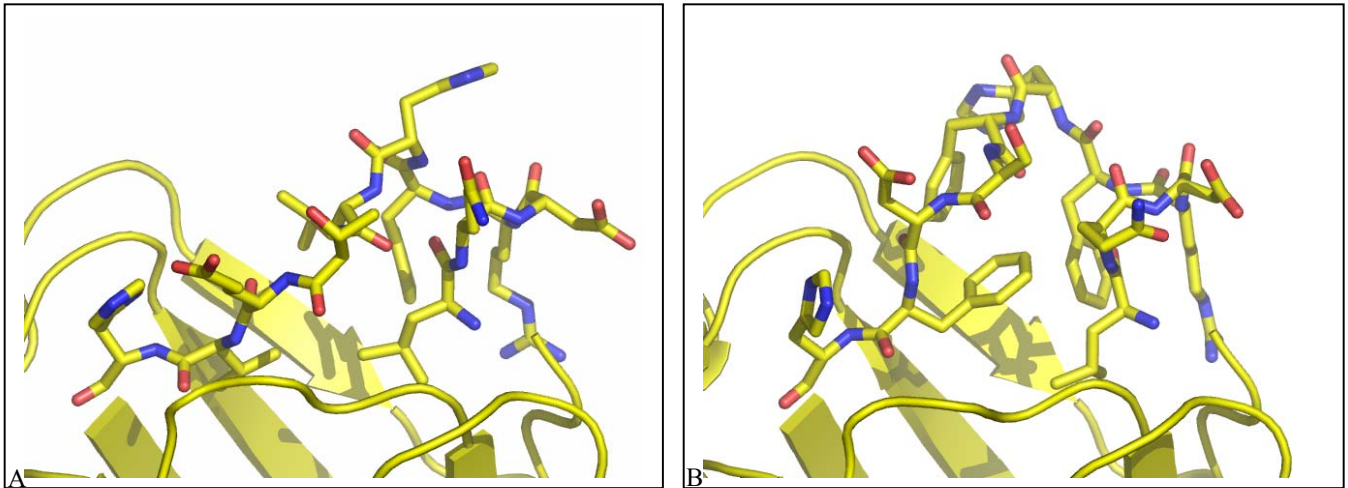


Figure 2.13 : Comparaison de la boucle entre les brins c et d du motif 3 (résidus 112 à 122). Les protéines non cryo- précipitables (A : γ B de bœuf) ne présentent pas l'encombrement stérique des résidus Phénylalanines regroupées en un cœur hydrophobe des protéines cryo- précipitables (B : γ E de rat).

Toutes les structures, qu'elles appartiennent aux γ cristallines de basse ou de haute température de transition, dévoilent une surface extrêmement « chargée » sur laquelle se trouve un grand nombre de paires d'ions, spécialement du côté C- terminal (voir figure 2.14).

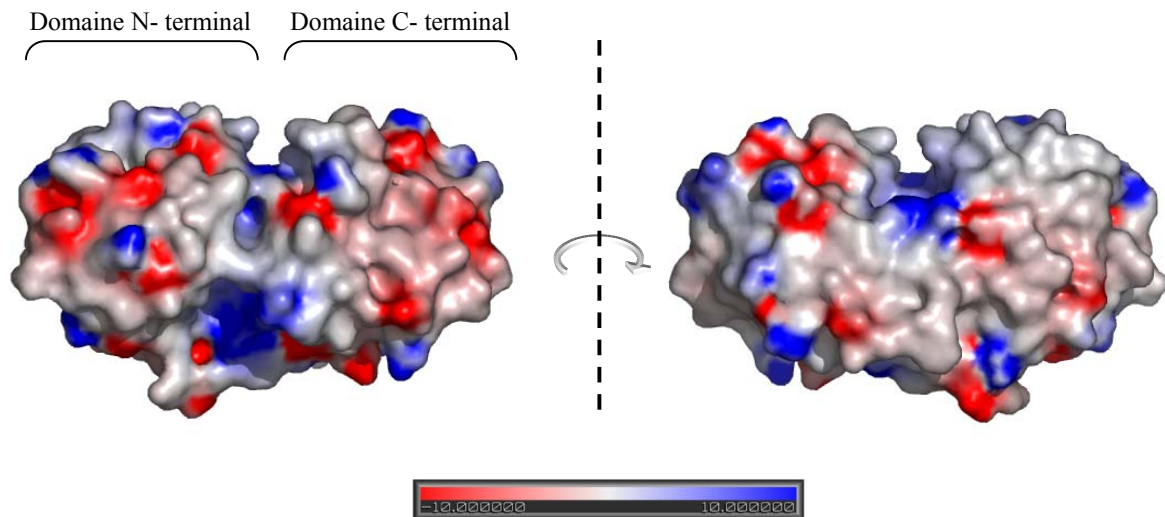


Figure 2.14 : Représentations de la surface moléculaire de la γ F cristalline de bœuf colorée par le potentiel électrostatique, générées par le programme GRASP. Ces vues montrent le grand nombre de charges présentes sur la surface.

D'un point de vue général, ils mentionnent que le résidu 163 est systématiquement conservé pour les non cryo-précipitables (lysine) par rapport aux cryo-précipitables (arginine). Or, ce résidu se situe à la surface du motif 4, pouvant ainsi permettre des interactions différentes avec d'autres protéines. Par ailleurs, il semble exister une certaine tendance à avoir un plus grand nombre d'histidines chez les γ cristallines cryo-précipitables (voir tableau 2.2). Mais, les auteurs mentionnent deux différences dans la nature des contacts cristallins (protéine-protéine). En premier lieu, il semblerait que les protéines cryo-précipitables possèdent une proportion légèrement plus importante d'interactions Arg/Arg et Arg/Lys entre les molécules. Ensuite, ces protéines présenteraient une relativement plus faible proportion à former des paires ioniques dites « conventionnelles ». D'un autre point de vue, le rapport (Arg + His) / Lys, plus important chez ces mêmes protéines, résulte en grande partie, des résidus de surface chargés et appariés, qui vont considérablement réduire la force répulsive d'hydratation qui existe entre les molécules à haute concentration. Bien que toutes les observations apportent des éléments de réponses aux différents comportements de transition de phase des γ cristallines, il est certain que ces derniers ne peuvent s'expliquer qu'en considérant l'ensemble de toutes ces observations.

Ainsi le repliement dense des γ cristallines, une forte symétrie et le regroupement de nombreuses charges en surface permettent une haute stabilité à la protéine, mais aussi la capacité d'interagir avec les molécules environnantes (solvant ou protéine).

2.1.4.5 Evolution de la super-famille $\beta\gamma$ cristalline

Les monomères de γ cristallines et les β cristallines sous forme dimérique (voire oligomérique) ont souvent été un intéressant cas d'étude dans la discussion de la transition évolutive des protéines monomériques vers les protéines oligomériques (Lubsen et al. 1988; Wistow et Piatigorsky 1988; D'Alessio 2002).

Comme il a déjà été précisé, les cristallines β et γ sont composées de deux domaines N et C-terminaux. Chaque domaine est formé par deux motifs en clé grecque. Dans les deux cas, les quatre motifs et les domaines sont organisés symétriquement autour d'axes pseudo dyades entre les domaines et les motifs. Cependant, les deux domaines des γ cristallines sont différents, d'un point de vue topologique, vis à vis des β cristallines. Les β cristallines se trouvent sous forme de

dimère du fait que le domaine N-terminus d'une chaîne interagit de façon non covalente avec la partie C-terminale d'une autre chaîne et inversement (voir figure 2.15). Le plus important est de noter que les structures des cristallines, que ce soit les γ ou les β , sont composées d'une répétition de ce motif en clé grecque qui a souvent été considéré comme le motif ancestral de la famille des β/γ cristallines.

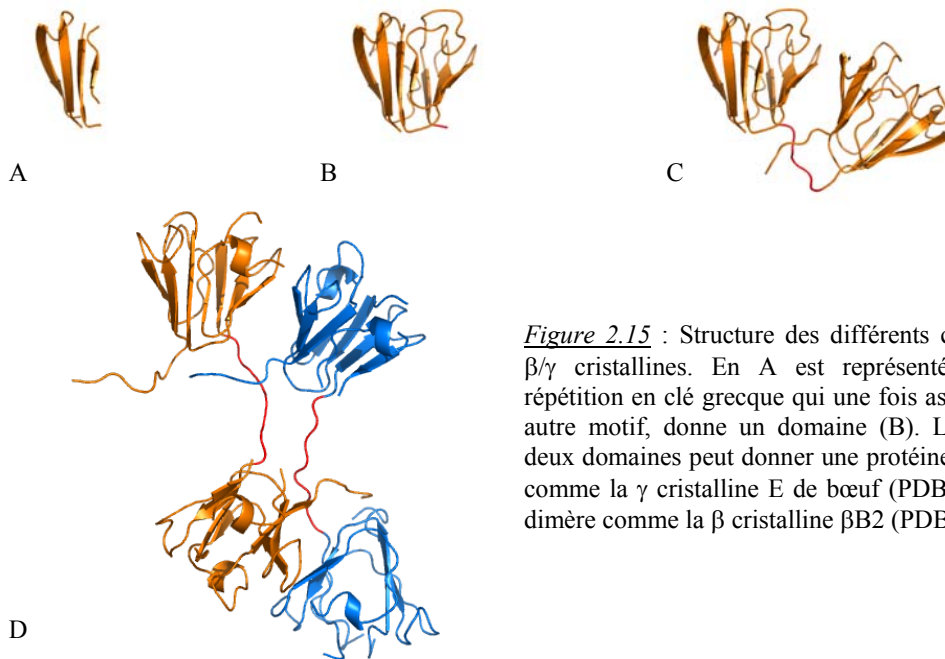
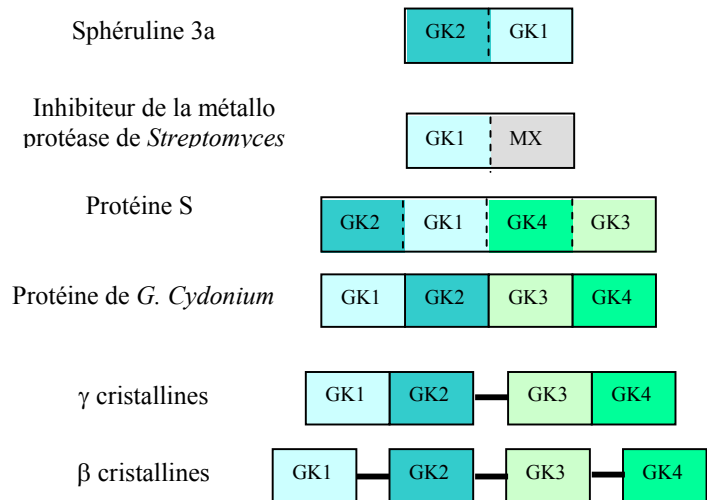


Figure 2.15 : Structure des différents constituants des β/γ cristallines. En A est représenté, le motif de répétition en clé grecque qui une fois assemblé avec un autre motif, donne un domaine (B). L'association de deux domaines peut donner une protéine monomérique, comme la γ cristalline E de bœuf (PDB 1M8U), ou un dimère comme la β cristalline β B2 (PDB 1BLB).

Au cours de l'évolution, pour conserver la précision de l'appareil optique qu'est le cristallin, les β et γ cristallines ont été le lieu d'une très forte pression de sélection. Ceci est observable par le faible taux de substitution au niveau des gènes des cristallines des vertébrés, spécialement pour les β cristallines, et par un taux inhabituel de substitutions similaires (Lubsen et al. 1988) excepté pour le peptide connecteur inter domaine et pour les extensions N- et C-terminales.

Au niveau des gènes des β cristallines, chaque motif en clé grecque (GK1, GK2, GK3 et GK4) est codé par des exons séparés (voir figure 2.16). Toutefois les gènes des γ cristallines sont composés de deux paires de motifs adjacents (GK1/GK2 et GK3/GK4), chacune étant codée par un exon et correspondant à un domaine. En fait, il existe trois exons dans les gènes des γ cristallines mais le premier code pour 3 acides aminés. Dans les deux cas, le peptide connecteur est interrompu par un intron.

Figure 2.16 : Arrangement des séquences codantes pour les motifs en clé grecque dans les gènes des protéines assimilées à la famille de β/γ cristallines. Les motifs sont montrés dans des boîtes et la numérotation est relative à la similarité de séquence qu'il existe avec les gènes des β/γ cristallines. Un domaine est formé par deux motifs adjacents. Les barres noires représentent les introns inter motifs et inter domaines. Les lignes en pointillés indiquent que la présence d'intron n'est pas reconnue (d'après d'Allessio 2002).



Un domaine similaire à celui des cristallines a été trouvé dans une protéine d'une moisissure de vase avec une haute identité de séquence et une haute similarité structurale : la sphéruline 3a (Clout et al. 2001). Or, le domaine GK1 par analogie, se trouve en C-terminal et non en N-terminal. Une autre protéine comportant un motif similaire à celui des cristallines a aussi été remarquée dans une protéine inhibitrice d'une métalloprotéase de *Streptomyces* (SMPI) (Ohno et al. 1998), avec une similarité de séquence au niveau du premier motif N-terminal, mais aucune ressemblance dans le second motif (MX).

Des repliements de deux domaines, similaires à celui des β/γ cristallines ont été aussi découverts dans la protéine S de la spore d'une bactérie (Wistow et al. 1985) où l'arrangement des motifs en clé grecque est différent des β/γ cristallines indiquant une divergence possible dans l'évolution des protéines entre les bactéries et les vertébrés à partir du motif unique (D'Allessio 2002). A l'opposé, une protéine spécifique de la différenciation de l'épiderme (EDSP) d'un batracien, présente un arrangement similaire à celui des β/γ cristallines (Wistow et al. 1995) ; tout comme une séquence protéique d'éponge marine *Geodia cydonium* mais sans une présence d'intron (Krasko et al. 1997). Enfin, la protéine AIM1, dont l'expression est en relation avec la suppression de mélanome humain, possède en son extrémité C-terminale, dans le modèle proposé, 12 motifs en clé grecque disposés en six fois deux motifs ordonnés similairement aux β/γ protéines (Ray et al. 1997).

Lubsen et al. (1988) ont exclu la possibilité que le gène de type γ , soit l'ancêtre immédiat des β cristallines en se basant sur la présence d'introns entre les domaines GK1 et GK2 du premier domaine et les motifs GK3 et GK4 du second domaine, chez les γ cristallines. Ils expliquent qu'il est impossible d'avoir, deux fois, une mutation identique, la perte d'intron, dans l'évolution de deux gènes homologues codant pour un domaine après leur divergence et avant leur fusion. Mais D'Alessio (2002) explique que cette thèse peut être sa propre antithèse : la probabilité d'avoir ce type de mutation est dépendante de facteurs extrinsèques (nature de l'agent mutagène, conditions environnementales...) et intrinsèques (séquences en acides nucléiques, conséquences de la mutation sur les structures secondaires et super secondaires...). Donc la probabilité qu'une altération génétique, de type insertion ou délétion d'intron, peut intervenir deux fois dans l'évolution de gènes homologues dans une certaine famille de gènes, est plus importante quand les gènes sont proches dans l'évolution.

En conclusion, même s'il existe certaines divergences de point de vue, il est certain que l'évolution de cette super famille protéique des β/γ cristallines, est le résultat de duplication du motif ancestral en clé grecque, de délétions d'introns et de fusion d'exons conduisant, dans un premier temps, à la formation d'un domaine à deux motifs puis, dans un second temps, à deux domaines sous forme monomérique comme pour les γ cristallines et sous forme de dimère, comme pour les β cristallines comme le propose D'Alessio (2002) en figure 2.17.

En ce qui concerne le fait que les cristallines de type γ vont tendre à former des monomères et non des dimères comme les β cristallines, il a été à de nombreuses fois discuté le rôle majeur du peptide connecteur entre les deux domaines (Mayr et al. 1994; Trinkl et al. 1994). Ces expériences de mutagenèse dirigée ont démontré que les séquences, ainsi que la longueur de peptide, serait la cause de ce comportement d'oligomérisation. Par exemple, la proline 80 forcerait le peptide connecteur en une conformation ouverte qui est retrouvée chez les β cristallines. La séquence des extensions N- et C-terminales chez les β cristallines, joue apparemment un rôle influençant la conformation du peptide.

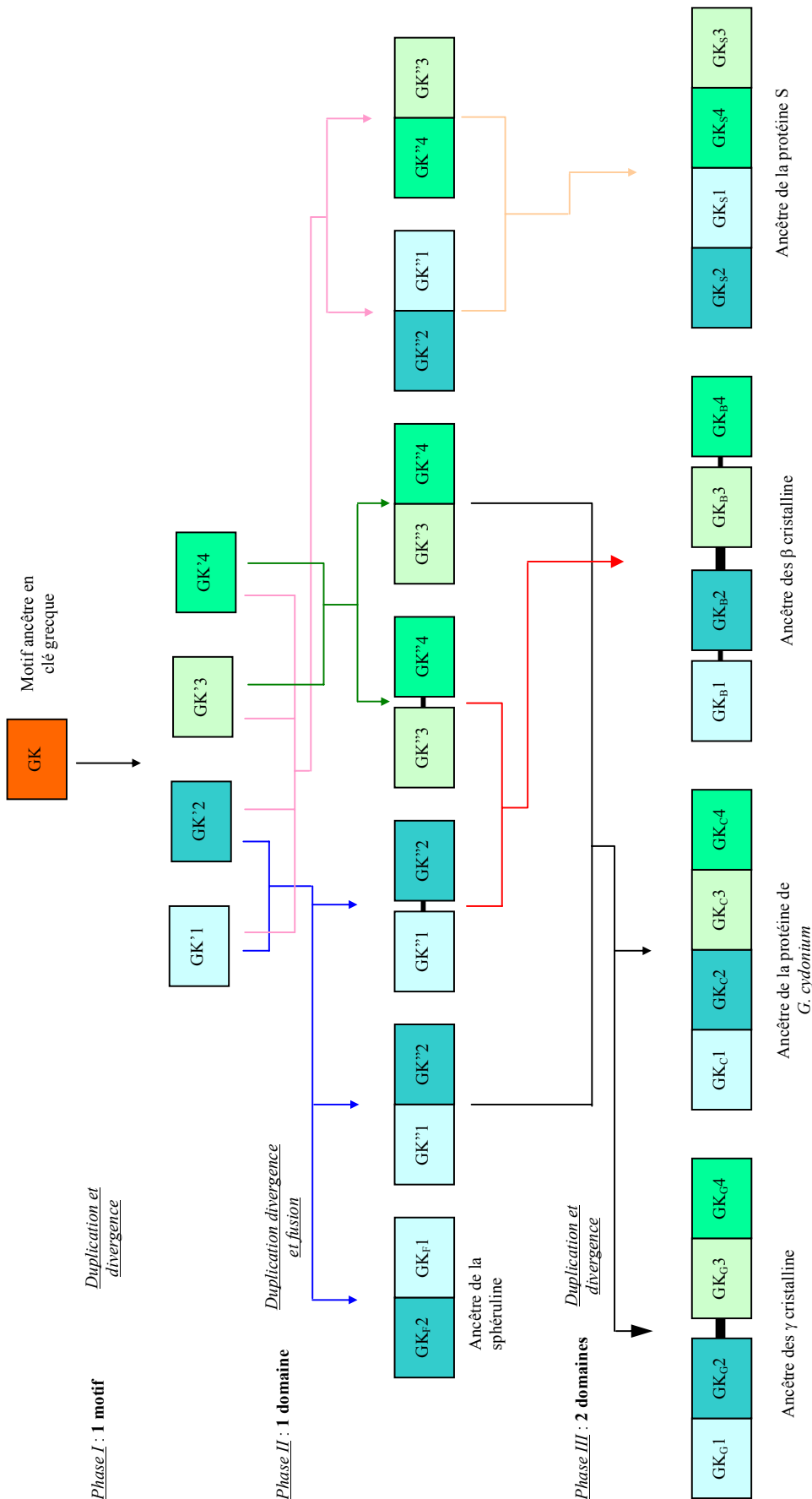


Figure 2.17 : Résumé schématique des principaux événements dans l'évolution pour aboutir aux différents types de gènes de la superfamille des β/γ cristallines. GK représente le motif ancestral en clé grecque. Les différentes phases qui se succèdent sont composées de duplication, de divergence et de fusion des gènes. Les indices indiquent F, S, C, G et B se réfèrent respectivement aux protéines ancestrales de la sphéruline 3a, de la protéine S, de la protéine de *G. cydonium*, de la γ cristalline et de la β cristalline. Les segments en noirs représentent les introns inter-motifs (trait fin) et les introns inter domaines (trait épais). (D'après D'Allesio, 2002).

2.1.5 L'eau et les γ cristallines

Les γ cristallines, qui occupent principalement le centre du cristallin (voir tableau 2.1), contribuent grandement à concentrer la lumière, grâce à la mise en place d'un gradient de l'indice de réfraction entre le cortex et le nucléus. Un gradient de contenu en eau, a été démontré comme étant, inversement proportionnel au précédent cité (Rink 1978). En effet, grâce à leur surface extrêmement chargée, les γ cristallines ont la possibilité de former des interactions plus ou moins stables avec le solvant environnant qu'est l'eau. Cette solvataion peut être visible en cristallographie et représentée par différents niveaux d'hydratation. Plusieurs γ cristallines ont déjà révélé de très grands réseaux d'hydratation autour des protéines (Najmudin et al. 1993; Kumaraswamy 1996). Par exemple, la γ B cristalline a mis en évidence un réseau de 64 molécules d'eau ininterrompu parmi les 394 molécules d'eau référencées (voir figure 2.18).

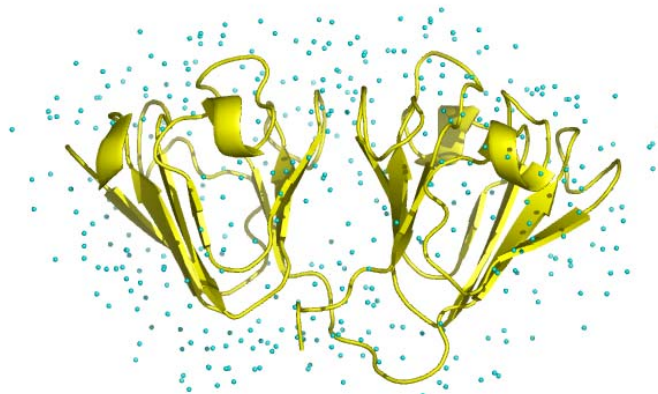


Figure 2.18 : La γ B cristalline (en jaune : 1AMM) est entourée de 394 molécules structurées (représentées par des sphères de couleur cyan) stabilisées par des liaisons hydrogène.

2.1.5.1 L'eau et la biologie

La vie dépend de l'eau. L'eau est un constituant essentiel de nombreux organismes : il représente jusqu'à 70% de la masse des cellules. Elle y joue un rôle prépondérant dans le métabolisme de ces dernières. La plupart des processus biologiques dans l'organisme ont besoin d'être dans un environnement aqueux :

- Elles permettent le transport de métabolites et le transfert de protons au sein de tout l'organisme. Par exemple, le composant majeur des 5 à 6 litres de sang du corps humain est l'eau.
- Elles favorisent le repliement tridimensionnel (au niveau de la structure tertiaire), ainsi que l'appariement des protéines (au niveau quaternaire) en générant des zones hydrophobes par regroupement de certains acides aminés. Les acides aminés dont les radicaux sont hydrophobes ont plus d'affinité entre eux qu'avec les molécules d'eau entourant la protéine. La chaîne a donc tendance à se replier de façon à les regrouper entre eux au centre de la molécule, sans contact direct avec l'eau. Cet effet hydrophobe est responsable de nombreuses interactions de reconnaissance entre molécules.
- La formation de couches d'hydratation et le remplissage de cavité interne rendent plus stable la structure 3D des protéines. Cette stabilisation est en grande partie due à l'élaboration de liaisons hydrogène.
- Les molécules d'eau sont aussi essentielles au bon fonctionnement des protéines. Généralement, les protéines perdent leurs fonctions quand elles sont déshydratées. Dans de nombreux cas, les processus enzymatiques nécessitent une ou plusieurs molécules d'eau pour réaliser une réaction de type chimique pendant laquelle, des atomes d'hydrogène peuvent être échangés.

2.1.5.2 L'hydratation des γ cristallines

Dans le cas des γ cristallines, bien que celles-ci ne présentent pas de réelles activités enzymatiques, il est néanmoins certain que les molécules d'eau jouent un rôle important dans leur stabilité et par la suite, dans leur capacité à s'ordonner dans les cellules du cristallin. Or, les γ cristallines ne possédaient pas la même quantité d'eau structurée. Ceci a été démontré à la fois par des données cristallographiques (voir tableau 2.2), mais aussi par des études de calorimétrie (Bettelheim et al. 1994). L'analyse par calorimétrie différentielle à balayage (Differential Scanning Calorimetry, DSC) est une technique qui permet de mesurer les variations d'enthalpie qui résultent de changements d'état d'un corps soumis à une variation de température. De ce fait,

il est possible de calculer la quantité d'eau qui ne gèle pas autour d'une protéine. Cette eau correspond, en fait, à une couche d'hydratation extrêmement structurée par une association très forte avec des résidus polaires de la molécule ou qui est prise au piège dans des cavités. Dans l'étude du groupe de Bettelheim, ils ont tenté de calculer la masse d'eau structurée autour des γ cristallines du bœuf. En résumé, les auteurs montrent que les γ cristallines de bœuf peuvent être départagées en deux groupes : la γ C et la γ D qui peuvent être entourées de grande quantité d'eau ; et la γ B et la γ E dont la capacité est plus réduite. L'ordre des γ cristallines pouvant immobiliser des molécules d'eau est le suivant : γ D > γ C > γ E > γ B. Mais cet ordre est différent des profils de transition de phase ou des observations faites en cristallographie. Il s'avère donc nécessaire de prendre en compte la forme totale des protéines et leur surface accessible pour expliquer les différents états de solvatation. Par exemple, la cristalline γ E de bœuf est plus sphérique et plus compacte que la cristalline γ B. Une accessibilité moins importante de la γ E par rapport à la γ B devrait réduire, ainsi, la possibilité d'avoir des résidus polaires en surface pouvant interagir avec des molécules d'eau.

Les différentes propriétés d'hydratation des γ cristallines et leurs interprétations sur la base de différences structurales sont étroitement liées aux comportements de transition de phase et d'agrégation. Ainsi il est certain, qu'en plus de leur agencement structural, les propriétés de solvatation des γ cristallines et leurs interactions spécifiques de type homo- ou hétérogènes, sont indispensables. La transparence et le pouvoir réfractif du cristallin dépendent de cette hydratation. Cette dernière est un phénomène complexe, qui doit prendre en compte :

- la **distribution** très ordonnée des cristallines dans le cristallin,
- leur **concentration** jusqu'à 850 mg/ml (Philipson 1969),
- et les **interactions** protéine- protéine et eau- protéine des cristallines.

2.2 Pourquoi envisager de la cristallographie aux neutrons ?

Il est fortement supposé que les interactions eau – γ cristalline sont impliquées dans les séparations de phase conduisant à l'opacification du cytoplasme des fibres du cristallin observée dans certaines formes de cataractes (Siezen et al. 1985; Thomson et al. 1987; Broide et al. 1991). Toutes les structures cristallographiques des γ cristallines ont montré de nombreuses paires d'ions

en surface des protéines conduisant à une densité de charge de la surface très importante mais dont la charge totale était faible. Les charges positives et négatives sont bien équilibrées avec les arginines et les aspartates formant des interactions. Ce grand nombre de groupes chargés situés à la surface, a aussi la possibilité d'interagir avec des molécules d'eau. Pour comprendre quel est le rôle de ces groupements et de ces molécules d'eau, il est important d'identifier les caractères électrostatiques ou covalents des interactions entre les groupes guanidiniums et carboxylates, mais aussi de se rendre compte de la géométrie des molécules responsables des séparations de phases. Cette approche serait une base structurale pour étudier les effets des interactions protéine-eau lorsqu'un point critique est approché.

Une des meilleures techniques pour une approche structurale est la diffraction par un cristal unique. De plus, le choix de ces protéines pour une étude cristallographique est essentiellement dirigé par le fait que le contenu en solvant (l'eau) est sensiblement identique dans le cristallin (~ 60%) (Rink 1978) et dans un cristal de protéine (~50%).

La détermination de structures de macromolécules biologiques par la cristallographie de rayons X ou de neutrons est devenue très bénéfique pour la compréhension et l'interprétation de nombreux processus biologiques tels que les mécanismes d'une catalyse enzymatique, les détails d'une interaction protéine-ligand, certaines fonctions moléculaires, et les problèmes de repliement protéique. La diffraction de rayons X par un cristal unique est une technique extrêmement populaire de nos jours : plus de 24000 structures de protéines et acides nucléiques ont été obtenues par cette méthode. En comparaison, les structures d'environ 4000 macromolécules, principalement des protéines ne dépassant pas 150 résidus, ont été résolues par spectroscopie de résonance magnétique nucléaire (RMN), technique fournissant un modèle de la molécule en solution. La diffraction de rayons X est une technique très puissante puisqu'elle permet d'obtenir, dans le meilleur des cas, un modèle où les atomes de la protéine sont distincts les uns des autres (résolution inférieure à la distance inter-atomique). Malgré la prééminence de la technique aux rayons X, les études de diffraction aux neutrons peuvent avoir aussi un rôle déterminant dans la compréhension des arrangements biologiques du fait que ces techniques possèdent certains avantages que n'a pas la diffraction aux rayons X. La capacité de localiser les positions des atomes d'hydrogènes, et donc par conséquent, d'orienter les molécules d'eau

situées à l'intérieur de la protéine ou à sa proche surface, est un des avantages les plus marquants de cette technique de diffraction des neutrons.

2.2.1 Visualisation des hydrogènes par la cristallographie aux neutrons

Les interactions qu'ont les rayons X et les neutrons avec les atomes d'un cristal, diffèrent sur quelques points. En effet, la principale différence est causée par le fait que les rayons X et les neutrons ne sont pas diffractés par les mêmes composants des atomes (Wlodawer 1982). Pour les rayons X, les centres de diffusion sont les nuages électroniques des atomes. L'information obtenue sera relative à la structure et la dynamique de ce nuage électronique. Ceci a pour conséquence que les amplitudes des rayons X diffusés sont proportionnelles aux nombres d'électrons entourant chaque atome (voir tableau 2.3). Le pouvoir diffusant de chaque atome, défini par le facteur de forme f_i , dépend cette fois de l'angle de diffusion, de la longueur d'onde utilisée, des vibrations thermiques de l'atome, et est proportionnel au nombre d'électrons. Par conséquent, l'hydrogène, avec un seul électron, va faiblement diffuser les rayons X et donc ne sera pas visible hormis une ultra haute résolution comme pour l'aldose réductase résolue à 0,66 Å (Howard et al. 2004). Généralement, à partir d'un jeu de données de diffraction des rayons X, à haute et basse résolution, la localisation et l'orientation détaillée de l'hydratation et de la protonation peuvent être seulement supposées par la stéréochimie des atomes environnants (C, N ou O) et les liaisons hydrogène.

<i>Elément</i>	<i>Isotope</i>	<i>Nombre d'électrons</i>	<i>Facteur de forme f_i (à $2\theta=0$) (10^{-12} cm)</i>	<i>Longueur de diffusion cohérente pour les neutrons (en fm)</i>	<i>Longueur de diffusion incohérente pour les neutrons (en barns)</i>
H	^1H	1	0,28	-3,74	80,26
	^2H (D)	1	0,28	6,67	2,05
C	^{12}C	6	1,69	6,64	0,001
N	^{14}N	7	1,97	9,36	0,50
O	^{16}O	8	2,25	5,80	0,00
P	^{31}P	15	4,23	2,84	0,005
S	^{32}S	16	4,50	2,84	0,007

Tableau 2.3 : Caractéristiques des principaux atomes constitutifs des protéines avec leur facteur de forme, leur longueur de diffusion cohérente et incohérente pour les neutrons (Dianoux and Lander 2002).

En revanche, les centres de diffusion pour les neutrons, sont les noyaux atomiques.

L'intensité diffusée par les nucléus, peut être séparée en deux contributions : une contribution

cohérente et une contribution incohérente. La contribution cohérente de la diffusion de neutrons permet donc une localisation directe des atomes d'hydrogènes mais permet aussi de discerner les hydrogènes de son isotope, le deutérium, en utilisant leur différence de signe de longueur de diffusion cohérente (voir tableau 2.3). D'un autre côté, chaque élément a sa propre interaction caractéristique avec les neutrons mais généralement les interactions diffèrent entre les isotopes d'un même élément. Il est donc aisé de discerner et d'identifier les différents atomes d'une protéine. La contribution incohérente, dans le cas de la diffraction, contribue fortement au bruit.

En cristallographie aux rayons X, le pouvoir diffusant tombe rapidement avec l'angle de diffusion (voir figure 2.19). Il est donc difficile de localiser un atome ou groupement d'atomes, légèrement désordonné, comme les hydrogènes des molécules d'eau du fait du faible pouvoir diffractant et de ce léger désordre. Le signal est perdu dans le bruit pour la plupart du temps hormis pour les ultra hautes résolutions. En neutrons, le facteur de forme d'un atome au repos ne diminue pas. Une résolution de 2 Å est satisfaisante pour visualiser la plupart des atomes d'hydrogènes.

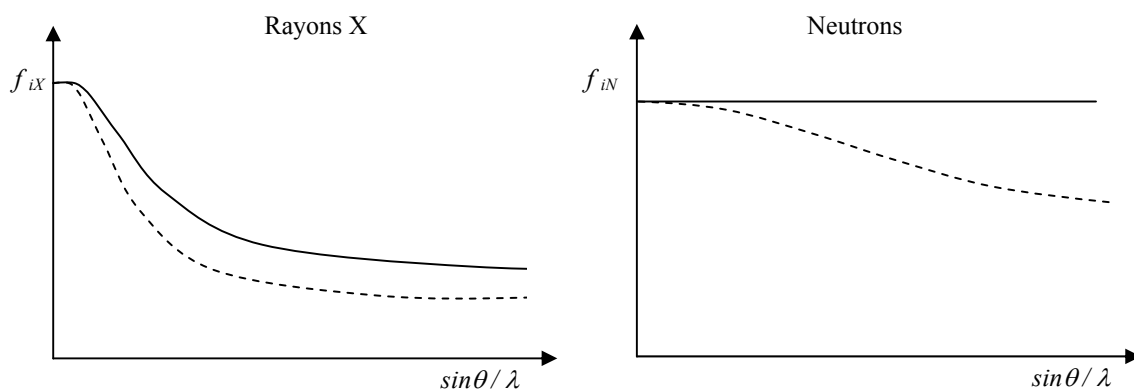


Figure 2.19 : Facteurs de forme en cristallographie aux rayons X (à gauche) et aux neutrons (à droite). En trait plein, sont représentés des atomes au repos ($B=0$) et en pointillé sont représentés des atomes vibrants sous l'effet thermique ($B>0$).

La longueur de diffusion (en tableau 2.3) est une mesure qui détermine comment un atome donné va diffracter les neutrons. Le facteur de diffusion négatif de l'hydrogène, qui est de -3,8 fermi, est causé par un changement de phase de l'onde diffusée qui est à 180° différent de la majorité des autres atomes. En pratique, cela va se traduire par des contours de densités négatives dans les cartes neutroniques de Fourier, en contraste avec les autres éléments (incluant le

deutérium) qui vont apparaître avec des densités positives (voir figure 2.20a). De plus, sur certaines cartes, la dispersion de la densité neutronique des hydrogènes et des atomes voisins directs (C, N, S...) peut donner un chevauchement significatif qui découle de l'annulation entre la contribution négative des atomes d'hydrogène et la contribution positive des autres atomes (Shu et al. 2000). Bien que cette annulation dépend de la résolution, elle résulte en des absences de densité proche des atomes de carbone et pousse le peu de densité négative restante, de l'hydrogène en dehors de sa véritable position (voir figure 2.20b).

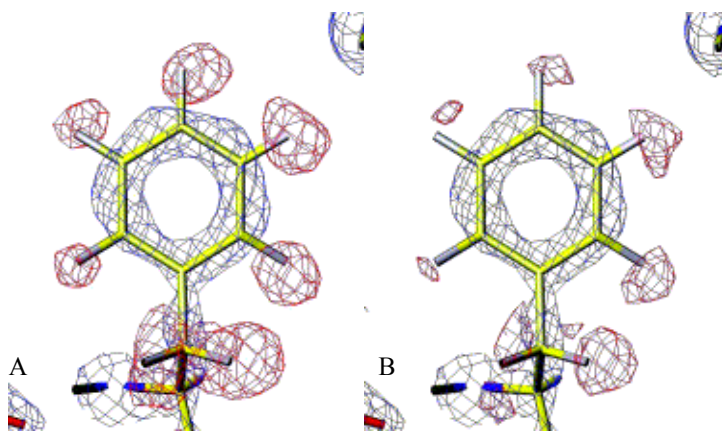


Figure 2.20 : Phénylalanine 43 de la myoglobine de cachalot résolue par cristallographie des neutrons à 1,5 Å (Ostermann et al. 2002). En bleu, est représentée la carte $2F_o - F_c$ de densité neutronique positive et en rouge, la même carte mais de densité négative. En A se trouve une carte dont les densités négatives ont été corrigées, sinon il y aurait annulation des cartes et mauvais positionnement des hydrogènes comme en B. La correction est faite en soustrayant les cartes de densité $2F_o - F_c$ calculées avec et sans les hydrogènes.

2.2.2 Diffusion incohérente des atomes d'hydrogène

Du fait de sa forte contribution au bruit, la diffusion incohérente est une des causes majeures rendant difficiles les collectes de données, de bonne qualité, en cristallographie aux neutrons. La diffusion des atomes d'hydrogène est quasiment responsable à elle seule de la diffusion incohérente. Du fait que la diffusion incohérente de ces atomes est 40 à $8 \cdot 10^5$ fois plus importante que les autres atomes et qu'elle est de 8 à 40 fois plus importante que la diffusion cohérente de ces mêmes atomes, ceci va être la source de deux problèmes. Premièrement, les neutrons diffusés de manière incohérente ne sont plus disponibles pour la diffusion cohérente, ce qui équivaut à une absorption : l'information risque d'être perdue. Et deuxièmement, les neutrons vont diffuser de façon isotrope, et donc le niveau de bruit de fond va être augmenté. Le rapport $I/\sigma(I)$ (signal sur bruit de fond) sera donc diminué au cours de l'expérience et toutes les tâches de diffraction seront difficilement repérables. Il est intéressant de remarquer qu'en cristallographie des rayons X, c'est le support du cristal (la boucle et la solution cryogénique) et l'air environnant

qui sont la principale contribution du bruit de fond. Contrairement à la cristallographie aux neutrons, ce n'est pas l'échantillon qui est le responsable majoritaire de la diffusion incohérente.

2.2.3 La réduction du bruit de fond en cristallographie des neutrons

Etant donné que le bruit de fond est essentiellement dû aux atomes d'hydrogène de l'échantillon, il est donc possible de surmonter cet effet en remplaçant les hydrogènes par son isotope, le deutérium, dont la longueur de diffusion incohérente est 40 fois moins importante (voir tableau 2.3). Par la même occasion, le problème d'annulation des cartes de densité neutronique n'existera plus car la longueur de diffusion cohérente du deutérium est positive, tout comme les autres constituants du cristal.

D'un point pratique, la substitution peut être effectuée, en partie, par croissance ou trempage d'un cristal dans une solution deutériée. La croissance cristalline dans des solutions où le précipitant est deutérié, a été réalisée, en outre, pour la rubredoxine d'une achaeobactérie marine *P. furiosus* et un de ses mutants (Chatake et al. 2003). De plus nombreux cas de trempage d'un cristal hydrogéné dans une solution deutériée ont été publiés. C'est le cas de la concanavoline A de haricot (Habash et al. 2000), de la ribonucléase A (RNase-A) du pancréas de bœuf (Wlodawer et al. 1987), de l'endothiapepsine (Coates et al. 2001) et de la myoglobine (Phillips et Schoenborn 1981). Dans ce type d'expérience, seuls les hydrogènes échangeables vont être remplacés par des deutériums. Les atomes d'hydrogènes échangeables sont, les atomes reliés à des atomes d'oxygène, d'azote ou de soufre et sont donc :

- (i) les hydrogènes de toutes les liaisons amides de la chaîne principale (excepté pour les prolines) ;
- (ii) les hydrogènes des chaînes latérales des résidus tyrosine, tryptophane, cystéine, sérine, thréonine, lysine, arginine, histidine, asparagine et glutamine au niveau des O-H, des N-H et des S-H ;
- (iii) les hydrogènes des chaînes latérales des acides aminés acides, l'acide glutamique et l'acide aspartique, quand ils sont sous forme protonnée.

Les hydrogènes non échangeables sont donc ceux qui sont reliés aux atomes de carbones des chaînes principales et latérales. L'échange des hydrogènes par des deutériums peut être relativement long surtout si les hydrogènes sont impliqués dans des structures secondaires comme les feuillets β des γ cristallines.

Pour une protéine comme la γ E cristalline de rat, il y a 356 hydrogènes échangeables (soit 26% des hydrogènes de la protéine) et 1010 hydrogènes non échangeables (soit 74% des hydrogènes de la protéine). Bien que le remplacement des hydrogènes échangeables va réduire considérablement le bruit de fond, il est important d'augmenter l'enrichissement en deutérium pour diminuer le bruit de fond et rendre plus aisée la lecture des cartes de densité neutronique. Pour se faire, il est possible de deutérer en totalité l'échantillon protéique en produisant *in vivo* la protéine. Cette méthode qui consiste à cloner et à exprimer dans un système bactérien, capable de croître sur un milieu complètement deutérié. Cette approche a été utilisée pour préparer un échantillon perdeutérié de myoglobine (Shu et al. 2000) et **c'est un des buts principaux de cette thèse.**

2.2.4 Faible flux de neutrons

Un autre type de limitation est associé à la cristallographie aux neutrons des protéines : le faible flux de neutrons qui arrivent au niveau de l'échantillon. Le nombre moyen de neutrons incidents au niveau du cristal est de $3 \cdot 10^7$ neutrons.cm⁻².sec⁻¹ à ILL pour le LAue Diffractometer (www.ill.fr) alors que le nombre moyen de photons à l'ESRF qui frappe un cristal est d'environ 10^6 à 10^7 fois plus intense. Pour palier à ce problème, des temps d'exposition plus longs (quelques semaines à quelques mois) sont nécessaires pour obtenir des jeux de données avec des statistiques raisonnables. Mais il est important de souligner que l'exposition d'un cristal aux neutrons ne produit pas de dommages radiatifs comme les rayons X.

A cause de ce faible flux de neutrons, il est nécessaire d'utiliser des cristaux de taille importante pour collecter des données de diffraction. Dans le passé, des cristaux « énormes » ont été utilisés dans les premières expériences de diffraction neutronique. Par exemple, Wlodawer (1980) a utilisé un cristal de 30 mm³ de RNase A pour collecter des données aux neutrons.

Depuis, l'évolution des appareils et le développement d'un quasi Laue diffractomètre (LADI) à l'ILL, qui permet d'utiliser une gamme de longueur d'onde, ont limité les collectes de données à quelques jours.

2.2.5 La diffraction polychromatique dite diffraction de Laue

En 1912, Friedrich, Knipping et Laue ont obtenu le premier cliché de diffraction polychromatique. Ils avaient exposé un cristal de sulfate de cuivre par une gamme continue de longueur d'onde de rayons X. Sur LADI, en illuminant un cristal par un rayon polychromatique de rayons X ou de neutrons (de longueur d'onde compris entre λ_{\min} et λ_{\max}), un grand nombre de réflexions va être généré et enregistré en un seul cliché (voir figure 2.21).

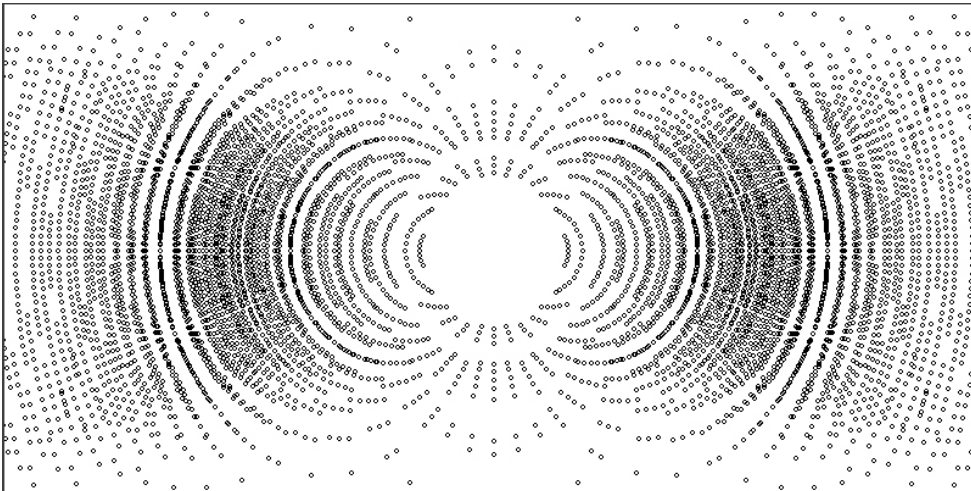


Figure 2.21 : Cliché de diffraction quasi-Laue théorique de la γE cristalline de rat pour une gamme de longueur d'onde comprise entre 2,5 et 4 Å, et une résolution de 2 Å (d'après le site www.ill.fr/YellowBook/LADI/). Le volume de la maille cristalline est d'environ 166 300 Å.

En utilisant les rappels cristallographie du chapitre 5, il est possible de faire une construction d'Ewald qui contient plusieurs sphères de rayon compris entre $1/\lambda_{\min}$ et $1/\lambda_{\max}$. Les points du réseau réciproque qui se trouveront sur une de ces sphères vont donner lieu à de la diffraction (voir figure 2.22). Avec un détecteur approprié, il est possible d'enregistrer en un seul cliché, tous les nœuds de l'espace réciproque qui remplissent les conditions de diffraction (voir chapitre 5) et qui sont à l'intersection de l'une des sphères.

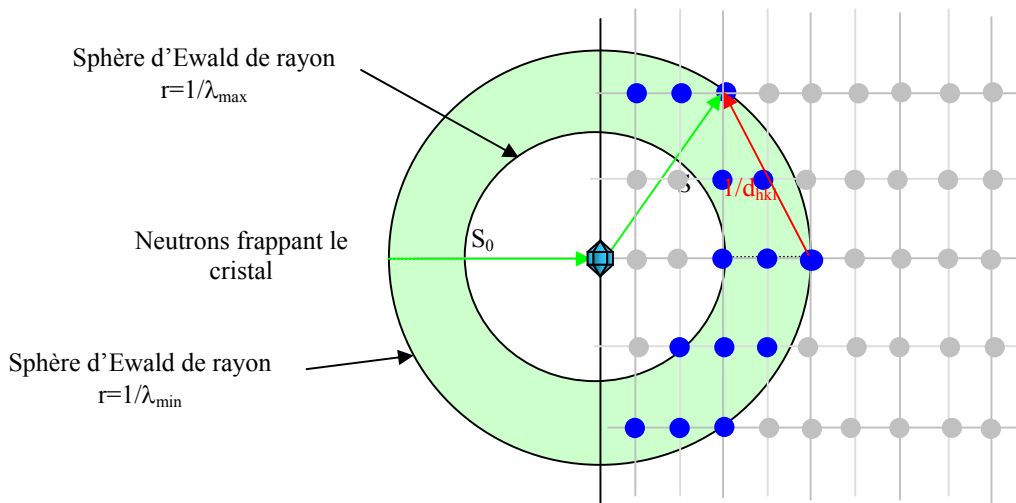


Figure 2.22 : Construction d'Ewald pour la diffraction de Laue. Le cristal va être illuminé par des neutrons d'une gamme continue de longueur d'onde comprises entre λ_{\min} et λ_{\max} . Tous les points du réseau réciproque qui vont se trouver sur une des sphères dont le rayon est compris entre $1/\lambda_{\max}$ et $1/\lambda_{\min}$, et qui remplissent les conditions de diffraction (ces nœuds sont représentés en bleu), vont former des tâches de réflexions sur le détecteur.

En cristallographie aux rayons X, cette méthode de Laue a été développée pour obtenir en un temps très court (de l'ordre de la picoseconde à la seconde) un jeu de données complet. Il est donc possible de suivre tout changement structural au cours d'une réaction catalytique mesurable dans l'échelle de temps indiquée précédemment. Pour la cristallographie des neutrons, la technique de Laue est utilisée pour augmenter le flux de neutrons utile sur l'échantillon, afin de diminuer la durée de collecte (Bon et al. 1999).

Cependant, la technique de Laue a une limitation dans le traitement des réflexions. Pour intégrer l'intensité diffractée de chaque tâche de Bragg, il faut que ces dernières soient suffisamment espacées, donc il n'y ait pas de recouvrement spatial des tâches. Plusieurs paramètres vont affecter le nombre et le recouvrement des réflexions sur un cliché de diffraction (Bon 1998):

- la taille de la maille cristalline. Plus la maille cristalline est petite, plus la maille réciproque est large, donc plus les tâches de diffraction seront espacées.
- plus la gamme de longueur d'onde est grande, plus le nombre de réflexions est grand.
- plus la limite de résolution est bonne, plus le nombre de réflexions est important.
- plus la distance détecteur-cristal est importante, plus les tâches de diffraction seront espacées.

- la divergence du faisceau et la mosaïcité du cristal influent beaucoup sur la forme et la largeur des tâches de Bragg.

2.3 Bilans

Les différents états de protonation des résidus chargés se trouvant en surface des γ cristallines et l'orientation de ces groupements de surface, ainsi que le réseau hydraulique qui existe autour des γ cristallines, sont des considérations essentielles pour comprendre les différents comportements d'agrégation de ces protéines lors de l'opacification du cristallin. La méthode de diffraction neutronique est certainement la technique la plus appropriée pour observer les différents états des groupements ammoniums et carboxyliques, tout comme la position et l'orientation des molécules d'eau entourant la γ cristalline.

En contrepartie de ces informations, cette technique possède quelques limitations qui peuvent être surmontées.

Le premier inconvénient est la diffusion incohérente provenant de l'hydrogène, qui peut être considérablement réduite en remplaçant ce dernier par son isotope, le deutérium. De façon, à substituer tous les atomes de la protéine, nous avons développé des techniques de sur-expression de protéines recombinantes en système bactérien, et qui seront discutées dans le chapitre suivant.

Mais l'inconvénient majeur de la cristallographie aux neutrons est que le faible flux de neutrons provenant de réactions nucléaires est beaucoup moins intense que celui des photons X d'une source synchrotron. Pour palier à ce problème, l'utilisation de cristaux de volume supérieur à 1 mm^3 est nécessaire. Dans le chapitre 7, seront discutées les difficultés et les différentes techniques de cristallisation utilisées pour obtenir des cristaux de protéine perdeutériée, en vue d'expérience de diffraction neutronique. Ce chapitre est aussi consacré aux conclusions et aux perspectives de ce travail.

Bibliographie :

- Basak, A., Bateman, O., Slingsby, C., Pande, A., Asherie, N., Ogun, O., Benedek, G.B., and Pande, J. 2003. High-resolution X-ray crystal structures of human gammaD crystallin (1.25 Å) and the R58H mutant (1.15 Å) associated with aculeiform cataract. *J Mol Biol* **328**: 1137-1147.
- Bax, B., Lapatto, R., Nalini, V., Driessen, H., Lindley, P.F., Mahadevan, D., Blundell, T.L., and Slingsby, C. 1990. X-ray analysis of beta B2-crystallin and evolution of oligomeric lens proteins. *Nature* **347**: 776-780.
- Benedek, G.B. 1971. Theory of the transparency of the eye. *Appl. Opt.* **10**: 459-473.
- Bettelheim, F.A., Reid, M.B., and Garland, D. 1994. Hydration of gamma-crystallins. *Exp Eye Res* **58**: 219-223.
- Bon, C. 1998. Etude structural et dynamique de l'eau à l'intérieur et autour du lysozyme triclinique d'oeuf de poule. Université Joseph Fourier, Grenoble.
- Bon, C., Lehmann, M.S., and Wilkinson, C. 1999. Quasi-Laue neutron-diffraction study of the water arrangement in crystals of triclinic hen egg-white lysozyme. *Acta Crystallogr D Biol Crystallogr* **55**: 978-987.
- Broide, M.L., Berland, C.R., Pande, J., Ogun, O.O., and Benedek, G.B. 1991. Binary-liquid phase separation of lens protein solutions. *Proc Natl Acad Sci U S A* **88**: 5660-5664.
- Castan, A., and Enfors, S.O. 2002. Formate accumulation due to DNA release in aerobic cultivations of *Escherichia coli*. *Biotechnol Bioeng* **77**: 324-328.
- Chatake, T., Ostermann, A., Kurihara, K., Parak, F.G., and Niimura, N. 2003. Hydration in proteins observed by high-resolution neutron crystallography. *Proteins* **50**: 516-523.
- Chirgadze, Y., Driessen, H., Wright, G., Slingsby, C., Hay, R.E., and Lindley, P.F. 1996. Structure of the bovine Eye lens gammaD crystallin at 1.95Å. *Acta Crysta* **D52**: 712-721.
- Clout, N.J., Kretschmar, M., Jaenicke, R., and Slingsby, C. 2001. Crystal structure of the calcium-loaded spherulin 3a dimer sheds light on the evolution of the eye lens betagamma-crystallin domain fold. *Structure (Camb)* **9**: 115-124.
- Coates, L., Erskine, P.T., Wood, S.P., Myles, D.A., and Cooper, J.B. 2001. A neutron Laue diffraction study of endothiapepsin: implications for the aspartic proteinase mechanism. *Biochemistry* **40**: 13149-13157.
- Crabbe, M.J., and Goode, D. 1995. Protein folds and functional similarity; the Greek key/immunoglobulin fold. *Comput Chem* **19**: 343-349.
- D'Alessio, G. 2002. The evolution of monomeric and oligomeric betagamma-type crystallins. Facts and hypotheses. *Eur J Biochem* **269**: 3122-3130.
- de Jong, W., Caspers, G., and Leunissen, J. 1998. Genealogy of the alpha-crystallin--small heat-shock protein superfamily. *Int J Biol Macromol.* **22**: 151-162.
- Descartes, R. 1637. *La dioptrique*.
- Dianoux, A.J., and Lander, G. 2002. *Neutron Data Booklet*.
- Habash, J., Raftery, J., Nuttall, R., Price, H.J., Wilkinson, C., Kalb, A.J., and Helliwell, J.R. 2000. Direct determination of the positions of the deuterium atoms of the bound water in -concanavalin A by neutron Laue crystallography. *Acta Crystallogr D Biol Crystallogr* **56 (Pt 5)**: 541-550.
- Heon, E., Priston, M., Schorderet, D.F., Billingsley, G.D., Girard, P.O., Lubsen, N., and Munier, F.L. 1999. The gamma-crystallins and human cataracts: a puzzle made clearer. *Am J Hum Genet* **65**: 1261-1267.
- Horwitz, J. 1992. Alpha-crystallin can function as a molecular chaperone. *Proc Natl Acad Sci U S A* **89**: 10449-10453.
- Howard, E.I., Sanishvili, R., Cachau, R.E., Mitschler, A., Chevrier, B., Barth, P., Lamour, V., Van Zandt, M., Sibley, E., Bon, C., et al. 2004. Ultrahigh resolution drug design I: details of interactions in human aldose reductase-inhibitor complex at 0.66 Å. *Proteins* **55**: 792-804.

- Ingolia, T.D., and Craig, E.A. 1982. Four small *Drosophila* heat shock proteins are related to each other and to mammalian alpha-crystallin. *Proc Natl Acad Sci U S A* **79**: 2360-2364.
- Jacob, M., Schindler, T., Balbach, J., and Schmid, F.X. 1997. Diffusion control in an elementary protein folding reaction. *Proc Natl Acad Sci U S A* **94**: 5622-5627.
- Kantorow, M., and Piatigorsky, J. 1994. Alpha-crystallin/small heat shock protein has autokinase activity. *Proc Natl Acad Sci U S A* **91**: 3112-3116.
- Kim, K.K., Kim, R., and Kim, S.H. 1998. Crystal structure of a small heat-shock protein. *Nature* **394**: 595-599.
- Klemenz, R., Frohli, E., Steiger, R.H., Schafer, R., and Aoyama, A. 1991. Alpha B-crystallin is a small heat shock protein. *Proc Natl Acad Sci U S A* **88**: 3652-3656.
- Krasko, A., Muller, I.M., and Muller, W.E. 1997. Evolutionary relationships of the metazoan beta gamma-crystallins, including that from the marine sponge *Geodia cydonium*. *Proc R Soc Lond B Biol Sci* **264**: 1077-1084.
- Kumaraswamy, V.S., Lindley, P. F., Slingsby, C., Glover, I. D. 1996. An eye lens protein-water structure: 1.2 angstrom resolution structure of gamma B-crystallin at 150K. *Acta Crystallogr D Biol Crystallogr* **52**: 611.
- Liu, C., Asherie, N., Lomakin, A., Pande, J., Ogun, O., and Benedek, G.B. 1996. Phase separation in aqueous solutions of lens gamma-crystallins: special role of gamma s. *Proc Natl Acad Sci U S A* **93**: 377-382.
- Lubsen, N.H., Aarts, H.J., and Schoenmakers, J.G. 1988. The evolution of lenticular proteins: the beta- and gamma-crystallin super gene family. *Prog Biophys Mol Biol* **51**: 47-76.
- Mayr, E.M., Jaenicke, R., and Glockshuber, R. 1994. Domain interactions and connecting peptides in lens crystallins. *J Mol Biol* **235**: 84-88.
- Najmudin, S., Nalini, V., Driessen, H., Slingsby, C., Blundell, T., S., M.D., and Lindley, P.F. 1993. Structure of the bovine eye lens protein gammaB(gammaII)-crystallin at 1.47 Å. *Acta Crystallogr D Biol Crystallogr* **49**: 223-233.
- Norledge, B.V., Hay, R.E., Bateman, O.A., Slingsby, C., and Driessen, H.P. 1997. Towards a molecular understanding of phase separation in the lens: a comparison of the X-ray structures of two high Tc gamma-crystallins, gammaE and gammaF, with two low Tc gamma-crystallins, gammaB and gammaD. *Exp Eye Res* **65**: 609-630.
- Ohno, A., Tate, S., Seeram, S.S., Hiraga, K., Swindells, M.B., Oda, K., and Kainosho, M. 1998. NMR structure of the *Streptomyces* metalloproteinase inhibitor, SMPI, isolated from *Streptomyces nigrescens* TK-23: another example of an ancestral beta gamma-crystallin precursor structure. *J Mol Biol* **282**: 421-433.
- Ostermann, A., Tanaka, I., Engler, N., Niimura, N., and Parak, F.G. 2002. Hydrogen and deuterium in myoglobin as seen by a neutron structure determination at 1.5 Å resolution. *Biophys Chem* **95**: 183-193.
- Pande, J., Ogun, O., Nath, C., and Benedek, G. 1993. Suppression of phase separation in bovine gamma IV crystallin solutions: effect of modification by charged versus uncharged polar groups. *Exp Eye Res* **57**: 257-264.
- Philipson, B. 1969. Distribution of protein within the normal rat lens. *Invest Ophthalmol* **8**: 258-270.
- Phillips, S.E., and Schoenborn, B.P. 1981. Neutron diffraction reveals oxygen-histidine hydrogen bond in oxymyoglobin. *Nature* **292**: 81-82.
- Ray, M.E., Wistow, G., Su, Y.A., Meltzer, P.S., and Trent, J.M. 1997. AIM1, a novel non-lens member of the betagamma-crystallin superfamily, is associated with the control of tumorigenicity in human malignant melanoma. *Proc Natl Acad Sci U S A* **94**: 3229-3234.
- Rink, H. 1978. The Water content in bovine lenses during Aging. *Intrdiscipl. Topics Geront.* **12**: 271-277.
- Shu, F., Ramakrishnan, V., and Schoenborn, B.P. 2000. Enhanced visibility of hydrogen atoms by neutron crystallography on fully deuterated myoglobin. *Proc Natl Acad Sci U S A* **97**: 3872-3877.

- Siezen, R.J., Fisch, M.R., Slingsby, C., and Benedek, G.B. 1985. Opacification of gamma-crystallin solutions from calf lens in relation to cold cataract formation. *Proc Natl Acad Sci U S A* **82**: 1701-1705.
- Stafford, M.J. 2001. The histology and biology of the lens. *Optometry Today Optics Today*: 23-30.
- Thomson, J.A., Schurtenberger, P., Thurston, G.M., and Benedek, G.B. 1987. Binary liquid phase separation and critical phenomena in a protein/water solution. *Proc Natl Acad Sci U S A* **84**: 7079-7083.
- Trinkl, S., Glockshuber, R., and Jaenicke, R. 1994. Dimerization of beta B2-crystallin: the role of the linker peptide and the N- and C-terminal extensions. *Protein Sci* **3**: 1392-1400.
- Trokel, S. 1962. The physical basis for transparency of the crystalline lens. *Invest Ophthalmol* **1**: 493-501.
- van Montfort, R.L., Basha, E., Friedrich, K.L., Slingsby, C., and Vierling, E. 2001. Crystal structure and assembly of a eukaryotic small heat shock protein. *Nat Struct Biol* **8**: 1025-1030.
- Wistow, G., Jaworski, C., and Rao, P.V. 1995. A non-lens member of the beta gamma-crystallin superfamily in a vertebrate, the amphibian *Cynops*. *Exp Eye Res* **61**: 637-639.
- Wistow, G., Slingsby, C., Blundell, T., Driessen, H., De Jong, W., and Bloemendal, H. 1981. Eye-lens proteins: the three-dimensional structure of beta-crystallin predicted from monomeric gamma-crystallin. *FEBS Lett* **133**: 9-16.
- Wistow, G., Summers, L., and Blundell, T. 1985. *Myxococcus xanthus* spore coat protein S may have a similar structure to vertebrate lens beta gamma-crystallins. *Nature* **315**: 771-773.
- Wistow, G.J., and Piatigorsky, J. 1988. Lens crystallins: the evolution and expression of proteins for a highly specialized tissue. *Annu Rev Biochem* **57**: 479-504.
- Wlodawer, A. 1980. Studies of Ribonuclease-A by X-ray and Neutron Diffraction. *Acta Crysta* **B36**: 1826-1831.
- Wlodawer, A. 1982. Neutron diffraction of crystalline proteins. *Prog Biophys Mol Biol* **40**: 115-159.
- Wlodawer, A., Deisenhofer, J., and Huber, R. 1987. Comparison of two highly refined structures of bovine pancreatic trypsin inhibitor. *J Mol Biol* **193**: 145-156.

Chapitre 3

Production de γ cristallines totalement deutériées

Abstract

The aim of this chapter is to develop reliable and economical protocols for production of fully deuterated biomolecules by bacteria. This requires, firstly, an expression vector, a pET system, containing the cDNA of the γ crystallin gene, able to remain stable in the bacteria during high cell density culture (with a kanamycin resistant gene). Secondly, the preparation of deuterium-tolerant bacterial strains is necessary and performed by a multi stage adaptation with plating on a solid deuterated minimal medium and repeated liquid deuterated culture. An *E. coli* strain, *BL21 (DE3)*, containing this plasmid is adapted to growth on deuterated glycerol medium.

In order to achieve high cell densities of cultivations of bacterial culture it is usually necessary to restrict the growth rate in order to avoid production of inhibiting by-products from high consumption rates of substrate. The fed batch technique is suitable to overcome this problem. With this method, the adapted strain shows good, consistent growth rate (0.24 h^{-1}) in deuterated medium but lower than a hydrogenated culture (0.34 h^{-1}); however yields of a very well adapted system look similar at the end of the culture (1.85 g of cell past per g of used glycerol). Expression of hydrogenated and deuterated recombinant γ E-crystallin is achieved in fed batch cultivations with identical expression levels.

3.1 Expression de γ cristallines en bio-réacteur

Nombreuses sont les études sur les γ cristallines. Pour la plupart d'entre elles, les protéines sont extraites, directement, à partir des yeux qui les produisent de façon naturelle. C'est ainsi que les cristallins de veaux, de bœufs, de rats, et d'hommes ont servi de sources riches en γ cristallines puisqu'il est possible d'en trouver jusqu'à 400 mg/ml. Le problème avec l'étude qui suit, est la nécessité d'obtenir une protéine où la totalité des hydrogènes est remplacée par des atomes de deutériums, pour des raisons qui ont été précédemment citées. La production de protéines deutériées chez un mammifère n'a pas été encore établie et il semble que cela soit encore difficilement envisageable même si l'ingestion d'eau lourde (D_2O , The Merck Index, 1978, Composé 2896) n'est considérée comme toxique pour les organismes dits « supérieurs » qu'en grosse quantité (Kushner et al. 1999). Il est donc difficilement envisageable d'isoler des γ cristallines totalement marquées d'organisme eucaryote de type mammifère ou de lignée cellulaire mammifère.

Afin d'obtenir de grandes quantités de protéines recombinantes, des systèmes de surproduction bactériens ou lévariens sont largement utilisés au sein de nombreux laboratoires tels que le laboratoire de deuteration de l'ILL. La production à grande échelle de protéine recombinante est facilitée par l'utilisation de fermenteurs. Malheureusement la substitution isotopique de protéines en bactérie ou en levure reste encore une technique difficile et non courante, du fait des rendements de culture beaucoup moins importants en D_2O qu'en H_2O . Il ne faut pas non plus oublier le problème financier qu'engendre l'utilisation de produits deutériés (voir tableau 3.1). Une culture de 2l en D_2O peut alors s'avérer coûter jusqu'à plus de 2000 euros.

Composés deutériés	Prix
D_2O	~ 350 euros/l
Glucose deutérié $C_6D_{12}O_6$	~ 570 euros/g
Glycérol deutérié $C_3D_8O_3$	~ 110 euros/g
Méthanol CD_4O	~ 25 euros/ml
Acide succinique deutérié $C_4D_6O_4$	~ 60 euros/g

Tableau 3.1 : Comparatif des prix des produits deutériés pouvant être utilisés en fermentation (d'après le catalogue Cambridge Isotope Laboratories).

En effet, chaque composé deutérié possède un coût relativement élevé. Pour minimiser les coûts, il est possible d'utiliser de plus petit volume, ainsi que la technique du " fed-batch " dans un fermenteur pour obtenir des densités cellulaires importantes. Ainsi, pour palier à ces problèmes et pour obtenir des productions de biomasse élevée, toute l'équipe du laboratoire de deuteration de l'ILL contribue aux développements de techniques de marquage (de protéines ou d'acides nucléiques) en utilisant principalement les cultures de haute densité cellulaire (HCDC, « High Cell Density Culture ») en fermenteur dans lesquelles, il est possible de contrôler de nombreux paramètres de culture.

3.1.1 Notions essentielles pour les cultures cellulaires

Que ce soit en flasque, en fermentation de type " batch ", ou en fermentation de type " fed- batch ", il est capital de connaître quelques rudiments sur les cultures cellulaires afin de produire une biomasse importante bien que les nombreux processus physiologiques impliqués ne soient pas encore bien décrits et définis.

3.1.1.1 Apports nutritifs et conditions de culture

Le catabolisme et l'anabolisme sont les voies obligatoires à la synthèse de petits éléments qui seront utilisés pour la génération des macromolécules (Pirt 1975; Enfors et Häggström 2000). Ces petits éléments sont des acides nucléiques, des acides aminés, des acides gras et des sucres, qui seront ensuite utilisés et assemblés en constituants macro moléculaires de la cellule tels que l'ADN, l'ARN, les protéines ou encore les membranes lipidiques.

Les différents composants nécessaires à une bonne croissance cellulaire sont une source de carbone (glucose, glycérol, acide succinique ou méthanol), une source d'azote (sous ammoniacale NH_3 ou apportée par d'autres composés), du soufre, du phosphore, des éléments en plus petite quantité (HPO_4^- , Cl^- , SO_4^- , K^+ , Na^+ , Ca^{++} , Mg^{++} , Fe^+ , Mn^{++} , Co^{++}), de l'eau (légère ou lourde), de l'oxygène et des vitamines. Tous ces composés seront donc présents au cours de la fermentation. Etant donné que le but de la perdeutériation est d'effectuer un marquage spécifique au niveau des hydrogènes des protéines, il n'est pas possible d'utiliser des extraits bactériens ou de levures, commercialisés, ne pouvant être totalement deutériés (trop coûteux). Le milieu doit être défini et minimum.

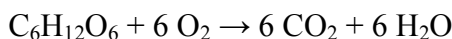
3.1.1.1.1 La source de carbone

Le carbone est l'un des éléments les plus abondants de la bactérie. Le CO₂ est la seule source de carbone pour les bactéries autotrophes. Les bactéries hétérotrophes dégradent une grande quantité de substances hydrocarbonées (alcool, acide acétique, acide lactique, polysaccharides et sucres divers). Il est important de se rendre compte que le carbone apporté servira à la fois de source pour la production de nouvelles macromolécules mais aussi pour l'apport énergétique de la cellule. Généralement, la quantité de carbone est le facteur limitant principal dans une culture cellulaire.

Le glucose est couramment utilisé comme source de carbone dans les milieux de culture définis (Riesenberg et al. 1991; Vanatalu et al. 1993...; Hochuli et al. 2000). Mais le glucose est un composé relativement onéreux quand il est sous forme deutérié. Mais, il est possible d'utiliser le glycérol, 5 fois moins cher, pour remplacer la source carbonique. Celui-ci est intégré au niveau de la glycolyse par transformation en D-glycéraldéhyde puis en D-glycéraldéhyde-3 Phosphate par respectivement une glycérol hydrogénase et une triokinase (voir annexe A).

3.1.1.1.2 Cas de l'oxygène

Dans un métabolisme aérobie, l'oxygène joue le rôle de dernier accepteur d'électron ou d'hydrogène. Ce processus est réalisé par une enzyme oxydase. Derrière son rôle d'accepteur d'électron, l'oxygène peut, grâce à l'oxydase, être incorporé dans le catabolisme des substrats carbonés. Une estimation de la quantité totale d'oxygène requise pour le métabolisme énergétique du métabolisme énergétique peut être obtenue à partir de la stœchiométrie de l'oxydation. Par exemple, pour l'oxydation d'une molécule de glucose durant la croissance d'*E. coli*, il est possible d'écrire :



Quand l'apport en oxygène est limité, le substrat carboné est oxydé partiellement, ce qui conduit à la sécrétion de sous-produits (acétate, formate, éthanol). La sécrétion de ces sous-produits métaboliques permet la formation d'énergie avec un changement de métabolisme cellulaire (Riesenberg et al. 1991; Varma et al. 1993). Ce changement de métabolisme n'est pas bon pour la continuité de la culture cellulaire. Durant une culture

cellulaire, le taux d'oxygène consommé par les cellules est en équilibre avec l'oxygène apporté au système. Mathématiquement, cet équilibre peut être décrit par :

$$qO_2 \cdot X_v = k \cdot C \quad (3.1)$$

où qO_2 est la consommation d'oxygène par cellule, X_v est la densité cellulaire viable, C est la concentration d'oxygène dissout et k est le coefficient de transfert de masse qui est caractéristique de l'équipement utilisé. Ce transfert de masse dépend de la taille des bulles, du débit et de l'agitation. A l'ILL, les fermenteurs sont munis d'agitateurs mécaniques ou magnétiques et de cannes permettant un afflux gazeux d'air comprimé ou d'un mélange d'oxygène.

Ainsi pour optimiser la production d'énergie lors d'une culture, la concentration en oxygène est un paramètre important qu'il faut surveiller avec précision. Pour éviter que les bactéries poussent dans une phase anaérobique, le taux d'oxygène dissout est maintenu à 30 %. Ceci est effectué par des mesures en continu d'oxygène dissout dans le milieu.

3.1.1.1.3 Effet de la température

La température affecte la cinétique des réactions cellulaires, la nature du métabolisme, l'apport en nutriments et donc la composition de la biomasse. La plupart des bactéries peuvent vivre et se multiplier dans un intervalle de température de 30 °C. La place de cet intervalle de température sur le thermomètre permet de classer les bactéries en trois groupes : les psychrophiles ou cryophiles qui préfèrent des basses températures (de 0 °C à 25 °C), les mésophiles, au milieu (de 10 °C à 45 °C) et les thermophiles (de 40 °C à 90 °C). Mais l'intervalle de température pour une croissance optimale est plus étroit et est défini par la loi d'Arrhénius :

$$\mu_{\max} = A \times e^{\frac{-\Delta E}{RT}} \quad (3.2)$$

où μ est le taux de croissance, A est une constante dépendante de l'organisme, R est la constante des gaz parfaits, T est la température et ΔE est l'énergie d'activation.

Au cours d'une culture, la température se situe généralement aux abords d'un intervalle d'Arrhénius, où la croissance est à son optimum, mais ce paramètre dépend aussi des conditions et du devenir de la culture. En effet, une culture en fermenteur, où le taux de croissance doit être contrôlé, n'aura pas forcément la même température qu'une culture en flasque. De même une culture utilisée afin de produire de la biomasse, ne sera pas tempérée de la même manière si celle ci produit des protéines recombinantes plus ou moins solubles.

De plus, la température joue un rôle très important au niveau de la solubilité de l'oxygène. Plus la température est réduite, plus la solubilité de l'oxygène dans l'eau est augmentée. Donc, à des températures plus faibles, l'apport en oxygène est plus facile pour la culture. Il faut donc faire un compromis entre la solubilité de l'oxygène, qui doit basse, et la température optimum de croissance cellulaire, qui se situe vers 37°C pour *E. coli*.

3.1.1.1.4 Le pH

Des bactéries sont trouvées dans des conditions extrêmes de pH (surtout les Archaeobactéries), de pH 1,0 (sources sulfuriques acides) à pH 11,0 (lacs alcalins). *Escherichia coli*, pousse à des pH qui tendent vers la neutralité, quoique légèrement alcalins. Par exemple, elle pousse sans problème entre pH 6,0 et pH 8,0 et peut encore pousser, quoique difficilement, à des valeurs plus extrêmes telles pH 5,0 et pH 9,0. Ces valeurs de pH reflètent, bien sûr, les pH extérieurs de la cellule. Les variations de pH cytoplasmique, qui se situe aux alentours de 6,8 (Stryer 1988), sont beaucoup plus étroites que celles du pH extérieur.

3.1.1.2 Paramètres de la croissance

Dans un premier temps, les paramètres de croissance d'une culture vont être définis à partir d'une simple culture en batch où les conditions requises pour obtenir de la biomasse sont : un *inoculum* viable, une source d'énergie, des nutriments qui fournissent le matériel essentiel à la formation de biomasse, l'absence d'inhibiteur de croissance, et des conditions physico-chimiques idéales. A partir de là, dans une culture cellulaire il est possible d'observer une variation de biomasse (dx) pendant un intervalle de temps (dt) et qui est proportionnelle à la biomasse x présente au temps t soit :

$$dx = dt \cdot \mu x \quad \text{ou} \quad \mu x = \frac{dx}{dt} \quad (3.3)$$

Le coefficient différentiel (dx / dt) exprime le taux de croissance de la population. Le paramètre μ qui représente le taux de croissance par unité de biomasse, est appelé taux de croissance spécifique. Si celui-ci est constant, l'intégration de l'équation précédente devient alors :

$$\ln x = \ln x_0 + \mu \times t \quad \text{ou} \quad \ln \left(\frac{x}{x_0} \right) = \mu \times t$$

soit :

$$x = x_0 \times e^{\mu \times t} \quad (3.4)$$

La croissance qui est en accord avec cette loi sera appelée croissance exponentielle constante. Cette loi ne reste valable que si les conditions de culture restent constantes. Et la relation entre le taux de croissance spécifique et le temps Γ que met la biomasse pour doubler est obtenue en fixant $x = 2x_0$ et $t = \Gamma$ soit :

$$\Gamma = \frac{\ln 2}{\mu} \quad (3.5)$$

Le rendement Y d'une culture sera défini par le quotient suivant :

$$Y = \frac{\Delta x}{\Delta s} \quad (3.6)$$

où Δx est la variation de la biomasse consécutive à l'utilisation d'une quantité de substrat Δs . Plus précisément, le rendement est exprimé par la limite de $\Delta x / \Delta s$ quand Δs tend vers 0 et ce qui donne :

$$Y = \frac{dx}{ds} \quad (3.7)$$

Dans ce cas, étant donné que si x et s sont les concentrations en biomasse et en substrat respectivement, il est nécessaire d'introduire un signe négatif car x et s vont varier en sens contraire, donc le rendement devient :

$$Y = -\frac{dx}{ds} \quad (3.8)$$

Le maximum de densité de biomasse, et donc de rendement, qui peut être obtenu avec un milieu donné est déterminé par une de ces cinq conditions suivantes :

- épuisement de la quantité du substrat limitant ;
- limitation de facteurs de croissance (exemple : O_2)
- accumulation trop élevée d'un produit inhibiteur ;
- niveau de densité cellulaire maximum atteint ;
- lyse cellulaire avec libération de nombreux produits.

Pour estimer la biomasse d'une culture, plusieurs méthodes sont utilisables mais celles-ci sont toujours basées sur l'un des principes de mesure suivants : la masse de la culture, la masse d'un des composé de la biomasse (exemple contenu en azote de la cellule), la masse d'un produit formé, la capacité métabolique (production de gaz par exemple), le comptage de cellule après étalement reproductible sur boîte gélosée et les méthodes colorimétriques. Le choix de la méthode pour mesurer la biomasse est crucial pour avoir un contrôle de la culture et les facteurs qui influencent ce choix sont : les propriétés de la biomasse étudiée (la taille des cellules peut variée), les propriétés du milieu de culture

(certains milieux absorbent à 600 nm), la précision requise, et la rapidité de la mesure. Au laboratoire de deutériation de l'ILL, les méthodes les plus couramment utilisées sont : la mesure de masse mouillée, la mesure de masse sèche, la mesure de densité optique à 600 nm et la mesure de turbidité par une sonde type Wedgewood.

3.1.1.3 Les phases de croissance : cas d'une culture en « batch »

Dans une culture en batch, tous les substrats sont accessibles en très grande quantité pour donner un taux de réactions non réprimé (en respect avec les substrats). La concentration de la biomasse augmente de manière exponentielle et le taux de croissance spécifique est à son maximum. Ceci est visible jusqu'à ce que des facteurs, tels que des inhibiteurs ou une faible concentration en substrat, réduisent le taux de croissance. Seuls les gaz comme l'oxygène, les agents contrôlant le pH et l'anti-mousse sont ajoutés lors d'une culture en batch. A la différence, lors d'une culture en flasque, le pH et l'oxygénation ne sont pas contrôlés. Au cours de cette culture, il est possible de discerner six différentes phases qui reflètent des changements dans la biomasse et dans l'environnement des cellules (voir figure 3.1).

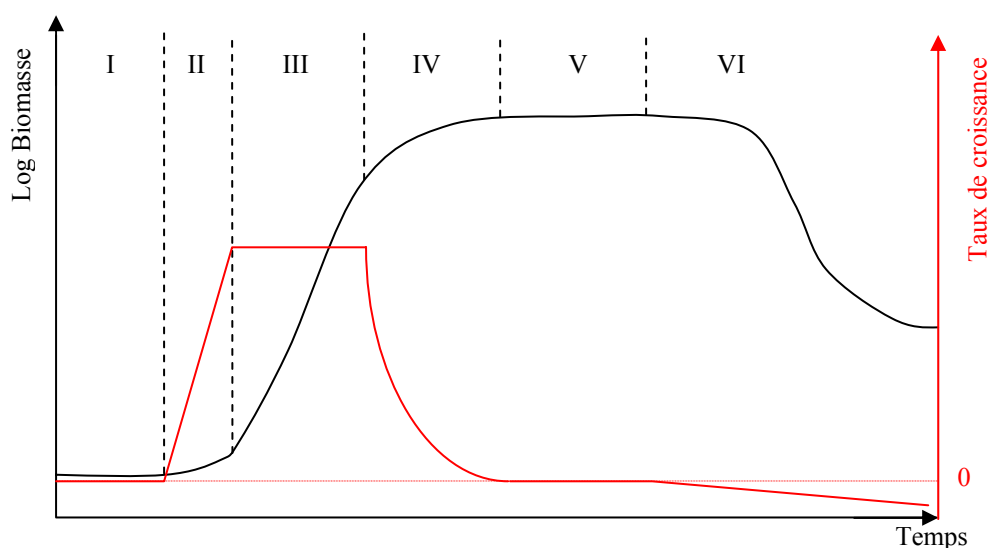


Figure 3.1 : Courbe de croissance d'une culture en batch avec 6 phases : I, la phase de latence ; II, la phase d'accélération ; III, la phase exponentielle ; IV, la phase de décélération ; V, la phase stationnaire ; et VI la phase de déclin. Le taux de croissance spécifique est représenté en rouge.

Une culture de haute densité en batch est toujours initiée par l'addition d'un *inoculum* à t_0 (~10% du volume de culture). La phase de latence I correspond à un temps pendant lequel le taux de croissance spécifique est nul ($\mu = 0$). La durée de cette phase dépend de l'âge des bactéries et de la composition du milieu. C'est le temps nécessaire à la bactérie

pour synthétiser les enzymes adaptées au nouveau substrat. En fermenteur la phase de latence est quasi nulle étant donné que l'*inoculum* a été réalisé avec le même milieu. Pendant la phase d'accélération II, le taux de croissance spécifique est en augmentation. Ce dernier atteint un maximum ($\mu = \mu_{\max}$) au cours de la phase de croissance exponentielle III. Cette phase dure tant que la vitesse de croissance est constante. Le temps de doublement des bactéries est le plus court.

La masse cellulaire est représentée par des cellules viables (mortalité nulle). Puis vient la phase de ralentissement IV où la vitesse de croissance régresse. Il y a un épuisement du milieu de culture et une accumulation des déchets. Il existe un début d'autolyse des bactéries. Pendant la phase stationnaire V, le taux de croissance spécifique devient nul ($\mu = 0$). Les bactéries qui se multiplient compensent celles qui meurent. Enfin, la phase de déclin VI : le taux de croissance est négatif ($\mu < 0$). Toutes les ressources nutritives sont épuisées. Il y a accumulation de métabolites toxiques. Il se produit une diminution d'organismes viables et une lyse cellulaire sous l'action des enzymes protéolytiques endogènes. Cependant, il persiste une faible croissance par libération de substances libérées lors de la lyse (croissance cryptique).

3.1.1.4 La culture en fed batch

La culture en fed batch (voir figure 3.2) prend l'avantage d'une culture poussant en phase exponentielle (voir précédemment) tout en minimisant la fabrication de sous produits toxiques en limitant l'apport continu d'un substrat.

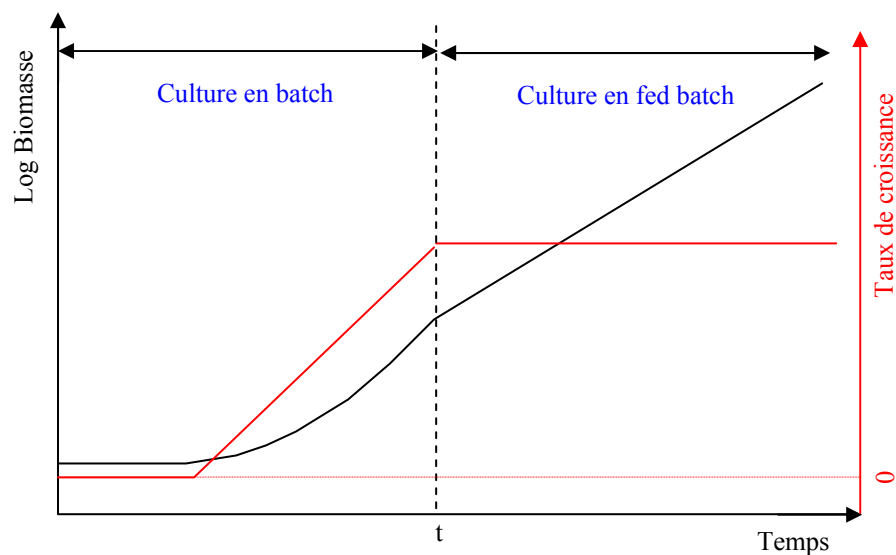


Figure 3.2 : Courbe de croissance d'une culture en fed batch. L'apport en source carbonique au bout du temps t permet de maintenir un certain taux spécifique de croissance.

Ainsi l'ajout contrôlé de ce substrat permet de restreindre le taux de croissance. La plupart du temps, le substrat ajouté est la source de carbone du milieu défini. Celui-ci sera de très haute concentration (jusqu'à 30-50% w/v) pour réduire l'augmentation de volume de culture (et donc la dilution des cellules) et le coût de la culture (dans le cas d'une culture en D₂O). La concentration en substrat et le taux de croissance spécifique de la culture vont atteindre un état quasi stationnaire dont les niveaux seront essentiellement déterminés par l'apport en substrat. Cet apport F peut être calculé de deux manières différentes par les équations 3.9 et 3.10:

$$F = \frac{\mu_{obj} * V * OD_{600}}{Y * 0,4 * C} \quad (3.9) \quad \text{ou} \quad F = F_0 * 2 \exp^{\frac{t}{\Gamma * 1,5}} \quad (3.10)$$

où μ_{obj} est le taux de croissance spécifique à atteindre (en heure⁻¹), V représente le volume de la culture (en litre), OD_{600} est la densité optique à 600 nm, Y correspond au rendement de masse cellulaire par unité de source de carbone utilisée et C est le pourcentage de source carbonique, le glycérol par exemple, dans la solution d'apport (en g/ml). $0,4$ est le facteur de conversion de la densité optique à la masse cellulaire. Pour l'équation 3.10, F_0 est un coefficient d'apport en glycérol à déterminer suivant la souche et les conditions de fermentation. t est le temps passé au cours de la fermentation et Γ est le temps de doublement de la culture. Le facteur 1,5 est fixé de façon à travailler en condition sub-optimale afin d'éviter une accumulation de substrat dans le milieu.

3.1.1.5 La production de protéine recombinante

Au cours de cette thèse, plusieurs systèmes d'expression ont été étudiés et sont toujours en cours de développement. Le plus étudié et utilisé est un système procaryote bactérien comprenant la souche d'*Escherichia coli* BL 21 (DE3). La production de protéine recombinante dans la bactérie *Escherichia coli* est certainement le système d'expression le plus répandu dans le monde. Pour l'expression de protéine, par exemple une γ cristalline, l'ADN complémentaire (ADNc) est cloné dans un vecteur d'expression : un plasmide pET28a dans notre étude (Novagen) qui est transformé dans des cellules d'*E. coli*. Ce plasmide contient un gène de résistance à la kanamycine pour la sélection des cellules et une origine de répllication colE1. Le gène cible est placé sous contrôle d'un promoteur pour la polymérase T7 ARN, et est exprimé avec ou sans une queue poly-histidine en N-terminal pouvant être coupée par la thrombine. La souche *Escherichia coli* BL 21 (DE3) possède un gène DE3

codant pour la polymérase T7 ARN qui est sous contrôle d'un opéron lactose (ou analogue tel que l'IPTG).

3.1.2 La culture en milieu deutérié

Parce que les effets de l'isotope deutérium sont connus pour altérer la division cellulaire et les fonctions enzymatiques, la croissance cellulaire et l'expression de protéines dans un environnement deutérié ne sont pas des techniques courantes. De précédents travaux ont montré, cependant, qu'il était possible d'obtenir, par la méthode de « fed batch », des cultures deutériées d'*E. coli* (Vanatalu et al. 1993; Junemann et al. 1996) à des densités optiques raisonnables avec des milieux entièrement deutériés pour minimiser l'incorporation des hydrogènes. Pour acquérir de la biomasse, il est important de bien choisir le milieu deutérié. Un milieu riche, avec des extraits d'algues deutériées ou du glucose deutérié comme source de carbone, est utilisable mais à des coûts très prohibitifs. Sans adaptation à ce nouvel environnement et avec un milieu minimum, contenant de l'acide succinique ou du glycérol deutériés (qui sont moins coûteux voir tableau 3.1), la croissance bactérienne va être souvent ralentie, et les rendements de protéines recombinantes vont être moindres (Paliy et al. 2003) (voir tableau 3.2).

Souche bactérienne	Γ en LB en minutes	Γ en MM hydrogéné en minutes	Γ en MM deutérié en minutes
JM109	31	65	201
MRE600	-	57	129
JM109 _{adaptée}	-	-	136
MRE600 _{adaptée}	-	-	96

Tableau 3.2 : Temps de doublement Γ moyen de culture de différentes souches *E. coli*. MM signifie milieu minimum. Les bactéries sont capables de croître en milieu deutérié mais avec des taux de croissance spécifiques différents. Une souche non adaptée va pousser moins vite qu'une souche adaptée au milieu deutérié (Paliy et al. 2003).

Pour favoriser la substitution isotopique, il est donc impératif de rendre les souches plus tolérantes à la croissance en D₂O, par des séries d'adaptation. Les protocoles classiques pour l'adaptation d'*Escherichia coli* (Mann et Moses 1971; Venters et al. 1995) incluent des passages de milieux hydrogénés à des milieux deutériés, en augmentant progressivement la concentration en eau lourde dans le milieu de culture.

3.2 Matériels et méthodes

3.2.1 Clones de γ cristallines disponibles

Pour l'expression de γ E cristallines de rat (*Rattus norvegicus*), l'ADNc a été cloné dans le vecteur d'expression pET28a de manière à avoir une expression de la protéine native sans et avec une queue poly-histidines en N-terminal (voir annexe B.1). Pour le pET28 γ ENco, la queue poly-histidines a été supprimée du pET28a par digestion complète par les enzymes de digestion NcoI et NotI. Cette partie est remplacée par la séquence codant pour les γ cristallines avec le côté 5' par digestion complète du vecteur par NdeI et NotI et une insertion des séquences codantes pour les γ cristallines entre ces deux sites (le côté 5' au niveau de NdeI).

L'intégration de l'ADN plasmidique dans la cellule est une transformation chimique : 1 ng d'ADN plasmidique est additionné à 100 μ l de cellules compétentes BL21(DE3) et puis incubé sur la glace pendant 20 minutes. Après absorption du plasmide à travers les membranes cellulaires, les cellules sont placées à 45°C pendant 45 secondes, puis 2 minutes sur la glace. 900 μ l de milieu SOC (voir annexe C.1) est ajouté et les cellules sont incubées pendant 1 heure à 37°C sous agitation. Pour sélectionner les cellules qui ont été correctement transformées, la culture est centrifugée (10' à 4000 rpm) puis le culot est repris par 50 μ l de milieu SOC et la suspension est étalée sur une boîte contenant de la kanamycine (30 μ g/ml). La boîte est incubée à 37°C pendant une nuit. Après avoir vérifié l'expression des γ cristallines en milieu riche (milieu Luria-Bertani + kanamycine, voir annexe C), sur gel SDS PAGE, les cellules sont congelées sur cryobilles à -80°C (Coligan et al. 2003).

3.2.2 Adaptation au milieu minimum deutérié

Les différentes souches d'*E. coli* contenant les plasmides, préalablement congelées, ont été étalées sur une boîte gélosée (LB + 30 μ g/ml kanamycine) *via* une cryobille, possédant sur sa surface la souche en question, et incubées à 37°C pendant une nuit. Les colonies sont étalées de nouveau sur une boîte gélosée de milieu minimum (milieu Enfors hydrogéné + 30 μ g/ml kanamycine ; voir annexe C.2 pour composition) et incubées à 37°C. La dernière opération est répétée mais cette fois sur milieu Enfors deutérié solide (voir annexe C.2 pour composition et protocole de préparation). Les colonies qui pousseront sur ce milieu vont

servir à inoculer 10 ml de milieu Enfors deutérié dans une flasque fermée hermétiquement de façon à réduire les échanges avec l'eau gazeuse environnante. Cette flasque va incuber à 30°C sous agitation (180 rpm). Cette culture va servir à inoculer 10 ml de milieu Enfors deutérié, qui seront incubés comme précédemment. Cette opération sera répétée, au total, 4 fois. Enfin, la dernière culture sera, soit congelée et conservée à - 80 °C, soit servira d'inoculum à un plus grand volume (150 ml) en vue d'effectuer une fermentation. Les différentes étapes d'adaptation d'*E. coli* au D₂O sont résumées en figure 3.3 :

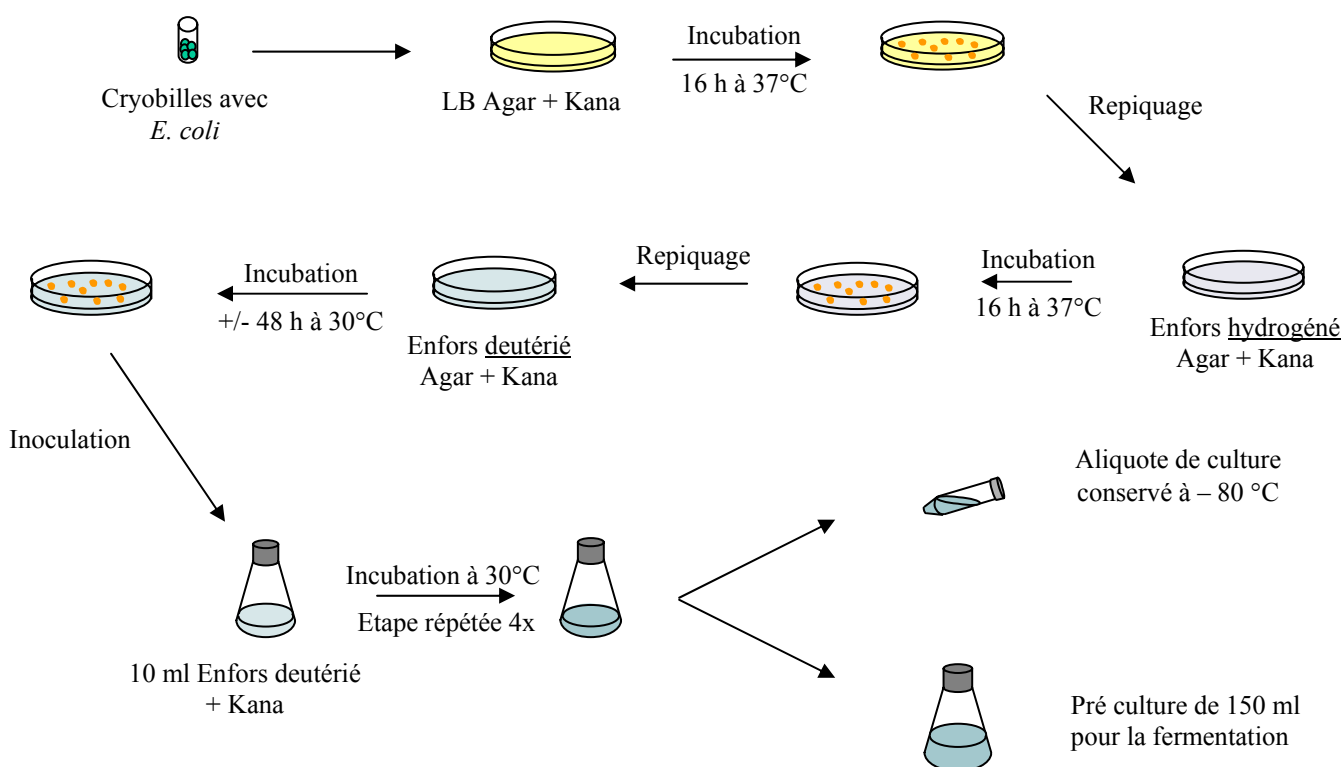


Figure 3.3 : Représentation schématique de l'adaptation d'*Escherichia coli* à la croissance en D₂O. Celle-ci est effectuée d'abord par une adaptation au milieu minimum sous forme hydrogénée, puis sous forme deutériée. L'adaptation est complétée en milieu liquide par la suite.

3.2.3 Expression en fermenteur

Pour une culture hydrogénée, la souche désirée est étalée sur une boîte gélosée (LB + 30 µg/ml kanamycine) via une cryobille contenant le clone désiré et incubée à 37°C pendant une nuit. Les colonies sont repiquées sur une boîte gélosée de milieu minimum (milieu Enfors hydrogéné + 30 µg/ml kanamycine) et incubées à 37°C. En flasque, 150 ml de milieu Enfors hydrogéné seront inoculés et incubés à 30°C pour servir de pré-culture au fermenteur. De même, pour une fermentation deutériée, un inoculum de 150 ml est réalisé à partir de la

souche déjà adaptée. Ces inoculum représentent 10 % du volume initial de la culture en fermenteur. Dans l'objectif d'obtenir de grandes quantités cellulaires, un fermenteur de type Labfors II (Infors) avec une capacité de 3,6 l (volume maximal de culture = 2,5 l ; voir figure 3.4) est utilisé avec 1,5 l de milieu Enfors + kanamycine (30 µg/ml). Ce type de fermenteur est un système entièrement géré par ordinateur et par le programme informatique IRIS (Infors). Les différents paramètres de fermentation sont mesurés, enregistrés et visualisés sur l'écran de l'ordinateur.

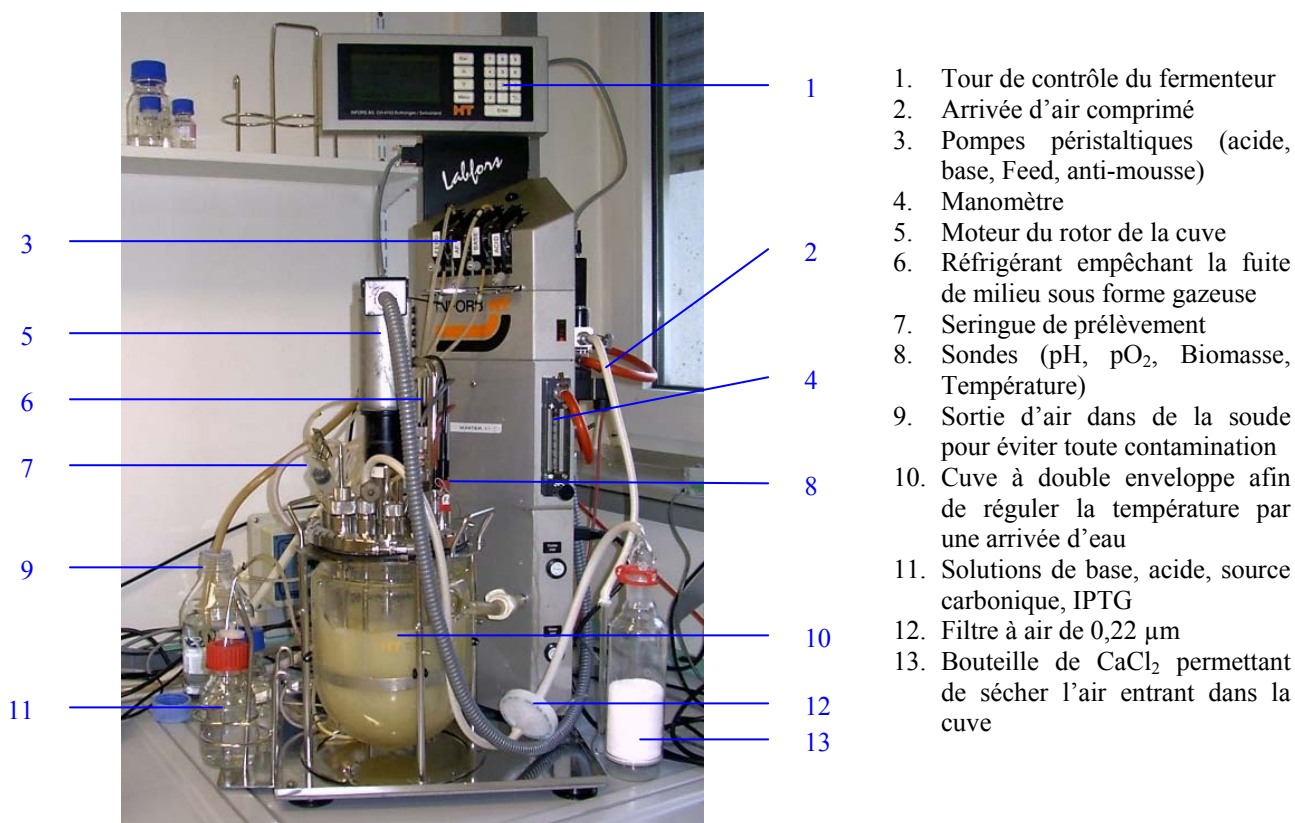


Figure 3.4 : Photographie d'une unité de fermentation du Labfors II. Le fermenteur est composé d'une tour enregistrant toutes les valeurs qui seront ensuite analysées par l'ordinateur.

Le temps que la préculture soit à prêtre, le fermenteur est rempli de façon stérile par 1,35 l de milieu Enfors, le pH est ajusté à 6,9-7 par ajout de soude et le milieu est porté à 30°C. Quand la préculture a atteint une densité optique de 0,6-0,8, le fermenteur est inoculé. Le pH de la préculture ne doit pas être trop acide (<6) car il va révéler la présence de cellules en phase stationnaire. Le pH, la quantité d'oxygène dissout par l'agitation et l'apport d'air comprimé, la température, la biomasse sont contrôlés tout au cours de la fermentation. Le déclenchement de l'apport en source carbonique peut être effectué de deux façons différentes : soit des indications physiologiques (changement de pH, augmentation de la pO₂) informent du moment important ; soit ce dernier est déterminé au cours du temps par les

précédentes fermentations, quand celles ci sont reproductibles (densité optique et pH de la préculture). L'apport en source carbonique est effectué par une solution contenant le précédent milieu mais avec une quantité plus importante de source carbonique (voir annexe C.2). Cet apport est calculé et contrôlé par l'équation 3.9. Quand la DO_{600} atteint 10-12, la synthèse de γ cristallines est induite par l'addition d'isopryl- β -D-thiogalactopyranoside (IPTG) à des concentrations finales de 0,5 mM à 1 mM, soit par ajout direct, soit par un ajout lent et continu. La fermentation continue au moins pendant 1 jour avec un apport en source carbonique plus faible et constant. Après vérification sur gel SDS PAGE de la production de γ cristallines au cours de la fermentation, la biomasse cellulaire est récoltée et centrifugée à 5000 rpm pendant 30 minutes. Le culot est par la suite repris par une solution Hepes 5mM pH 6,4 afin d'éliminer la majorité des sels de la culture. La suspension est à nouveau centrifugée (5000 rpm 30 minutes) et les cellules sont conservées à - 80°C après congélation par de l'azote liquide.

3.3 Résultats et discussions

3.3.1 Choix du plasmide

Dans les cultures de hautes densités cellulaires, l'instabilité plasmidique est fréquente sous certaines conditions (Park et al. 1990). En effet, les bactéries ont tendance à perdre leur plasmide quand il n'y a pas de pression de sélection. L'instabilité plasmidique peut être affectée par de nombreux facteurs incluant les caractéristiques génétiques des bactéries hôtes, le nombre de copies du plasmide, sa forme, le taux de mutation, les conditions de culture (Tierny et al. 1999). Un autre facteur important doit être considéré : la différence des taux de croissance des cellules ayant et n'ayant pas de plasmides. Généralement, l'existence d'un ADN extra chromosomique, tel qu'un plasmide multi-copies, à l'intérieur d'une cellule est gênant pour les cellules hôtes et ceci est encore plus véridique lors du stress produit par l'expression d'une protéine recombinante en grande quantité. Ainsi, sans pression de sélection, le taux de croissance de cellules libres de plasmide est plus important que celui de cellules en ayant un. Par conséquent, lors d'une culture avec ces deux types cellulaires, la sous population ayant un plasmide va décroître et l'autre va devenir dominante. En fermentation, lors de l'utilisation de plasmide conférant la résistance à l'ampicilline, il se produit un phénomène similaire. En fait, le gène de résistance à l'ampicilline code pour une protéine appelée la β

lactamase qui inactive l'ampicilline. Chez *E. coli*, la β -lactamase inactive l'ampicilline dans l'espace périplasmique des cellules. En culture haute densité, cette accumulation extracellulaire de β -lactamase risque de rendre inefficace l'ampicilline présente dans le milieu et donc favoriser involontairement la croissance des cellules n'ayant pas de plasmide. La résistance à la kanamycine est différente : l'enzyme cytoplasmique produite, l'aminoglycoside phosphotransférase, inactive, par phosphorylation, la kanamycine. La résistance au chloramphénicol est aussi cytoplasmique : le chloramphénicol va être inactivé par acétylation par l'acétyltransférase produite à l'intérieur de la cellule. Dans ces deux cas, la pression de sélection est maintenue, et les cellules, contenant le plasmide avec le gène d'intérêt, vont devenir la population ultra dominante de la culture.

Le choix des vecteurs comportant les résistances à la kanamycine ou au chloramphénicol est donc évident mais il est toujours nécessaire d'ajouter de l'antibiotique pendant la fermentation pour que celui ci reste toujours à une concentration constante.

3.3.2 Adaptation des bactéries à la croissance en D₂O

La souche d'*Escherichia coli*, MRE 600, est une souche bactérienne dont la croissance est la moins affectée par le D₂O, d'après Paliy et al (2003). Malheureusement cette souche n'est pas la plus adéquate pour la sur-expression de protéines. Le système T7 de la souche BL21 (DE3), dépourvue de protéases *lon* et *ompT*, est sûrement le plus couramment utilisé pour l'expression de protéines recombinantes. Ainsi, cette souche a été sélectionnée pour exprimer des protéines dans des environnements hydrogénés et deutériés.

En milieu deutérié et sans adaptation, la croissance de cette souche est extrêmement ralentie et, peut être dans certains cas, totalement nulle (mort cellulaire). Les protocoles classiques d'adaptation d'*E. coli* à un milieu totalement deutérié sont généralement une suite de cultures liquides dans lesquelles la concentration de D₂O est augmentée progressivement jusqu'à un pourcentage total (Mann et Moses 1971; Vanatalu et al. 1993; Venters et al. 1995). Ces adaptations ont été, en fait, décrites comme pouvant être le résultat de mutations spécifiques (De Giovanni 1960; Butler et Grist 1984) ou d'un processus d'ajustement métabolique par rapport au deutérium (Orgel 1964; Flaumenhaft et al. 1965). Au cours de notre étude, le premier étalement sur un milieu minimum solide et hydrogéné est considéré comme une accommodation du métabolisme bactérien aux substrats nouvellement disponibles. Mais le second étalement sur milieu minimum solide, cette fois-ci deutérié, peut

correspondre à la sélection de ces « mutants » ou colonies utilisant plus aisément le deutérium. Une étude protéomique et génomique est en cours de réalisation au laboratoire de deutériation pour déterminer exactement quels sont les changements intervenant face à ce type de stress. Mais il est certain que notre protocole sur milieu solide permet une sélection de colonies pouvant croître plus efficacement en milieu deutérié. L'adaptation à ce milieu et à cet environnement stressant est poursuivie par la suite en milieu liquide pour que les taux de croissance soient du même ordre en milieu hydrogéné et deutérié.

3.3.3 Culture en mode « batch fed batch »

Une culture de type « batch fed batch » est représentée en figure 3.5. Pendant ce type de culture, il est important de contrôler certains paramètres. De ce fait, le pH, la température, la quantité d'oxygène dissout peuvent être examinés et ajustés automatiquement au cours de la fermentation. Pour maintenir le pH à une certaine valeur (valeur optimum de croissance 6,8 pour *E. coli*), des ajouts de solutions titrantes sont effectués. Dans ce cas précis, la solution compensant l'acidification progressive du milieu est de la soude 4 %. Les différentes observations ont montré que l'addition de base est révélatrice de l'état physiologique de l'ensemble de cellules. Les ajouts répétés sont souvent signe d'un métabolisme fort conduisant à une augmentation de la biomasse : les cellules sont dans un bon état physiologique. Un ralentissement de ces additions démontre une diminution du métabolisme avec un taux de croissance spécifique moins élevé. La température est régulée par injection d'eau plus ou moins tempérée dans la double enveloppe du fermenteur. Elle n'est pas fixée à l'optimum de croissance d'*E. coli* de manière à pouvoir plus facilement contrôler le taux de croissance et donc éviter un « emballement » de la machinerie cellulaire. Une plus faible température permet aussi une plus grande solubilité de l'oxygène dans l'eau. La quantité d'oxygène dissout (pO_2) est maintenue à 30 % de saturation par un ajout constant d'air comprimé (non représenté) ou par une agitation variable du milieu suivant les besoins. Les fluctuations de la pO_2 peuvent en partie s'expliquer par le fragile équilibre qu'il existe entre l'apport en oxygène et la consommation de ce dernier par les cellules. Il est très important de veiller à ne pas avoir un manque d'oxygénation de la culture pour ne pas passer en anaérobiose.

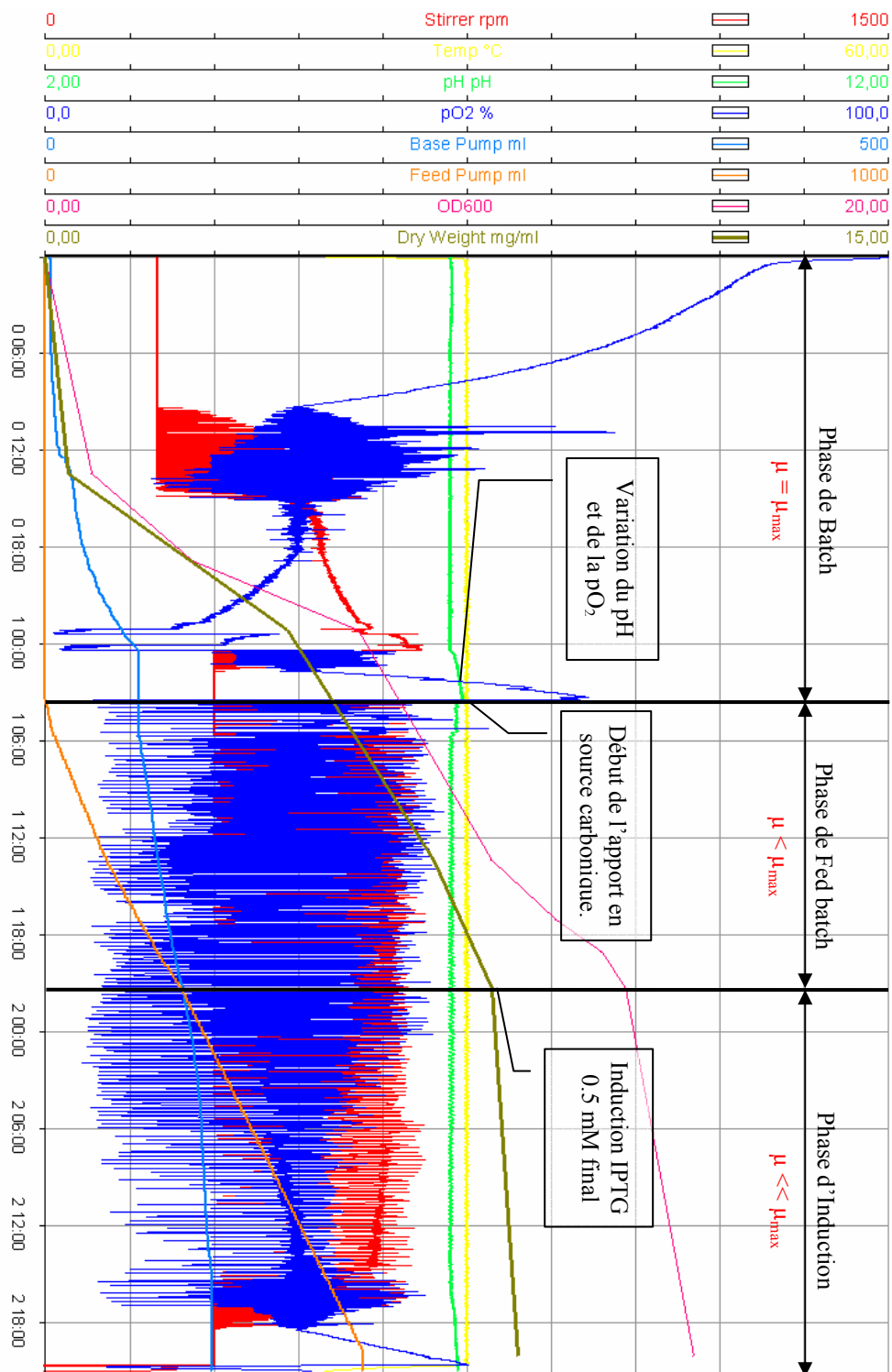


Figure 3.5 : Profil typique de fermentation d'*Escherichia coli* BL21 (DE3) en milieu minimum Enfors deutérié avec ses paramètres classiques. La fermentation est décomposée en 3 phases : une phase de batch où μ est à son maximum, une phase de fed batch où l'apport en glycérol limite le taux spécifique de croissance μ et une phase où μ est extrêmement réduit.

D'autres paramètres indirects, révélant l'évolution de la biomasse, peuvent être visualisés : la densité optique à 600 nm, la masse mouillée ou sèche de la culture. Cette fermentation de type "batch fed batch" peut se décomposer en trois phases :

3.3.3.1 La phase de « batch »

Juste après l'inoculation va suivre une courte phase de latence pendant laquelle les cellules vont commencer à consommer l'oxygène présent dans le milieu (la pO_2 en bleu va diminuer). L'agitation du milieu (en rouge) va augmenter et équilibrer cette utilisation. Le milieu va tendre à s'acidifier mais l'ajout de base (en cyan) va le stabiliser à pH 6,8 (en vert). Il est remarquable de voir que cet ajout va se faire de façon quasi exponentielle. Les cellules utilisent la source carbonique présente dans le milieu et vont atteindre un taux de croissance spécifique μ maximum. Pendant cette phase, la croissance, et donc la formation de sous-produits toxiques, ne sont pas sous contrôle. En revanche, la formation de biomasse suivie par la DO_{600} (en rose) et la masse sèche (en marron), suit une courbe exponentielle. A un certain moment (~ 1 jour et 3 heures), le pH et la pO_2 vont augmenter indiquant un changement de métabolisme des cellules qui n'ont plus assez de glycérol disponible. Les signes d'alcalinisation du milieu ne sont pas toujours présents mais ils sont révélateurs de la nouvelle utilisation de sous-produits tel que l'acétate ou le lactate. Or il est essentiel que la culture ne passe pas dans une phase stationnaire, et il est donc préférable que l'apport en source carbonique se fasse pendant la phase exponentielle de la culture. Quand les cultures s'effectuent de manière reproductible (même inoculum, même pH, mêmes conditions de fermentation), ce moment fatidique est aisément déterminé.

3.3.3.2 La phase de « fed batch »

Le glycérol est ajouté à la culture suivant l'équation 3.9, de façon à limiter le taux de croissance et donc la synthèse des sous produits toxiques. De cette manière, μ est sous contrôle, il est constant mais est inférieur à son maximum. Cet apport est sub-optimal pour éviter toute accumulation de substrat et de sous produits toxiques dans le milieu. La biomasse augmentera, donc, de façon linéaire. L'ajout de base suit la même restriction : son ajout devient linéaire.

3.3.3.3 La phase d'induction

L'expression de la protéine recombinante va se faire par addition d'IPTG à des concentrations variant de 0.5 mM à 1 mM. Cette quantité d'IPTG varie essentiellement en fonction de la densité cellulaire de la culture et du taux d'expression de la protéine. Pendant cette période, l'apport en source carbonique est continu, mais l'augmentation de la biomasse va diminuer voire dans certains cas être nulle. Dans le cas des γ cristallines, la production n'est pas toxique, elle n'entraîne pas de lyse cellulaire. Anderson *et al.* (1996) estiment que la production de protéines recombinantes ne demande pas des besoins énergétiques considérables, diminuant ainsi le taux de croissance. Mais, ils pensent que l'utilisation d'IPTG conduirait à la production de certaines protéines de réponse à un stress (Andersson *et al.* 1996).

D'un point de vue plus pratique, il arrive, dans certains cas, que la synthèse chimique de l'IPTG utilise des solvants nocifs et toxiques pour les cellules bactériennes (communication personnelle Jacques Winter, INFORS France). L'IPTG aurait donc un « effet létal » et Andersson et ses collaborateurs conseillent l'utilisation d'analogue de l'IPTG tel que le lactose ou une addition plus lente de l'IPTG. L'utilisation d'homologues deutériés n'est pas envisageable dans notre étude (analogue deutérié trop coûteux), bien que dans certains cas comme chez Cserjan-Pushmann et ses collaborateurs, le rendement Y est 2,1 fois plus grand lors d'une induction au lactose.

Par une addition plus fractionnée de l'IPTG, il est possible d'observer encore une augmentation de la biomasse pendant cette phase. Il est arrivé que l'expression soit plus importante (Cserjan-Puschmann *et al.* 2002) du fait d'une régulation plus fine du taux d'expression du gène recombinant et du maintien de la productivité par le système pendant une grande période. L'induction est effectuée donc pendant environ 24 heures par de petites injections d'IPTG. La sur-expression de γ cristallines est visualisable par une bande aux environs des 21 kDa sur gel SDS (voir figure 3.6).

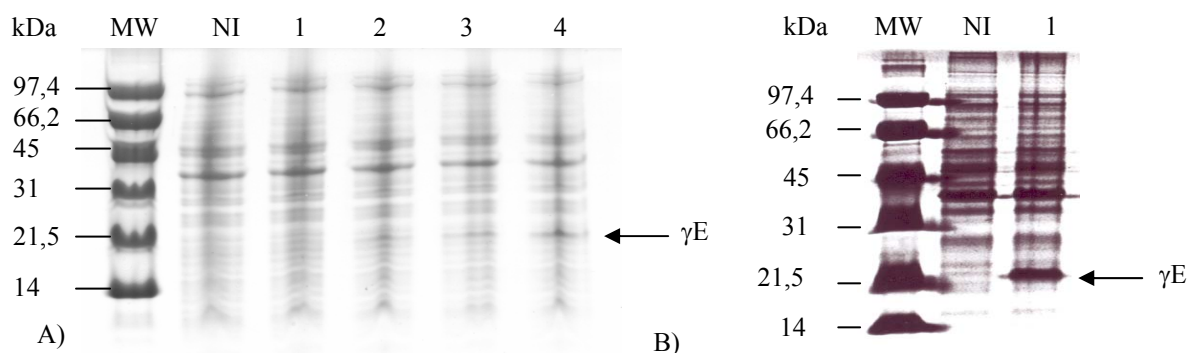


Figure 3.6 : Expression de la protéine recombinante hydrogénée γE cristalline de rat dans *E. coli. BL21 (DE3)*. Gels SDS PAGE Tris Tricine 12% des protéines totales extraites à partir des cellules transformées avec le pET28 $\gamma ENco$ codant pour la γE cristalline. A) Echantillons prélevés pendant 4 heures après l'induction avec 0.5 mM IPTG. B) Echantillon prélevé après 30 heures pour la même induction. Les protéines cellulaires totales sont analysées sans (NI) et avec traitement par 0,5 mM IPTG (A1-4 et B1). Le temps d'induction des échantillons est de 1, 2, 3, 4 heures (A2-4) et de 30 heures (B1). Les protéines sont visualisées par coloration au Coomassie.

Cette bande est la protéine majoritaire dans l'extrait total cellulaire. L'induction, pendant un si long temps (~ 30 heures), ne montre pas de dégradation importante de la protéine (bandes éventuelles de poids moléculaire plus faible) par le renouvellement protéique à l'intérieur même de la cellule. Cette observation, peu fiable, aurait pu être vérifiée uniquement à l'aide de western blot. Malheureusement aucun anticorps contre la γ cristalline n'était disponible.

3.3.4 Comparaisons cultures hydrogénées et deutériées

Les cultures hydrogénées et deutériées d'*E. coli BL21 (DE3)* ont été effectuées par la méthode de "batch fed batch". Il est relativement difficile de comparer toutes les cultures tellement certains paramètres peuvent être critiques pour les rendements et les taux de croissance. Il est aussi important de signaler que toute mesure de densité optique est une grande source d'erreurs (façons de prélever, de diluer), mais pour comparer deux cultures, ces erreurs sont normalement répétées dans tous les cas. Ainsi, il est quand même possible de dégager certaines observations de ces cultures. La figure 3.7 montre les courbes de croissance (Densité Optique à 600 nm) d'une fermentation de *BL21 (DE3)* dans 0 et 100 % D₂O dans le milieu Enfors.

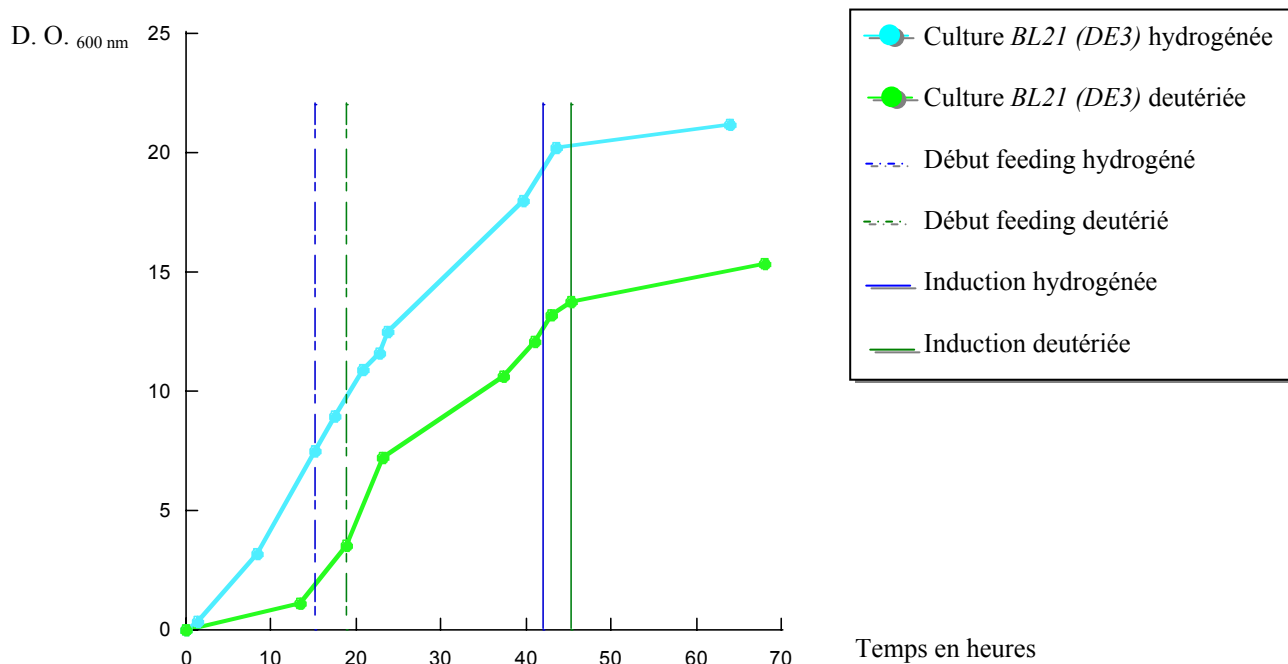


Figure 3.7 : Courbes de croissance d'*E. coli* BL21 (DE3) cultivées en batch, puis en fed batch et sous influence d'IPTG (0.5mM) dans un milieu Enfors hydrogéné et deutérié.

Les cultures présentent un profil de culture relativement similaire avec l'apparition des trois phases précédemment discutées. Au niveau de la culture en batch, là où le taux de croissance n'est pas contrôlé, la croissance en milieu deutérié est plus lente qu'en milieu hydrogéné car celle-ci présente une phase de latence plus longue (voir tableau 3.3). Cette différence est certainement due à un inoculum plus dense en milieu hydrogéné.

Paramètres	Culture hydrogénée			Culture deutériée		
Temps de fermentation en h	~ 64			~ 68		
Phase	I	II	III	I	II	III
μ en h^{-1}	0,62	0,34	0,05	0,44	0,24	0,07
Γ en h	1,12	2,04	~14	1,57	2,88	~10
Volume final de la culture en l	1,940			1,745		
Quantité de glycérol utilisée en g	70			43		
Biomasse mouillée en g	131			79,6		
Rendement Y	1,87			1,85		

Tableau 3.3 : Comparaison des résultats de fermentation d'une culture d'*E. coli* hydrogénée et d'une deutériée. La culture hydrogénée est une culture exécutée en routine où ce type de rendement est fréquent. La culture deutériée est la culture au meilleur rendement obtenu sur ce sujet. Les deux cultures présentent des taux de croissance spécifiques du même ordre.

La similitude entre les deux cultures est encore visualisable pendant la seconde phase, qui est contrôlée par l'utilisateur décidant de la quantité en source carbonique à apporter. Le "feeding" est commencé plus tardivement dans la culture deutériée du fait que celle-ci s'est retrouvée en phase exponentielle plus tardivement, ceci est toujours probablement dû à un inoculum moins dense.

Le ralentissement des deux cultures, dû à l'induction par l'IPTG, est visible aussi, et les taux spécifiques de croissance sont du même ordre. Mais après de nombreuses adaptations, il apparaît que les rendements d'une culture en milieu entièrement deutérié peuvent être similaires à une culture en milieu hydrogéné.

3.4 Bilans

Après avoir réalisé le clonage dans un vecteur d'expression comportant le gène de résistance à la kanamycine, le plasmide résultant a été transformé dans une souche *Escherichia coli*. Cette souche a permis l'expression des γ cristallines en milieu hydrogéné et deutérié. Après avoir développé une technique d'adaptation originale à la croissance en D₂O, cette expression a été effectuée en fermenteur afin de minimiser les coûts et produire de plus grandes quantités de protéines. Ayant produit de grandes quantités de pâte cellulaire bactérienne contenant nos protéines, nous nous sommes interrogés sur le niveau de deutérioration des protéines et les différences biochimiques et biophysiques qu'il pourrait exister entre les cristallines hydrogénées et les formes perdeutériées. Pour cela, nous avons purifié les protéines suivant des protocoles identiques, décrits dans le chapitre suivant. Par la suite, nous avons entrepris une étude de caractérisation par différentes techniques : la spectrométrie de masse, la détermination de l'état natif et du point isoélectrique, la digestion protéolytique, le dichroïsme circulaire et une étude infrarouge.

Au sein du laboratoire de deutérioration de l'ILL, de nombreuses études sur le marquage isotopiques sont réalisées, en particulier sur la perdeuration des protéines. Ces nombreuses études permettent d'établir des protocoles nécessaires à la production de protéines deutériées. Ces protocoles recourent à la fois la biologie cellulaire mais aussi moléculaire. Ainsi, la nécessité d'utiliser une pression de sélection qui agit de façon intracellulaire a permis la production de grande quantité de protéines. Le marquage isotopique par les deutériums est aussi favorisé par la technique, développée au cours de cette thèse, qu'est l'adaptation cellulaire au D₂O. Les γ cristallines étant des protéines eucaryotes, nous avons initié un travail

de production de protéine recombinante dans un système d'expression lévurien (*Pichia pastoris*) dans lequel l'expression de protéines mammifères semble être plus adéquate d'un point de vue repliement et modifications post-traductionnelles (essentiellement pour la glycosylation de protéine) (Clare et al. 1991; Cregg et al. 1993; Laroche et al. 1994). A ce jour, le clonage des gènes de γ C et γ E cristalline de rat, avec et sans queue poly-histidines, dans le vecteur, pPICZA, nécessaire pour la recombinaison homologue dans le génome de *P. pastoris* a été réalisé. L'intégration des gènes codant pour la γ E cristalline, dans le génome lévurien, est terminée. Mais l'expression en milieu minimum n'est pas très importante à cause d'un faible taux de recombinaison (pas de multi-recombinant) dans le génome et de la difficulté d'obtenir un milieu défini en eau lourde (beaucoup de précipitation). Cette étude n'ayant pas encore abouti à ce jour, les résultats ne seront pas présentés dans ce manuscrit.

Bibliographie :

- Andersson, L., Yang, S., Neubauer, P., and Enfors, S.O. 1996. Impact of plasmid presence and induction on cellular responses in fed batch cultures of *Escherichia coli*. *J Biotechnol* **46**: 255-263.
- Butler, L.O., and Grist, R.W. 1984. The effect of D2O on the growth and transforming activities of *Streptococcus pneumoniae*. *J Gen Microbiol* **130 (Pt 3)**: 483-494.
- Clare, J.J., Rayment, F.B., Ballantine, S.P., Sreekrishna, K., and Romanos, M.A. 1991. High-level expression of tetanus toxin fragment C in *Pichia pastoris* strains containing multiple tandem integrations of the gene. *Biotechnology (N Y)* **9**: 455-460.
- Coligan, J., Dunn, B., Speicher, D., and Wingfield. 2003. *Current protocol in protein sciences*, John Wiley and Sons, Inc ed.
- Cregg, J.M., Vedvick, T.S., and Raschke, W.C. 1993. Recent advances in the expression of foreign genes in *Pichia pastoris*. *Biotechnology (N Y)* **11**: 905-910.
- Cserjan-Puschmann, M., Grabherr, R., Striedner, G., Clementschitsch, F., and Bayer, K. 2002. Optimizing Recombinant Microbial Fermentation Processes
An Integrated Approach. *BioPharm July*: 26-34.
- De Giovanni, R. 1960. The effects of deuterium oxide on certain microorganisms. *Annals of the New York Academy of sciences* **84**: 644-647.
- Enfors, S.O., and Häggström, L. 2000. *Bioprocess Technology Fundamentals and Applications*, Hogskolestryckeriet, Royal Institute of Technology ed, Stockholm.
- Flaumenhaft, E., Bose, S., Crespi, H.L., and Katz, J.J. 1965. Deuterium isotope effects in cytology. *Int Rev Cytol* **18**: 313-361.
- Hochuli, M., Szyperski, T., and Wuthrich, K. 2000. Deuterium isotope effects on the central carbon metabolism of *Escherichia coli* cells grown on a D2O-containing minimal medium. *J Biomol NMR* **17**: 33-42.
- Junemann, R., Wadzack, J., Triana-Alonso, F.J., Bittner, J.U., Caillet, J., Meinnel, T., Vanatalu, K., and Nierhaus, K.H. 1996. In vivo deuteration of transfer RNAs: overexpression and large-scale purification of deuterated specific tRNAs. *Nucleic Acids Res* **24**: 907-913.
- Kushner, D.J., Baker, A., and Dunstall, T.G. 1999. Pharmacological uses and perspectives of heavy water and denatured compounds. *Can J Physiol Pharmacol* **77**: 79-88.
- Laroche, Y., Storme, V., De Meutter, J., Messens, J., and Lauwereys, M. 1994. High-level secretion and very efficient isotopic labeling of tick anticoagulant peptide (TAP) expressed in the methylotrophic yeast, *Pichia pastoris*. *Biotechnology (N Y)* **12**: 1119-1124.
- Mann, L.R., and Moses, V. 1971. Properties of *Escherichia coli* grown in deuterated media. *Folia Microbiol (Praha)* **16**: 267-284.
- Orgel, L.E. 1964. Adaptation to Wide-Spread Disturbance of Enzyme Function. *J Mol Biol* **116**: 208-212.
- Paliy, O., Bloor, D., Brockwell, D., Gilbert, P., and Barber, J. 2003. Improved methods of cultivation and production of deuterated proteins from *E. coli* strains grown on fully deuterated minimal medium. *J Appl Microbiol* **94**: 580-586.
- Park, S., Ryu, D., and Lee, S. 1990. Determination of kinetic parameters related to the plasmid instability: for the recombinant fermentation under repressed condition. *Biotechnology and Bioengineering* **32**: 404-414.
- Pirt, J.S. 1975. *Principles of microbes and cell cultivation*. Halsted Press.
- Riesenberg, D., Schulz, V., Knorre, W.A., Pohl, H.D., Korz, D., Sanders, E.A., Ross, A., and Deckwer, W.D. 1991. High cell density cultivation of *Escherichia coli* at controlled specific growth rate. *J Biotechnol* **20**: 17-27.
- Stryer, L. 1988. *Biochemistry*. Freeman and Company, New York.
- Tierny, Y., Hounsa, C.G., and Hornez, J.P. 1999. Effects of a recombinant gene product and growth conditions on plasmid stability in pectinolytic *Escherichia coli* cells. *Microbios* **97**: 39-53.
- Vanatalu, K., Paalme, T., Vilu, R., Burkhardt, N., Junemann, R., May, R., Ruhl, M., Wadzack, J., and Nierhaus, K.H. 1993. Large-scale preparation of fully deuterated cell components. Ribosomes from *Escherichia coli* with high biological activity. *Eur J Biochem* **216**: 315-321.
- Varma, A., Boesch, B.W., and Palsson, B.O. 1993. Stoichiometric interpretation of *Escherichia coli* glucose catabolism under various oxygenation rates. *Appl Environ Microbiol* **59**: 2465-2473.
- Venters, R.A., Huang, C.C., Farmer, B.T., 2nd, Trolard, R., Spicer, L.D., and Fierke, C.A. 1995. High-level 2H/13C/15N labeling of proteins for NMR studies. *J Biomol NMR* **5**: 339-344.

Chapitre 4

Caractérisations biochimiques et spectrales

Abstract

γ E-crystallin was expressed, in *E. coli*, with and without a poly-histidine tag, an amino acid repeat motif facilitating the purification. With a better yield, large quantities of hydrogenated and deuterated γ E-crystallin, expressed without a poly-histidine tag, were purified. The mean of perdeuteration level of non-exchangeable sites in the protein was found to be 98% by electrospray mass spectrometry but the hydrogenated protein solution was heterogeneous: the post translational modification level (N-terminal methionine removal) was higher in the deuterated system. The Methionine Amino Peptidase (MAP) was therefore co-expressed with hydrogenated γ E-crystallin to recover the homogeneity of the protein solution.

In the absence of known biochemical activity, the hydrogenated and deuterated γ E crystallins were characterised by native gel (same behaviour against DTT), isoelectric point determination (same migration) and limited proteolysis (high resistance to protease digestion in both cases). There are no major biochemical differences between these two proteins.

Further characterisations, such as Fourier transform infrared (FTIR) and circular dichroism (CD) spectroscopies, were used to investigate minimal differences. In FTIR, the perdeuteration had an effect on infrared spectra with a shift of the amide I band to lower wavenumbers by $\sim 10 \text{ cm}^{-1}$ and of ring vibration of tyrosine by $\sim 90 \text{ cm}^{-1}$. However, the perdeuteration may not have a drastic consequence on the secondary structure composition. In CD, hydrogenated and deuterated proteins in H_2O and in D_2O buffers have a similar secondary structure. By CD, chemical and thermal unfolding studies showed no difference except a very small stabilisation of the deuterated protein. The midpoint transition temperatures, in H_2O , are 81.2 and 81.8°C for the hydrogenated and deuterated proteins, and in D_2O , 83.2 and 84.5°C respectively.

4.1 Caractérisations des protéines hydrogénées et deutériées

Les effets et l'utilisation du D₂O sur les protéines sont relativement bien documentés (Schowen et Schowen 1982), et la perdeutériation ne semble avoir que de faibles effets sur la structure et l'activité biologique des protéines. Mais cette substitution est quand même atypique puisque, pour une protéine deutériée telle qu'une γ cristalline, le poids moléculaire va augmenter de 5% par rapport à la protéine protonée. De plus, d'un niveau macroscopique, des études expérimentales (Benjamin et Benson 1962) et théoriques (Scheiner et Cuma 1996) ont démontré que dans l'eau, la liaison deutérium était plus forte de $\sim 0,1$ à $0,2 \text{ kcal mol}^{-1}$ par rapport à la liaison hydrogène. Cette augmentation de la force de liaison serait attribuée à une plus importante masse du deutéron abaissant, ainsi, l'énergie vibrationnelle de la liaison. Il serait donc, surprenant que les propriétés physico-chimiques d'une protéine perdeutériée soient similaires à celles de la protéine hydrogénée, sans pour autant entraîner des changements drastiques. Il est essentiel de savoir si cette substitution isotopique n'a que des répercussions minimales sur les cristallines. Il serait idéal si les deux systèmes hydrogénés et perdeutériés ne diffèrent pas de manière significative. Dans le cas contraire, notre étude cristallographique ne permettrait pas de comparaison appropriée avec les protéines natives du cristallin. Ainsi certaines propriétés physico-chimiques et biologiques des deux protéines vont être évaluées de différentes manières, dans ce chapitre.

En premier lieu, la purification des deux espèces hydrogénées et deutériées et leur caractérisation par spectrométrie de masse ont été effectuées. Les γ cristallines n'ayant pas d'activité biologique reconnue, une comparaison de protéolyse ménagée entre les deux formes a été réalisée pour vérifier si l'accessibilité des protéases était identique pour une protéine hydrogénée ou une protéine perdeutériée. Ensuite la caractérisation biochimique (gel natif et point isoélectrique) est poursuivie et a révélé le comportement migratoire identique des deux protéines. L'analyse moléculaire par spectroscopie infrarouge a montré certains effets de la perdeutériation. Enfin j'ai réalisé, en dichroïsme circulaire, une étude de stabilité par dénaturation chimique et thermique que ce soit dans un environnement hydrogéné ou deutérié.

4.2 Matériels et méthodes

4.2.1 Préparations des échantillons

4.2.1.1 La protéine γ cristalline exprimée avec queue poly-histidines

Pour chaque préparation de protéines, 10 g de patte cellulaire congelée ont été resuspendus dans 150 ml de tampon de lyse en présence d'antiprotéases (sans EDTA) (voir annexe D.1). La suspension a été filtrée sur buchner avec un filtre 0,9 μm pour éviter toute obstruction de l'appareil d'ouverture des membranes. Les cellules ont été lysées par désintégration par le CELL Disruptor (Constant Systems) à une pression de 1,3 kbar. Le matériel insoluble a été éliminé par centrifugation pendant 45 minutes à 15000 rpm à 15°C.

La partie soluble a été filtrée sur filtre millipore 0,45 μm puis chargée sur une colonne de Nickel (Ni-NTA Agarose de Qiagen) équilibrée par le tampon de lyse (0,75 ml/minute). Après lavage de la colonne par 10 volumes de tampon de lyse, la colonne a été lavée une seconde fois par 10 volumes de tampon de lyse supplémenté de 20 mM imidazole. Les protéines accrochées ont été ensuite éluées par du tampon de lyse contenant 250 mM Imidazole. Le tampon a été remplacé par une colonne PD-10 ou par dialyse par le tampon de digestion de la thrombine. La digestion a été effectuée pendant une nuit à 20°C. La capture de la thrombine biotinylée était faite suivant les recommandations du fournisseur. La purification a été poursuivie par une chromatographie d'exclusion (Superdex 200 de Amersham Biosciences), reliée au système ÄKTA prime d'Amersham Biosciences. Les fractions de 2 ml ont été analysées sur gel SDS PAGE et concentrées par centrifugation (Amicon Ultra-15, limite d'exclusion à 5 kDa). La concentration de la protéine a été mesurée en utilisant l'absorption optique à 280 nm avec un coefficient d'absorption molaire ξ de 2,063 (<http://www.expasy.org/tools/protparam.html>).

4.2.1.2 La protéine γ cristalline exprimée sans queue poly-histidines

La lyse bactérienne a été effectuée de la même façon que précédemment sauf qu'elle a été réalisée dans un tampon différent (voir annexe D.2).

La partie soluble, filtrée sur filtre millipore 0,45 μm puis a été chargée sur une colonne d'affinité échangeuse de cations (SP Sepharose de Amersham Biosciences). Le débit était de 0,75 ml/minute de façon à permettre une fixation optimale de la protéine à la résine.

Après lavage par du tampon de lyse (10 volumes), les protéines accrochées ont été éluées par un gradient de sels (0 à 0,3 M NaCl). Les fractions contenant la protéine γ cristalline ont été regroupées, et précipitées par du sulfate d'ammonium (70% de saturation). Le culot a été repris par 2 ml de tampon de lyse et la purification a été poursuivie sur colonne d'exclusion de 50 ml (Superdex 200 de Amersham Biosciences), reliée au système Biologic de BioRad. Elle a permis la séparation de molécules de poids moléculaire allant de 1.10^4 à 6.10^5 Da. 1 ml d'échantillon a été injecté sur la colonne à un débit de 0.8ml/minute. Des fractions de 2 ml ont été collectées, analysées sur gel SDS PAGE et concentrées jusqu'à un volume de 12 ml par centrifugation (Amicon Ultra-15, limite d'exclusion à 5 kDa). La solution protéique a été ensuite dialysée toute une nuit contre de l'eau dans une cassette de dialyse PIERCE (limite d'exclusion 7 kDa). Un traitement au DTT 5mM final a précédé la concentration finale de la protéine par centrifugation (Amicon Ultra-15, limite d'exclusion à 5 kDa).

4.2.1.3 Préparation de la protéine en milieu deutérié (D_2O)

Les protéines ont été purifiées et concentrées suivant les techniques précédemment décrites. Par la suite, les protéines ont été diluées dans l'eau lourde déionisée, dont le pD a été vérifié autour de la neutralité, puis à nouveau concentrées par centrifugation. Cette opération a été recommencée 4 à 5 fois pour obtenir une solution protéique dans un environnement complètement deutérié. Le potentiel de deutérium est défini par l'équation suivante :

$$pD = pH_{\text{mesuré}} + 0,3314n + 0,0766n^2 \quad (4.1)$$

où n est la fraction molaire du D_2O .

4.2.2 Spectrométrie de masse

En formant des ions en phase gazeuse à partir de molécules neutres, la spectrométrie de masse est une des techniques extrêmement efficace pour déterminer la masse moléculaire. Toutes les spectrométries de masse ont trois composants essentiels : une source ionique, un analyseur de masse et un détecteur. Les ions sont produits à partir de l'échantillon dans la source ionique par une technique spécifique (l'électrospray dans notre cas). Les ions sont séparés dans l'analyseur par le rapport de leur masse sur leur charge (m/z) et ils sont ensuite détectés par un amplificateur électronique. Les données produisent un spectre de masse où les ions sont représentés en fonction de leur abondance du rapport m/z .

Les poids moléculaires des protéines hydrogénées et deutériées ont été mesurés par spectrométrie de masse par électrospray à l'Institut de Biologie Structurale (Grenoble). Les protéines ont été diluées dans 0,1% TFA et dessalées sur cartouche C4 protein Macrotrap (Michrom). Puis, elles ont été éluées avec 60% of CH₃CN/H₂O 9/1 (v/v) 0.1% TFA. Les solutions micromolaires (2,5 pmol/μl) ont été injectées directement dans un spectromètre électrospray triple quadripôle (AP III+, Perkin Elmer -Sciex). L'étalonnage du spectromètre a été réalisé avec une solution de polypropylène glycol. Les spectres de masse ont été enregistrés dans une échelle de taux de masse sur charge (m/z) de 500 à 1500 par étape de 0,5. Le spectromètre de masse fonctionne avec une aiguille ionisante à haut voltage (5 kV) et un potentiel à l'orifice de 80V. Les données ont été traitées avec le programme Masspec 3.3 qui permet la reconstruction des profils de masse moléculaire à partir de la déconvolution algorithmique des spectres de masse m/z contenant les différents états de charge.

4.2.3 Clonage de la Méthionine Amino Peptidase (MAP) dans le pLysS, expression et purification de la γ E cristalline

Le clone pET28bMAP contenant la Méthionine Amino Peptidase *d'E. coli* mutée en l'acide aminé 175 (sans conséquence sur l'activité) a été gracieusement fourni par le professeur Brian Matthews, Institute of Molecular Biology, Eugene, USA. Afin de fournir une origine de répllication et un marqueur de sélection différents de ceux du système pET28, le plasmide pLysS (Novagen) va servir de vecteur hôte à la MAP. Le plasmide pLysS possède une origine de répllication p15A compatible avec celle trouvée dans les plasmides dérivés du pBR322 (exemple le pET28), isolée à partir de *E. coli*.

Le vecteur pET28bMAP a été digéré par les enzymes SmaI et FspI et a créer des fragments dont un, contenait la séquence de la MAP qui est placée sous contrôle d'un promoteur pour la polymérase T7 ARN. Le pLysS a été digéré par les enzymes EcoRV et ScaI et a fourni des fragments dont un grand contenant le gène de résistance au chloramphénicol et l'origine de répllication p15A. Ces deux fragments (celui contenant la MAP et celui contenant la résistance au chloramphénicol) ont été assemblés et transformés dans des cellules Novablue (Novagen). Après sélection sur chloramphénicol, l'ADN plasmidique des clones positifs a été transformé par la méthode Hanahan (Sambrook et Russell 1998) dans la souche BL21 (DE3) contenant déjà le pET28 γ ENco.

L'expression a été vérifiée en milieu LB et en milieu minimum. Elle a été réalisée en fermenteur suivant le même protocole que celui décrit dans le chapitre 3. La purification des γ cristallines a été effectuée de la même façon que précédemment.

4.2.4 Gel natif

L'électrophorèse native ou non-dénaturante, c'est à dire en absence de produits dénaturants tels que l'urée ou des détergents, est une technique souvent utilisée pour la détermination du poids moléculaire d'une protéine, du nombre d'isoformes de cette protéine, et de la présence de protéines multimériques. Du fait que la mobilité dépend de la taille, de la forme, et de la charge inhérente à la protéine, les électrophorèses en condition non dénaturante fournissent un ensemble de paramètres de séparation différents de la taille (SDS PAGE) et de la charge (IEF).

Les électrophorèses en condition non-dénaturante ont été réalisées sur des échantillons en présence et absence de traitement réducteur au DTT pendant 20 minutes. Elles ont été réalisées, à température ambiante mais avec un système de refroidissement pendant 3 heures, sur des gels natifs 6 % pré-équilibré (Coligan et al. 2003). Les protéines ont été, ensuite, visualisées par coloration au bleu de Coomassie.

4.2.5 Détermination du point isoélectrique

La séparation de protéines suivant leur charge est effectuée par électrophorèse. La migration est effectuée dans un gradient de pH ; chaque molécule migre jusqu'à l'endroit où le pH est égal à son pHi. Un gel de forte porosité (polyacrylamide) est utilisé pour que la taille n'influence pas la migration. Le gradient de pH est généré par des ampholytes, molécules amphotères de synthèse introduites dans le gel.

La détermination du point isoélectrique des formes hydrogénées et deutériées a été réalisée avec l'appareil PhastSystem (Amersham Biosciences) sur des gels IEF dont le pH varie de 3 à 9 suivant les recommandations du manufacturier. La détection des protéines se fait par coloration au Coomassie par le kit PhastGel Blue R (Amersham Biosciences).

4.2.6 Protéolyse ménagée des γ cristallines

De manière à visualiser si l'accessibilité des enzymes protéolytiques est similaire pour les protéines hydrogénées et deutériées, des digestions enzymatiques de γ cristallines ont été suivies au cours du temps en conditions natives et dénaturantes. Les protéines utilisées pour former des fragments homologues de γ cristallines, sont des enzymes plus ou moins spécifiques à des séquences peptidiques. Les caractéristiques de chacune d'elles et les conditions utilisées sont résumées dans le tableau 4.1. Les ratios de γ cristallines par rapport à la protéase sont de 20/1 (sauf pour la pepsine 150/1), les temps d'incubation varient de 0 à 20 h. La réaction a été stoppée par ajout de bleu de charge chaud et chauffage pendant 5 minutes à 95°C. Les fragments générés ont été visualisés sur gel SDS-PAGE coloré par bleu de Coomassie.

Protéine	Trypsine	Subtilysine	Elastase	Pepsine
Spécificité*	P ₁ = Arg > Lys	P ₁ = acides aminés acides ou neutre	P ₁ ou P _{1'} = Ile > Val > Ala, Gly, Ser (et autres acides aminés neutres non aromatique)	P ₁ ou P _{1'} = Phe >> Leu >> Trp > Ala > autres acides aminés hydrophobes
Tampon	0,1 M NH ₄ HCO ₃	0,1 M NH ₄ HCO ₃	0,2 M Tris pH = 7	0,05 M C ₆ H ₈ O ₇
Conditions	pH = 7, T = 37°C	pH = 7, T = 37°C	pH = 7, T = 37°C	pH=2, 1 M Urée, T = 37°C

Tableau 4.1: Spécificités, tampons et conditions de protéolyses ménagées des enzymes Trypsine, Subtilysine, Elastase et Pepsine. * Nomenclature de coupure : P₂---- P₁----↓----P₁----P₂, ↓ marque le site de coupure.

4.2.7 Spectroscopie infrarouge par transformée de Fourier

L'eau lourde dont le pD a été vérifié va servir de solvant aux γ cristallines hydrogénées et perdeutériées. Il a été complètement échangé par plusieurs cycles de dilutions et de concentrations, comme il a été décrit dans le paragraphe 4.2.1.3. Les protéines ont été préparées jusqu'à une concentration avoisinant 12 mg/ml. Les échantillons ont été placés, à température ambiante, entre deux fenêtres de CaF₂ (Spectra-Tech. Waltham, MA) séparées par un espaceur de 1000 μ m en Téflon. Une centaine d'interférogrammes pour chaque échantillon, ont été enregistrés à température ambiante (~23°C), sur un spectromètre 615

Jasco (Great Dunmow, U.K.) au C.E.A. de Grenoble (Equipe de Vincent Forge). La résolution spectrale (largeur de fenêtre spectrale utilisée) est de 4cm^{-1} . Les transformées de Fourier des 100 interférogrammes moyennés de chaque échantillon et du solvant sont calculées et corrigées par apodization. L'apodization consiste à amener doucement l'interférogramme à zéro à ses extrémités pour éviter des artefacts dans la transformée de Fourier. Le spectre du solvant a été obtenu à partir du tampon lors du dernier échange H₂O/D₂O des protéines. L'absorption de vapeur d'eau résiduelle a été soustraite aux différents spectres. Le lissage des courbes a été effectué par le programme *Jasco* qui consiste à recalculer un interférogramme par transformée de Fourier inverse du spectre. Dans cet interférogramme recalculé, les hautes fréquences ont été enlevées par une méthode similaire à l'apodization. Ces hautes fréquences correspondent à du bruit, et par conséquent le spectre obtenu après nouvelle transformée de Fourier a été lissé. La décomposition des spectres a été effectuée en employant le logiciel d'ajustage de courbe fourni par le manufacturier Jasco.

4.2.8 Dichroïsme circulaire

Les protéines sont faites de résidus qui sont des molécules chirales et donc ont une activité optique. Quand une solution de protéines est illuminée par de la lumière polarisée, celle-ci est capable de changer la direction de polarisation de la lumière. De ce fait, l'intensité de la rotation dépend à la fois, de la longueur d'onde du rayon incident, et de l'orientation relative des acides aminés. Le spectre de dichroïsme circulaire va donc varier avec les éléments de structure secondaire présents de la protéine. Nous avons utilisé le dichroïsme circulaire de façon à visualiser les changements de structure ou de stabilité dus à la perdeutériation ou à la modification du solvant. Pour ce faire, les conditions chimiques et les températures de demi-dénaturation des différentes formes de γ cristallines en différents milieux ont été étudiées.

Les 200 μl échantillons à 0,5 mg/ml final ont été incubés à température ambiante. Pour enregistrer le spectre CD, une cuvette en quartz de 0,1 cm contenant la solution protéique a été placée sur le porte-échantillon d'un spectropolarimètre Jasco-810. Le balayage de longueur d'onde a été effectué de 180 à 260 nm à une vitesse de 50 nm/minute. Le ratio signal sur bruit a été augmenté en moyennant 10 spectres. Le programme utilise une somme fabriquée de modèles idéaux d'hélice α , de brin β , de coudes β et de pelotes statistiques qui va varier suivant le contenu de chaque élément de structure secondaire pour obtenir la

méthode des moindres carrés d'un spectre de dichroïsme. L'ellipticité résiduelle moyenne, θ , a été calculée en ($\text{deg.cm}^2.\text{dmol}^{-1}$) suivant l'équation :

$$[\theta] = \frac{\theta_{obs} * mrw}{10 * d * c} \quad (4.2)$$

où θ_{obs} est l'absorption dichroïque observée, mrw est le poids moyen d'un résidu (110 Dalton), d est la longueur du chemin optique en centimètre, et c est la concentration in mg/ml.

Une étude de dénaturation chimique avec des concentrations croissantes, de 0 à 6 M, de GuH(D)Cl, a été effectuée. L'incubation durait au minimum 20 minutes mais n'excédait pas 1 heure. Le Guanidinium DCl a été préparé par évaporation flash et la poudre lyophilisée est reprise dans de l'eau lourde.

Le signal CD a été enregistré, 120 fois, à un minimum de 217 nm, longueur d'onde significative des feuillets β , composant majoritaire de structure secondaire des γ cristallines. La perte de structure secondaire est suivie et normalisée par l'équation suivante :

$$\% \text{ de perte de structure II} = \frac{(\theta_{6M} - \theta_{0M}) - |(\theta_{6M} - \theta_{obs})|}{(\theta_{6M} - \theta_{0M})} * 100 \quad (4.3)$$

où θ_{6M} et θ_{0M} sont respectivement l'ellipticité à 6 M et à 0M GuH(D)Cl. θ_{obs} est l'ellipticité à la concentration en GuH(D)Cl désirée.

Des études de dénaturation thermique sur les mêmes échantillons ont été réalisées. Le signal CD a été enregistré à 217 nm pendant que la température augmentait de 20 à 96°C (20°C/heure). La perte de structure est visualisée par l'équation suivante :

$$\% \text{ de perte de structure II} = \frac{(\theta_{96^\circ\text{C}} - \theta_{20^\circ\text{C}}) - |(\theta_{96^\circ\text{C}} - \theta_{obs})|}{(\theta_{96^\circ\text{C}} - \theta_{20^\circ\text{C}})} * 100 \quad (4.4)$$

où $\theta_{96^\circ\text{C}}$ et $\theta_{20^\circ\text{C}}$ sont respectivement l'ellipticité à 96°C et à 20°C. θ_{obs} est l'ellipticité à la température observée.

4.3 Résultats et discussions

4.3.1 Lyse cellulaire

La première étape de purification a consisté en une lyse cellulaire. Après différents essais avec la sonication dans différents tampons, il s'est avéré que l'utilisation de cette technique entraînait l'insolubilité d'une grande partie des γ cristallines (environ 50%). Des essais ont été pratiqués avec la « French Press », les rendements étaient plus importants mais la perte de protéines d'intérêt était moindre voir nulle avec la méthode de lyse par broyage « One Shot » avec le désingrateur de cellules (CONSTANT SYTEMS). Cet appareil fait preuve de rapidité, de facilité d'utilisation, d'un broyage efficace et surtout d'un refroidissement efficace de l'échantillon d'où la diminution des pertes en protéine (Hallez 1997). Il est fort possible que la chaleur produite à la fois par les différentes impulsions du sonicateur, ou par la forte pression créée par la « french Press », soit le dénominateur commun à l'origine de l'agrégation des cristallines.

4.3.2 Purification de la protéine avec une queue poly-histidines

L'expression de protéine recombinante portant une queue poly-histidines peut être un avantage pour une purification. L'utilisation de chromatographie d'affinité sur colonne Nickel (Ni-NTA) est une bonne première séparation des cristallines avec les protéines d'*E. coli*. La protéine était absente ou en très faible quantité dans l'effluent et les lavages et, son élution était immédiate avec de l'imidazole (figure 4.1).

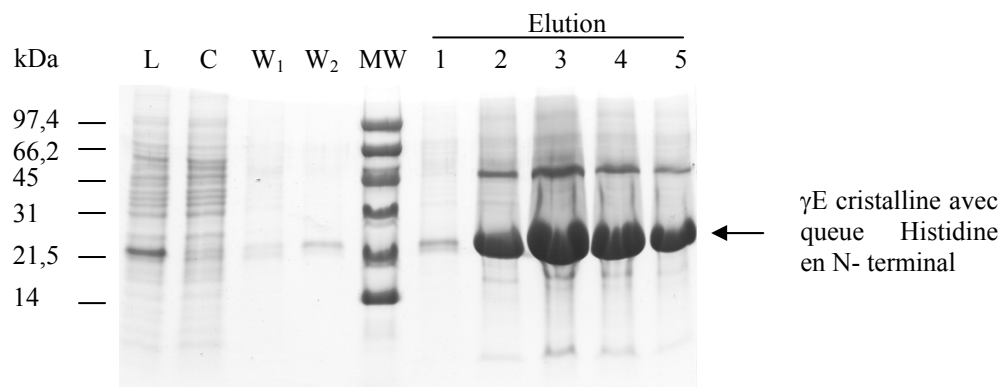


Figure 4.1 : Première étape de purification de la protéine γ E cristalline avec une queue poly-histidines. Chromatographie d'affinité sur colonne Nickel. La protéine en sortie de colonne présente déjà un bon niveau de pureté mais quelques impuretés résident encore dont une protéine dont le poids moléculaire est le double de la cristalline. L : Partie soluble du lysat ; C, effluent de la colonne Nickel, W₁₋₂, Lavages, MW, Marqueur de poids moléculaire, 1-5 : fractions d'élution

Lors de cette étape, une bande correspondant à un poids moléculaire deux fois plus important a été éluée en même temps que la cristalline. Il est possible qu'il y ait eu formation d'un dimère covalent en absence de DTT.

Le changement de tampon par dialyse ou par colonne de dessalage (PD10) pour se retrouver dans les conditions de digestion de la queue par la thrombine ne semblait pas convenir à la protéine. En effet, la variation de pH à travers son point isoélectrique théorique ($pI \approx 7,2$) entraînait un phénomène d'agrégation et donc de perte de protéine. Malgré cela, il reste assez de protéine pour effectuer la coupure de la queue histidine (voir figure 4.2).

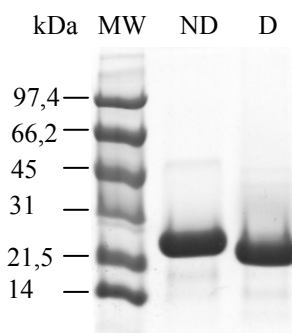


Figure 4.2 : Digestion de la queue poly-histidine par la thrombine. MW : Marqueur de poids moléculaire, ND : γ E cristalline non digérée, γ E cristalline digérée.

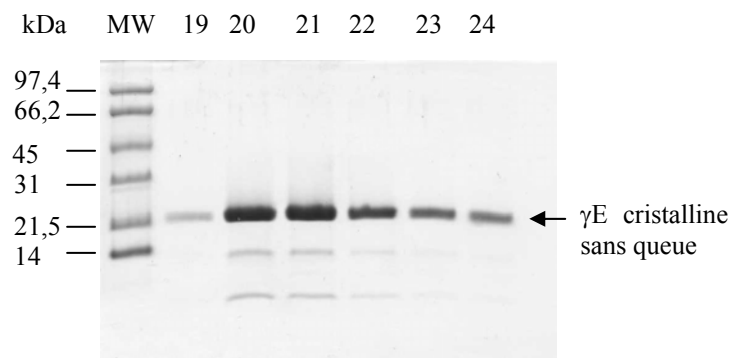


Figure 4.3 : Gel SDS PAGE de l'élué de la chromatographie d'exclusion S200. MW : Marqueur de poids moléculaire, 19-24 : fractions.

L'étape finale de purification, une chromatographie d'exclusion a permis de vérifier l'homogénéité de la protéine. La protéine γ E cristalline préparée de cette façon était pure comme le montre le gel d'élué des fractions de la colonne S200 (voir figure 4.3). Le « dimère covalent » précédemment observé à la sortie de la colonne Nickel, est désormais séparé de la cristalline. Un phénomène d'agrégation est aussi visible lors de la concentration des fractions d'intérêt, de ce fait le rendement de cette purification a été réduit. Lors de la concentration, les 2 bandes, de poids inférieur, visualisables sur le gel 4.3, n'étaient plus présentes.

A cause de la perte de protéine par précipitation lors du changement de tampon, j'ai décidé de purifier les γ cristallines en évitant tout changement de pH lors de la purification. Pour ce faire, j'ai envisagé une purification à partir de la protéine sans queue poly-histidines.

4.3.3 Purification de γ cristallines sans queue poly-histidine

Les protéines γ cristallines hydrogénées ou deutériées ont été purifiées sur colonne d'affinité échangeuse de cations, dans un tampon hydrogéné. Après l'expression et la lyse cellulaire, la protéine était soluble comme le montre la figure suivante. Les fractions contenant la γ cristalline présentaient un niveau de purification élevé mais des protéines contaminantes étaient présentes (voir Figure 4.4).

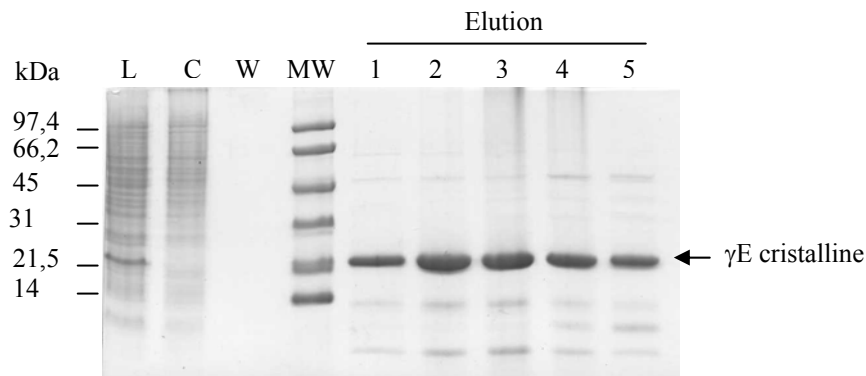


Figure 4.4 : Gel SDS PAGE de l'élué de la colonne échangeuse de cation SP. L : partie soluble du lysat; C : effluent de la colonne; W : lavage; 1-5 : fractions d'élué.

L'étape de précipitation à l'ammonium sulfate a permis d'ajouter une étape de purification et de réduire le volume de la solution protéique. Après la chromatographie d'exclusion, la majorité des protéines contaminantes était éliminée malgré la présence de deux protéines de poids inférieurs (voir Figure 4.5).

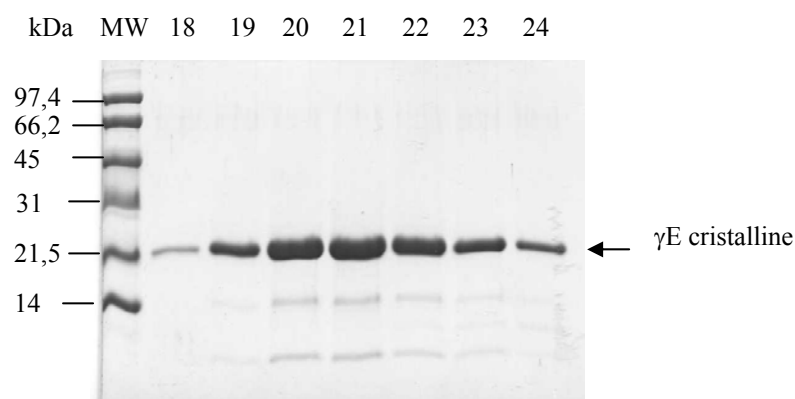


Figure 4.5 : Gel SDS PAGE de l'élué de la chromatographie d'exclusion S200. MW : Marqueur de poids moléculaire, 18-24 : fractions.

Ces bandes ont déjà été observées dans les purifications sur colonne Nickel et SP, et que ce soit pour la protéine γE , la γD ou les protéines natives de rat. Les poids moléculaires approximatifs de ces bandes sont de ~ 14000 Da et ~ 6000 Da. Il est fort probable que ces bandes soient des dégradations des γ cristallines au niveau de la boucle *cd* du motif 3 de la γ cristalline, puisque leur concentration varie avec celle de la cristalline (voir figure 4.3 et 4.5). Or, cette dégradation n'est pas visible sur le profil de la chromatographie d'exclusion (pas de pic au niveau des 14000 ou des 6000 Da). Il est possible que cette dégradation soit visible uniquement lors du dépliement de la protéine causé par le tampon dénaturant de tampon de migration ainsi que le traitement thermique. Ainsi, il y a certainement une digestion protéolytique *in vivo*, puisque la présence d'anti-protéases dans les tampons de purification devrait empêcher une protéolyse *in vitro*. En revanche le repliement général doit maintenir les polypeptides à leur place, pour obtenir une élution identique par chromatographie d'exclusion. Il est, en revanche, moins évident que ces fragments soient la conséquence d'utilisation de codons rares même si ces derniers sont nombreux dans la séquence de la γE cristalline de rat (voir annexe A.B.3). En effet, la présence de ces codons ne fournit pas de fragment de 6 et 14 kDa.

4.3.4 Evaluation de l'incorporation isotopique par spectrométrie de masse

Afin de vérifier le niveau de deutériation, ainsi que la pureté des protéines, le poids moléculaire des γ cristallines a été analysé par spectrométrie de masse. Après déconvolution des spectres enregistrés de la forme hydrogénée, deux pics pouvaient être visibles (voir figure 4.6A). Les masses enregistrées étaient de 22263 Da et de 21132 Da. La masse moléculaire théorique (ExPASy Proteomics Server) de la forme hydrogénée de la γE cristalline est de 21262,9 Da. La masse de 21132 Da, correspond à la perte de la méthionine N-terminal de la protéine. Ceci est une modification post-traductionnelle commune aux protéines procaryotes et eucaryotes, spécialement quand le second acide aminé de la chaîne polypeptidique est occupé par un acide aminé dont la chaîne latérale est petite, telle qu'une glycine pour les γ cristallines (Hirel et al. 1989). Cette modification a été observée lors de la caractérisation des γ cristallines de rat (Ben-Bassat et al. 1987; Siezen et al. 1988). Pour la protéine γE cristalline deutériée, le spectre présente un seul pic (voir figure 4.6B) à 22213 Da, qui correspond à la masse de la protéine deutériée sans méthionine N-terminal. Le nombre

théorique de deutériums incorporés devrait être de 1002. Or la différence obtenue entre les masses expérimentales deutériée et hydrogénée est de 982. Le niveau d'incorporation isotopique est alors de $(982 / 1002) * 100 = 98 \%$. Bien que le niveau d'expression des γ cristallines, ne diffère pas entre une culture hydrogénée et une culture deutériée, il est probable que cette différence réside au niveau de l'enzyme catalysant cette réaction : la Méthionine Amino Peptidase (MAP). En effet, le niveau de translation des protéines a pu diminuer lors d'une culture en D_2O , ce qui permettrait à la MAP d'effectuer toutes les modifications N-terminales.

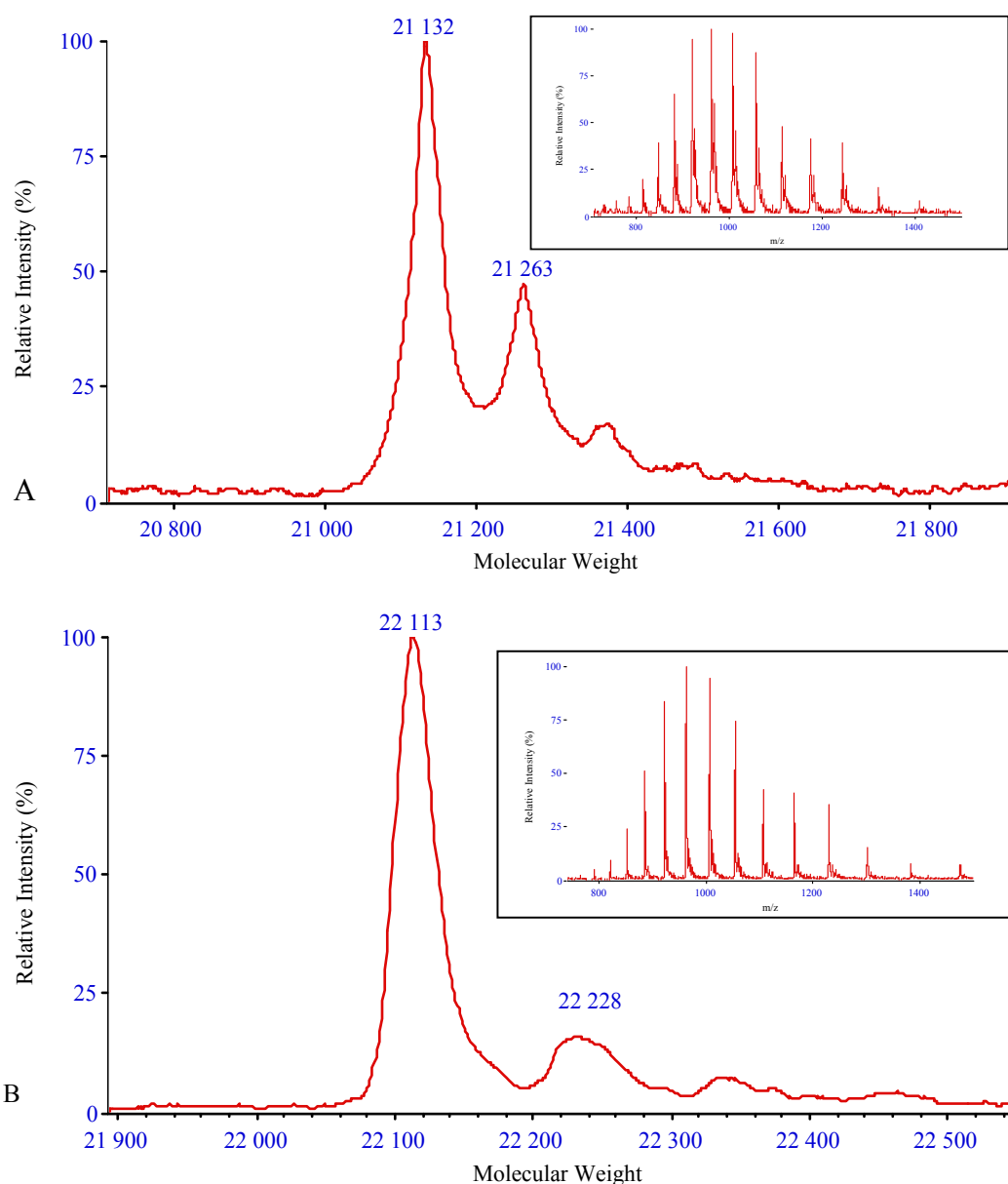


Figure 4.6 : Spectres de masses de la γE cristalline hydrogénée (A) et deutériée (B). En encart, sont représentés les spectres non déconvolués. Le spectre A présente 2 pics majoritaires correspondant aux formes sans et avec la méthionine N-terminale. Le spectre B exhibe un pic qui indique que la protéine se trouve sans méthionine N terminal quand elle est deutériée. Les pics résiduels sont des adduits du TFA avec la cristalline.

Etant donné qu'une donnée essentielle de la cristallogenèse est l'homogénéité des protéines, il est essentiel d'augmenter l'activité de cette protéine au niveau de la protéine hydrogénée. Nous avons décidé de produire la MAP en plus grande quantité en même temps que les cristallines. Cette co-expression doit permettre l'obtention des γ cristallines sans leur méthionine N-terminale. Le clone de la MAP d'*E. coli* a été gracieusement donné par le professeur Brian Matthews mais ce gène a été transféré dans un autre vecteur pour les raisons qui ont été décrites dans le chapitre 3.1. Le nouveau clonage, l'expression et la purification sont détaillés dans la partie Matériels et Méthodes. L'expression de la MAP ne semble pas affecter celle des cristallines en milieu minimum (voir Figure 4.7).

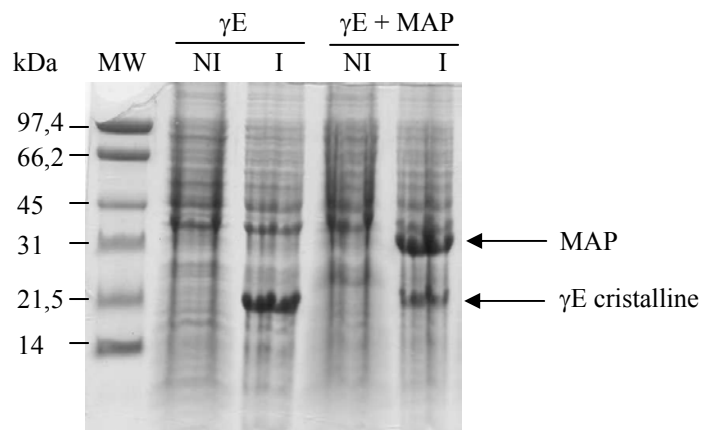


Figure 4.7 : Coexpression de la γ E cristalline et de la MAP visualisée en gel SDS PAGE. Les extraits totaux avant (NI) et après induction à l'IPTG des cellules contenant le plasmide pET γ ENco seul et les plasmides pET γ ENco et le pLMAP, montrent le même niveau d'expression des cristallines.

En revanche, la purification des cristallines va être légèrement modifiée sur la première étape, puisque une partie de la MAP va être purifiée en même temps que les cristallines (voir figure 4.8A). Mais, la MAP ne sera pas précipitée par ammonium sulfate et donc la purification reste donc acceptable, après la chromatographie d'exclusion (voir Figure 4.8B).

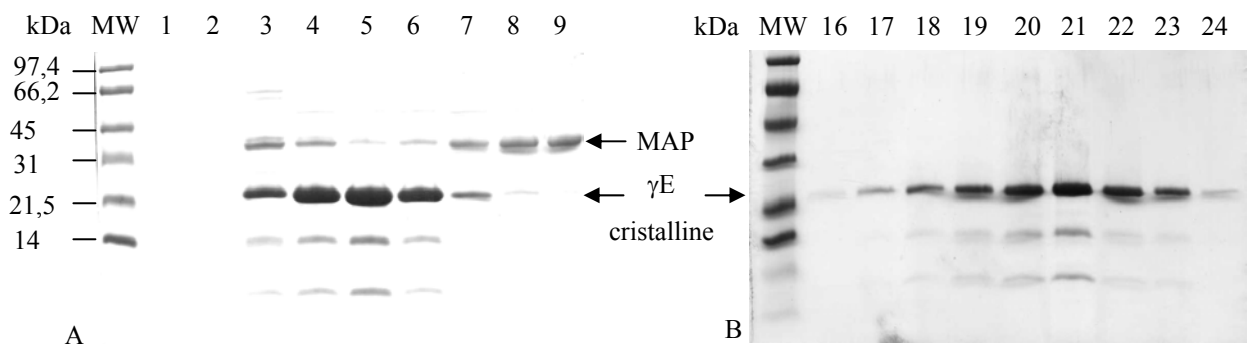


Figure 4.8 : Gel SDS PAGE de la purification de la γ E cristalline exprimée avec la MAP. Le gel A montre l'élué de la colonne SP, et le gel B montre l'élué de la chromatographie S200.

Le spectre de masse (voir figure 4.9) démontre que la purification est beaucoup plus homogène puisque la solution protéique présente une seule espèce de la cristalline sans la méthionine N-terminale.

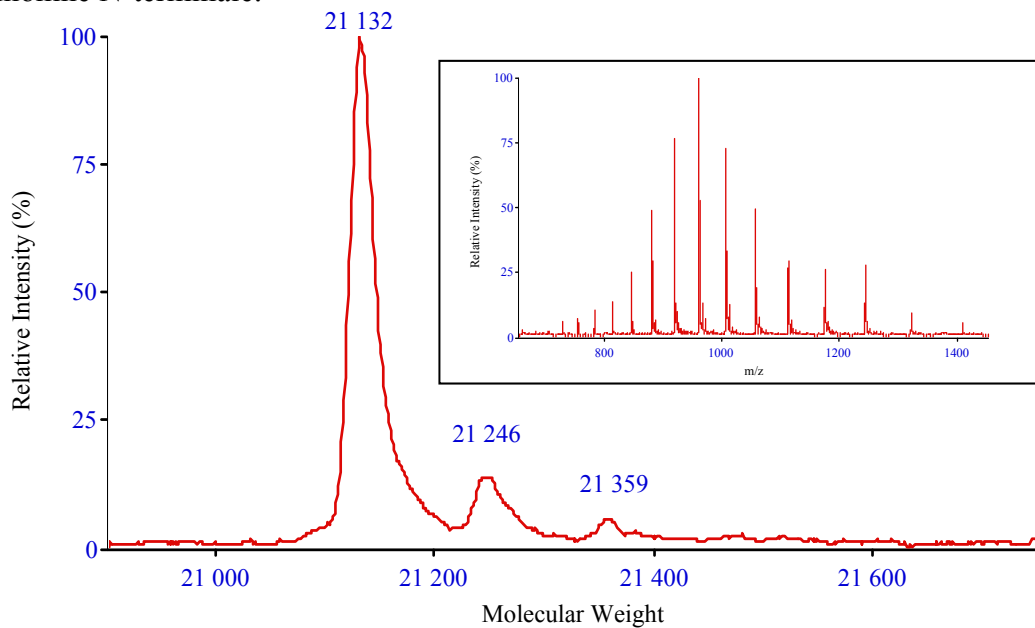


Figure 4.9 : Spectre de masse de la γ E cristalline purifiée et exprimée avec la MAP. En encart, est représenté le spectre non déconvolué. Le spectre montre un pic qui indique que la protéine se trouve sans méthionine N-terminale quand elle est exprimée avec la MAP. Les pics résiduels sont des adduits du TFA avec la cristalline.

Les γ cristallines pures vont pouvoir être caractérisées sous différents points de vue biochimiques et biophysiques.

4.3.5 Gel natif

Les conditions non dénaturantes d'un gel natif ont permis de déterminer les différents conformères des cristallines. Les protéines ont subi, ou non, un traitement réducteur pendant 20 minutes. Les gels natifs réalisés en chambre froide ne permettent pas l'entrée de la γ E dans le gel, la protéine semble être agrégée au fond des puits du gel. La migration a, donc, été réalisée à température ambiante et avec un voltage faible pour éviter de faire fondre le gel. Cette méthode a entraîné un inconvénient majeur qui est la diffusion des protéines à travers le gel. Malgré cette diffusion, il a été possible de visualiser un comportement migratoire bien spécifique (voir figure 4.10). Sans traitement DTT, pendant la purification, la protéine a migré suivant trois espèces que ce soit pour la cristalline hydrogénée ($H\gamma E$) ou deutériée ($D\gamma E$).

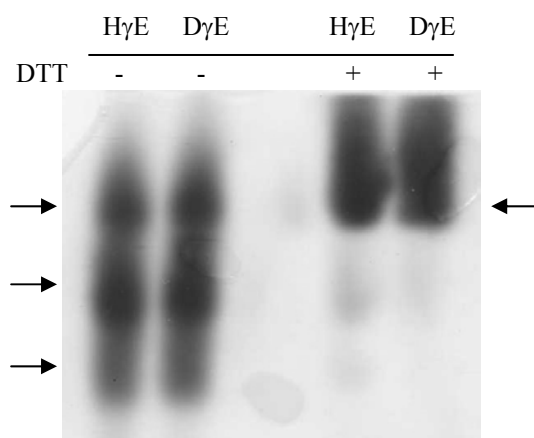


Figure 4.10 : Gel natif 6 % des γ E cristallines hydrogénées (H γ E) ou deutériées (D γ E). Sans traitement réducteur, les cristallines montrent 3 conformères. En présence de DTT, les cristallines n'ont plus qu'une conformation.

En condition réductrice, la γ E cristalline, qu'elle soit sous forme hydrogénée ou deutériée, migre sous condition native en une bande. Cette bande est au même niveau que l'un des isoformes présents en condition non réductrice. La deutériation n'altère pas le nombre de conformation que ce soit en condition réductrice ou non. L'accessibilité de l'agent réducteur, qu'est le DTT, n'est pas modifiée par cette substitution isotopique.

4.3.6 Point isoélectrique

Le point isoélectrique des γ cristallines a été déterminé pour les formes hydrogénées et deutériées. Il serait surprenant que la deutériation change la charge totale de la protéine. Cette expérience n'a pu être réalisée qu'en solvant hydrogéné, car les gels IEF deutériés n'existent pas. Les hydrogènes, ou deutériums, labiles des fonctions carboxylates, alcools, ou ammoniums, responsables de la charge globale de la protéine seront tous des hydrogènes. Le point isoélectrique expérimentale des protéines H γ E et D γ E est de $\sim 7,25$ alors que le théorique est de 7,14 (voir figure 4.11).

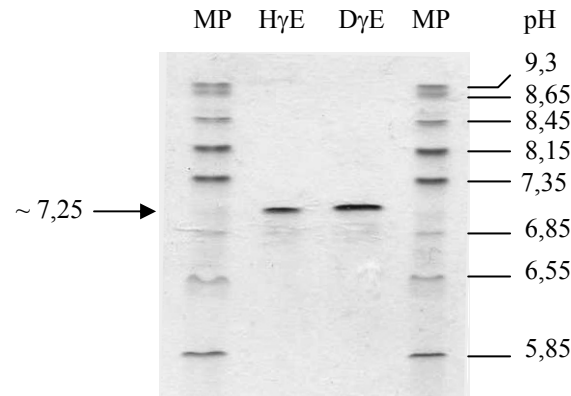
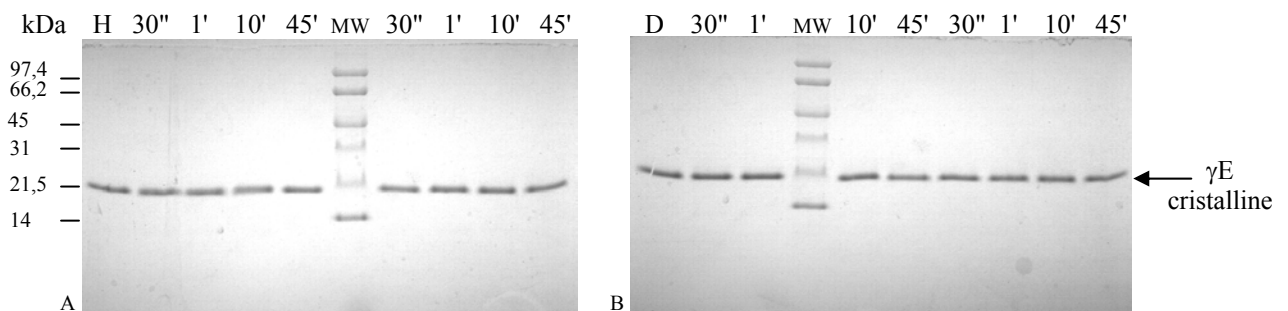


Figure 4.11 : Analyse en IEF des H γ E et D γ E. Le point isoélectrique ne varie pas entre les deux formes, il est de $\sim 7,25$. MP est le marqueur de pH.

La deutériation n'a pas d'effet sur la séparation d'un point de vue charge. Les charges de surface, accessibles au solvant, n'ont pas été modifiées de façon significative par la perdeutériation.

4.3.7 Protéolyse limitée

Les γ cristallines n'ont pas d'activité enzymatique connue. Une étude comparative de protéolyse limitée a été entreprise. Différentes protéases, ont été mises en présence des protéines hydrogénées et deutériées. Les figures 4.12A et 4.12B montrent que les γ E cristallines hydrogénées et deutériées sont résistantes à la digestion par la subtilysine. Il n'y a pas de digestion au cours du temps bien que le ratio protéase sur cristalline soit de 1/20.



Figures 4.12A et B: Digestion protéolytique par la subtilysine dans les rapports 1/150 et 1/50 (protéase / protéine) de la γ E cristalline hydrogénée (A) et deutériée (B) au cours du temps. H et D représentent les protéines sans enzyme, et les signes " signifient seconde et ' minute. La protéine est résistante à la digestion ce qui démontre la stabilité des γ cristallines. La perdeutériation n'a pas d'influence sur cette gamme de temps et à ces concentrations de protéase.

Les digestions protéolytiques par la trypsine et l'élastase sont identiques dans les conditions précédentes. La protéine commence à être hydrolysée, uniquement, par la subtilysine après 20 heures d'incubation (voir figure 4.13) que ce soit pour la protéine hydrogénée ou deutériée.

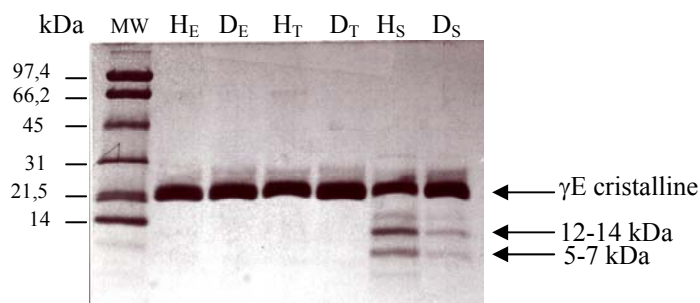


Figure 4.13 : Digestion protéolytique de la γ E cristalline hydrogénée (H) et deutériée (D) par l'élastase (E), la trypsine (T) et la subtilysine (S) après 20 heures d'incubation.

Les fragments générés sont de deux poids moléculaires différents. Approximativement, les fragments de 12-14 kDa et de 5 à 7 kDa peuvent correspondre à une coupure entre au niveau de la boucle *cd* du motif 3, comme dans le cas énoncé de purification. Sur ce dernier gel, il est possible de voir que la digestion semble légèrement plus importante chez la protéine hydrogénée.

Il a été décrit que le taux de coupure des liaisons C-H, N-H et O-H était 7, 8,5, et 10,6 fois plus rapide que le taux des analogues deutériés (Wade 1999). Il est certain que, lors d'une digestion protéolytique, ce ne sont pas ces liaisons qui sont mises en jeu. Mais la liaison peptidique peut se retrouver plus forte du fait de la présence d'atomes de deutériums voisins et de la délocalisation de leurs électrons sur les liaisons avoisinantes. Il a été aussi démontré que, dans certains cas comme celui de glutathione-S-transférase, la protéine deutériée était plus sensible à la coupure protéolytique (Brockwell et al. 2001). Dans ces circonstances, les barrières d'activation pour les fluctuations transitoires du dépliement local sont considérablement abaissées. Pour essayer d'abaisser ces barrières énergétiques, une protéolyse a été réalisée dans des conditions légèrement dénaturantes (pH acide, 1M urée) avec une enzyme dont la spécificité est moins élevée, la pepsine (voir figures 4.14). Pour la protéine hydrogénée ou la protéine deutériée, la protéolyse a produit des fragments de même taille qui doivent correspondre essentiellement à la libération des peptides de ~ 6 kDa, correspondant à la taille d'un motif en clé grecque.

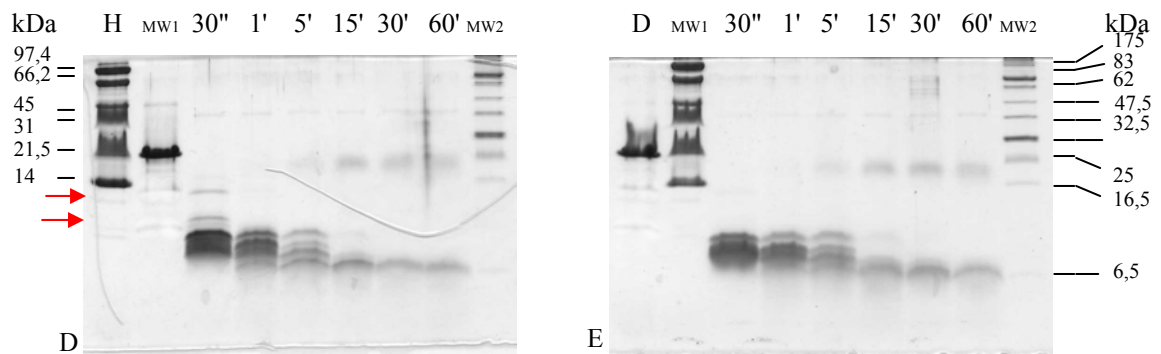


Figure 4.14: Digestion protéolytique ménagée de la γ E cristalline hydrogénée (D) et deutériée (E) par la pepsine en condition semi-dénaturante. H et D représentent les cristallines sans élastase, et les signes ''et 'signifient le temps d'incubation en seconde et minute. Il n'y a pas de différence remarquable sauf au niveau du temps d'incubation de 30'' où pour l'hydrogénée, il y a deux bandes marquées en rouge.

Peu de différences sont visualisables, mis à part deux peptides générés après 30 secondes d'incubations pour la protéine hydrogénée. Ces peptides sont absents sur le gel de protéolyse de la γ E cristalline deutériée. Il serait étonnant que les deux protéines n'engendrent pas les mêmes peptides. Donc il est plus approprié de penser que ces deux peptides ont été produits par les deux systèmes mais que ceux-ci ont été hydrolysés moins rapidement dans le cas de la protéine hydrogénée. Ainsi la protéine hydrogénée présenterait une stabilité légèrement plus importante, mais qui doit être démontrée par d'autres investigations.

La structure des γ cristallines est extrêmement stable et difficile à dégrader par des enzymes protéolytiques (David et Shearer 1986). Cette très forte résistance de la γ E cristalline à ces protéolyses ménagées est toujours aussi importante que la protéine soit hydrogénée ou deutériée. L'échange isotopique n'a donc, pas affecté drastiquement les forces mises en jeu dans la stabilité produite par la conformation spécifique des γ cristallines, vis à vis des protéolyses ménagées. Il est quand même intéressant de se rendre compte qu'il existe un faible effet concordant avec les observations déjà existantes sur la perdeutériation. Dans notre étude, cette différence montre une relative plus haute stabilité de la protéine sous forme hydrogénée, comme il a été montré, précédemment, au niveau d'une protéolyse ménagée de la Glutathione S-transferase de *S. japonicum* (Brockwell et al. 2001). Mais ces observations étaient beaucoup plus évidentes dans le cas de la GST.

Afin de confirmer ces premières observations, d'autres investigations sont nécessaires, et des études par spectroscopie devraient poursuivre ces résultats.

4.3.8 Spectroscopie infrarouge à transformée de Fourier

La spectroscopie infrarouge à transformée de Fourier (ou FTIR) est basée sur l'absorption d'un rayonnement infrarouge par la protéine analysée. Elle permet via la détection des vibrations caractéristiques des liaisons chimiques, d'effectuer l'analyse des fonctions chimiques présentes dans la molécule et de leur environnement. Lorsque la longueur d'onde (l'énergie) apportée par le faisceau lumineux est voisine de l'énergie de vibration d'une liaison protéique, cette dernière va absorber le rayonnement et la diminution de l'intensité réfléchie ou transmise va être enregistrée. Le domaine infrarouge entre 4000 cm^{-1} et 400 cm^{-1} ($2,5 - 25\ \mu\text{m}$) correspond au domaine d'énergie de vibration des liaisons des molécules. Toutes les vibrations ne donnent pas lieu à une absorption, cela va dépendre aussi de la géométrie de la protéine et en particulier de sa conformation. La position de ces bandes d'absorption va dépendre en particulier de la différence d'électronégativité des atomes et de leur masse. Par conséquent, à une protéine de composition chimique et de structure donnée va correspondre un ensemble de bandes d'absorption spécifiques permettant d'identifier la molécule et ses caractéristiques structurales.

La bande infrarouge, la plus analysée pour l'étude des protéines, est celle de l'amide I qui se situe vers 1600 et 1700 cm^{-1} et qui est principalement due aux vibrations du groupe C=O des résidus. Du fait que les groupements C=O sont impliqués dans des éléments de structures secondaires *via* les liaisons hydrogène au groupe N-H, une bande amide I renferme une multitude de bandes avec des fréquences spécifiques. Ces différentes composantes peuvent être attribuées à des éléments de structures secondaires (Jung 2000). Les pourcentages de structure secondaire peuvent être estimés à partir des surfaces relatives à ces bandes.

D'autres bandes spectrales peuvent être utilisées pour ce type d'analyses. La région de l'amide II (vers 1550 cm^{-1} en H_2O et vers 1450 cm^{-1}) est la conséquence de la courbure du groupe NH couplé à l'étirement de la liaison C-N. Cette région est rarement utilisée car elle est très voire trop sensible à la deutériation. La bande de l'amide III est trouvée entre 1220 et 1330 cm^{-1} et résulte de la courbure de la liaison N-H et du l'étirement de la liaison C-N. A cause de sa faible intensité, cette région est très difficilement exploitable.

4.3.8.1 Attribution des structures secondaires par FTIR

Cette expérience n'a pas été réalisée en eau légère puisque les modes de vibrations dans l'infrarouge de l'eau légère sont très intenses et peuvent chevaucher les bandes d'intérêt de l'échantillon, spécialement au niveau de l'amide I. Ce problème est généralement résolu en substituant l'eau légère par de l'eau lourde. Les modes de vibrations interviennent à des fréquences différentes de ceux observés pour l' H_2O , à cause de la dépendance de la masse de la fréquence de vibration (Stuart 1997).

La figure 4.15 montre les spectres d'absorption des protéines hydrogénée et deutériée dans une gamme spectrale de 1350 à 1800 cm^{-1} .

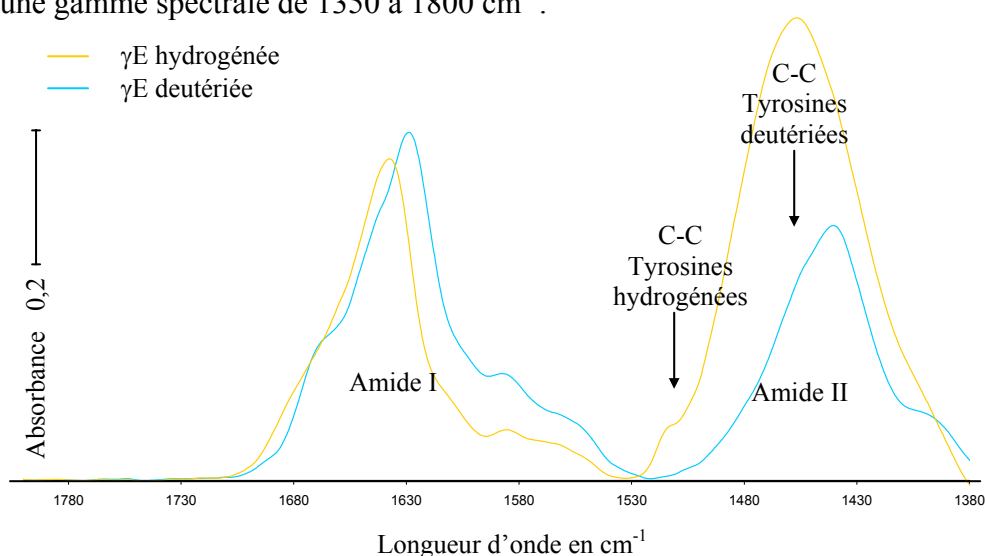


Figure 4.15 : Spectres d'absorption infrarouge des γE cristallines hydrogénée (—) et deutériée (—) en D_2O .

Le spectre de la protéine deutériée présente un déplacement général vers les longueurs d'onde plus élevées d'environ 10 cm^{-1} : la bande amide I est décalée de 9 cm^{-1} , alors que la bande amide II l'est de 15 cm^{-1} . Le maximum d'absorbance de la bande de l'amide I est de 1629 et 1638 cm^{-1} respectivement pour les protéines hydrogénée et deutériée. Les très larges bandes correspondant à l'amide II sont centrées aux environs de 1441 et 1456 cm^{-1} .

Dans le spectre de la protéine hydrogénée, il est possible de trouver une bande vers 1515 cm^{-1} , caractéristique de la vibration des liaisons C-C des noyaux aromatiques des tyrosines et des phénylalanines, qui sont au nombre de 16 et 14, respectivement, dans la γE cristalline. Celle-ci est en partie cachée par la bande de l'amide II. En revanche, dans le spectre de la protéine deutériée, cette bande est déplacée vers 1450 cm^{-1} (Hellwig et al. 2002).

Le nombre, la position et la surface des bandes obtenues permettent d'obtenir directement les quantités relatives des structures qu'elles représentent. Les pourcentages en hélices α , des structures β et des coudes β sont estimés en additionnant les surfaces de toutes les bandes assignées à chaque structure. Le spectre déconvolué de la protéine perdeutériée, ainsi que sa dérivée seconde pour la région amide I, sont montrés dans la figure 4.16.

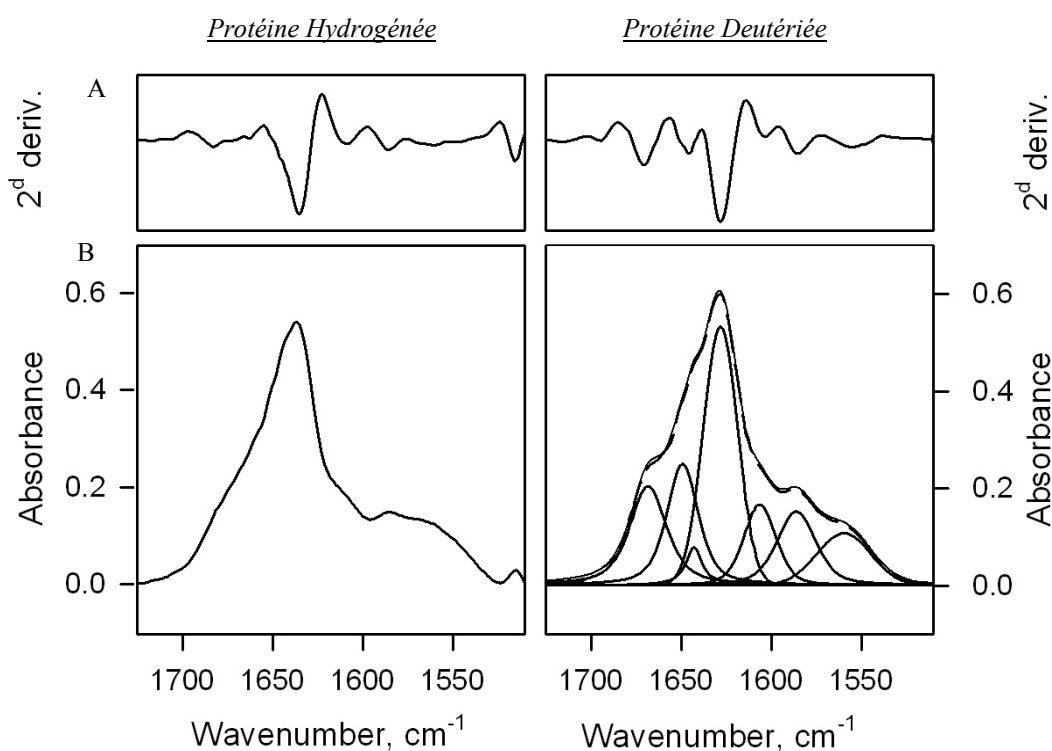


Figure 4.16 : A) Dérivées seconde des spectres infrarouges des γ E cristallines hydrogénée et deutériée. B) Spectres infrarouge déconvolués des mêmes protéines. La protéine hydrogénée (spectre de gauche) n'est pas représentée étant donné la mauvaise ligne de base de l'échantillon.

La déconvolution du spectre de la protéine hydrogénée n'est pas représentée car la ligne de base n'était pas correcte sur les différentes tentatives de collecte. En revanche, il est possible de se rendre compte que la forme générale du spectre est similaire, malgré un déplacement de $\sim 10 \text{ cm}^{-1}$ du spectre de la protéine deutériée vers les longueurs d'onde inférieure. Cette variation a été observée dans une étude sur le cytochrome P450 de *Pseudomonas putida*, entre les formes hydrogénée et deutériée (Meilleur et al. 2004).

Le tableau 4.2 donne la décomposition de la bande de l'amide I de la protéine deutériée en D_2O , avec l'attribution de chaque fréquence :

Fréquence en cm^{-1}	Largeur	Hauteur	Surface	Attribution	Pourcentage
1559	35,33	0,11	4,04	N-H résiduels	-
1585	25,22	0,15	4,94	COOD	-
1606	21,29	0,16	4,29	Chaîne latérales Arginine	-
1628	22,46	0,53	12,72	Brins β	45,8 %
1642	22,49	0,25	6,85	Désordonné	4,7 %
1649	10,82	0,08	1,32	Hélice α	24,7 %
1668	24,42	0,20	6,87	Coude et brins β	24,8 % coude β et brins β

Tableau 4.2 : Attribution des bandes servant la déconvolution de la bande spectrale amide I du spectre de la protéine deutériée.

La déconvolution et l'ajustement à la courbe ont permis de décomposer la région de l'amide I en 7 pics distincts assignables à des structures secondaires. L'attribution de ces pics à des structures secondaires est principalement marquée par une bande majoritaire dont le maximum est à 1628 cm^{-1} . Cette bande est caractéristique de structures secondaires β , comme celle présente à 1637 cm^{-1} dans le spectre de Lamba et al. (1993) pour la protéine hydrogénée. Ensuite différents pics sont reconnaissables. Trois bandes, centrées en 1559, en 1585 et en 1606 cm^{-1} ne contribuent pas à la bande de l'amide I. La première bande est essentiellement due à des N-H résiduels. En effet, la protéine a été exprimée par une bactérie, dans un milieu de culture deutérié, mais a été par la suite purifiée dans des solvants hydrogénés pendant plusieurs jours. Certains hydrogènes n'ont pas été ré-échangés par dialyse contre du D_2O à cause de la haute stabilité de la protéine. La bande, située aux alentours de 1585 cm^{-1} , correspond aux fonctions acides (COO^-) des chaînes latérales des aspartates et des glutamates qui sont chargées au pH utilisé. Enfin, le pic, localisé en 1606 cm^{-1} , est typique des vibrations des chaînes latérales des acides aminés.

La γE cristalline est composée essentiellement de structure secondaire en brins β , comme il a été précisé auparavant. Mais, il existe d'autres composantes dans les γ cristallines, telles que les hélices α ou les coudes β . L'absorption infrarouge caractéristique des hélices α se produit dans une région étroite mais bien définie entre 1650 et 1660 cm^{-1} (Lamba et al. 1993; Stuart 1997; Jung 2000). Si le déplacement de cette région de 10 cm^{-1} , suite à la deutériation est pris en compte cette bande se positionne entre 1640 et 1650 cm^{-1} . La surface de cette bande est considérable puisqu'elle représente jusqu'à 24,7% de région consacrée à l'amide I. La bande centrée en 1668 cm^{-1} est spécifique, en considérant le

déplacement de 10 cm^{-1} , de brins β (surtout les feuilletés anti-parallèles), mais cette composante peut aussi contenir une forte contribution de coudes β . Enfin, vers 1643 cm^{-1} , se situe une très petite bande significative des régions non structurées. En résumé, la composante majeure est la structure en brins β qui est représenté avec plus de la moitié des structures secondaires des γ cristallines (voir tableau 4.3), ensuite les hélices α sont présentes, suivent les coudes β et une très faible proportion non structurée.

4.3.8.2 Comparaison des données infrarouge avec d'autres analyses hydrogénées

Comme il a été énoncé, il n'a pas été possible d'attribuer des structures secondaires à la protéine hydrogénée en raison d'une mauvaise ligne de base. En revanche, il existe une analyse faite entre les données en spectroscopie infrarouge, en rayons X et en dichroïsme circulaire avec des γ cristallines hydrogénées (Lamba et al. 1993). Nos données infrarouges, ainsi que celles obtenues en dichroïsme circulaire sur les différentes formes isotopiques de la γ E cristalline de rat, peuvent être comparées dans le tableau 4.3.

Avant toute comparaison, il est important de noter que les γ cristallines hydrogénées citées dans le tableau 4.3, ne sont pas identiques à la γ E puisqu'elles ne proviennent pas du même organisme et n'ont pas une identité de séquence.

Bien que toutes les γ cristallines possèdent une forte quantité de brins β , cette composante est variable de 51 à 71 % suivant la technique d'analyse et la cristalline étudiée. Les pourcentages d'hélices α et de coudes β sont aussi très variables, voire même largement différents quand la structure deutériée est comparée à la γ cristalline étudiée en infrarouge, ou encore avec la même protéine mais étudiée en dichroïsme circulaire. Le pourcentage en hélices α est doublé pour la γ E cristalline aux dépens, essentiellement, des coudes β . Mais cette variabilité peut éventuellement s'expliquer par l'approximation faite sur la bande centrée en 1668 cm^{-1} . En effet, cette bande contient une part de brins β mais aussi des coudes β . L'approximation faite est 50 % pour les brins β et 50 % pour les coudes β . Il est certain que pour visualiser l'effet de la perdeutériation en résonance infrarouge par transformée de Fourier, sur la γ E cristalline, il aurait été plus aisé d'étudier la γ E cristalline hydrogénée dans les mêmes conditions.

<i>γ cristalline</i>	<i>Brins β</i>	<i>Hélices α</i>	<i>Coudes β</i>	<i>Autres</i>	<i>Méthodes</i>	<i>Référence</i>
<i>γ ii(γB) de bœuf</i>	53-71	19-27	19-31	-	Rayons X	(Wistow et al. 1983)
<i>γ ii(γB) de bœuf</i>	52	27	21	-	Rayons X	(Blundell et al. 1981)
<i>γ ii(γB) de bœuf</i>	56-61	0-5	11-15	24-28	DC	(Mandal et al. 1985)
<i>γ iii(γD) de bœuf</i>	46-58	0-5	13-18	29-31	DC	(Mandal et al. 1985)
<i>γ iv(γA ou E) de bœuf</i>	45-52	0-5	17-20	30-31	DC	(Mandal et al. 1985)
<i>γ de bœuf(non précisé)</i>	68	2	5	24	DC	(Siezen et Argos 1983)
<i>γ C de bœuf</i>	51	15	8	27	DC	(Fu et Liang 2002)
<i>γ de bœuf (non précisé)</i>	52	12	32	0	FTIR	(Lamba et al. 1993)
<i>γ E deutériée en D₂O (DγEd2o) de rat</i>	58,2	24,7	24,8	4,7	FTIR	Ce manuscrit
<i>HγEd2o</i>	55,6	16,7	21,4	6,3	DC	Ce manuscrit
<i>HγEd2o</i>	55,5	16,8	21,5	6,2	DC	Ce manuscrit
<i>DγEh2o</i>	56,2	16,8	21,4	5.6	DC	Ce manuscrit
<i>DγEd2o</i>	56,2	16,9	21,4	5.5	DC	Ce manuscrit

Tableau 4.3 : Estimation, en pourcentage, des structures secondaires de différentes γ cristallines de bœuf et de rat. Les valeurs sont issues de la publication de Lamba et al. (1993). Sont ajoutées les valeurs de la γ C de bœuf et celles de la γ E de rat étudiée dans cette thèse. La précision des valeurs de la spectroscopie infrarouge (FTIR) est de +/- 4%, et celles du dichroïsme circulaire (DC) sont de 5 %.

L'analyse effectuée sur le cytochrome P450cam, par Meilleur et al. (2004) en FTIR n'a pas démontré de changement structural bien qu'ils aient eu des problèmes d'attribution de suite au décalage isotopique. Ils avaient perdu 12 % d'hélices α au profit de régions désordonnées, d'hélices 3_{10} et de coudes β du fait que les bandes étaient très proches et se chevauchaient.

Dans la présente étude, nous avons certainement le même problème d'attribution car la protéine deutériée a gagné un certain pourcentage d'hélice α au profit des coudes β . Il sera démontré plus tard, que les structures cristallographiques des protéines hydrogénées et deutériées ne présentent aucune divergence de structures secondaires.

Meilleur et al. (2004) ont discuté de l'effet de la perdeutériation sur le spectre infrarouge. Le décalage des bandes amide I et II vers des longueurs d'onde plus faibles est expliqué par une augmentation de la masse. La région de l'amide résulte principalement de

l'étirement du groupement C=O des peptides. Dans un modèle simplifié d'un oscillateur harmonique, le remplacement isotopique des hydrogènes va conduire à une augmentation de la masse effective du carbone, induisant une masse réduite μ plus élevée dont la racine carré est inversement reliée à la fréquence de vibration (voir équation 4.5).

$$\mu = \frac{m_c m_o}{(m_c + m_o)} \quad \text{et} \quad \nu = \frac{1}{2\pi} * \sqrt{\frac{k}{\mu}} \quad (4.5)$$

où m_c et m_o représentent la masse effective du carbone et de l'oxygène, μ est la masse réduite et k , la constante de Boltzmann.

Un autre effet observable de la perdeutériation est le décalage de la vibration de la liaison C-C des noyaux aromatiques des tyrosines, centrée aux alentours des 1515 cm^{-1} dans la protéine hydrogénée, et qui migre vers 1450 cm^{-1} quand la protéine est perdeutériée. Cette dernière est partiellement cachée par celle de l'amide II.

Etant donné les problèmes concernant la protéine hydrogénée, les problèmes d'attributions des bandes pour la protéine deutériée, il est préférable d'envisager d'autres techniques d'études que la spectroscopie infrarouge, pour affirmer que la perdeutériation n'a pas eu d'effet sur les structures secondaires de la γE cristalline.

4.3.9 Dichroïsme circulaire

La spectroscopie en dichroïsme circulaire, est une technique idéale pour comparer les conformations des protéines. Dans l'UV lointain, de 190 à 240 nm, elle permet d'étudier préférentiellement les structures secondaires. Dans l'UV proche, de 240 à 340 nm, les spectres sont plus révélateurs de structures tertiaires locales aux environs des résidus aromatiques et des ponts di-sulfures. Les spectres dans l'UV proche sont souvent plus sensibles aux différences conformationnelles. En revanche, une étude dépendante de la température ou d'un agent dénaturant, peut être faite pour, à la fois, quantifier et comparer les stabilités thermiques.

Les spectres de dichroïsme circulaires dans l'UV lointain, à température ambiante (20°C), des protéines hydrogénées et deutériées dans l'eau légère et l'eau lourde sont montrés dans la figure 4.17. La forme des quatre spectres est très similaire, elle montre une ellipticité négative centrée aux alentours de 216-218 nm, caractéristique d'une conformation en feuillet

β . Ces spectres sont semblables à ceux des γ cristallines du bœuf étudiés par Mandal et ses collaborateurs (1987).

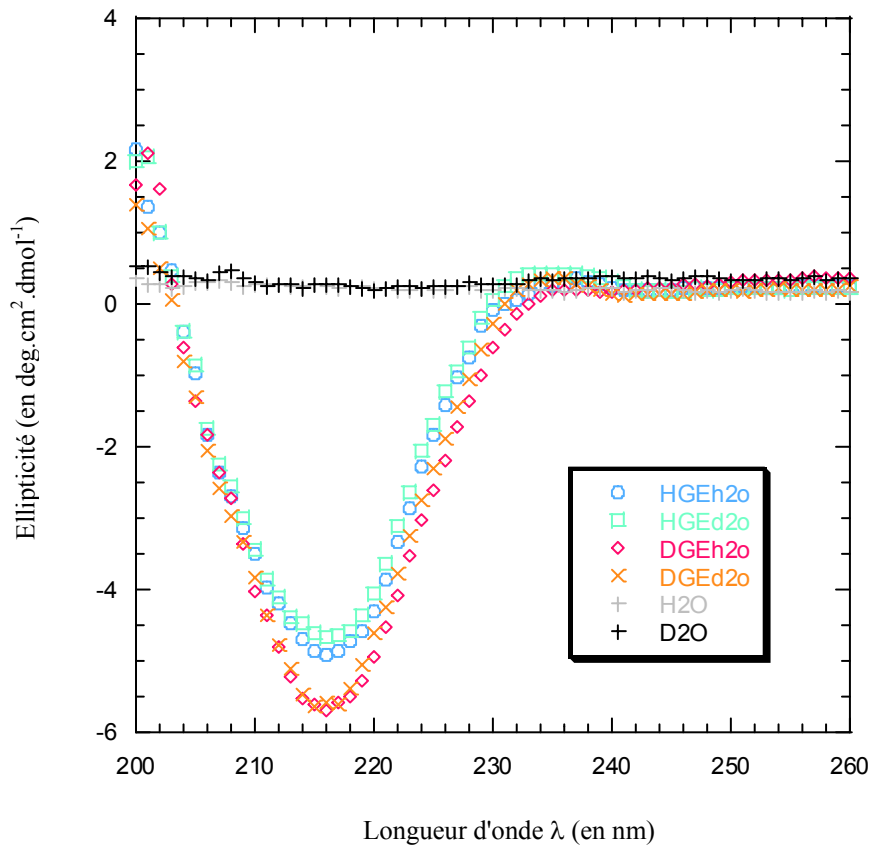


Figure 4.17: Spectres de dichroïsme circulaire dans l'UV lointain des différentes formes de la γ E cristalline. HGEh2o signifie la protéine hydrogénée dans l'eau, et HGEd2o, signifie la protéine hydrogénée en eau lourde et ainsi de suite.

Les pourcentages des structures secondaires ont été déterminés par le programme *Cdnn 2.1* et sont donnés dans le tableau 4.3. En comparant les spectres ainsi que les pourcentages en structures secondaires, il est possible de dire qu'il n'y a pas de changement de structure secondaire entre les différentes formes isotopiques de la protéine. Les dissemblances, qui peuvent être observées, sont probablement causées uniquement par une petite différence en concentration, entre les échantillons. Pour vérifier cette similarité, une étude de stabilité des différentes protéines est réalisée par des dénaturations chimiques et thermiques visualisées par dichroïsme circulaire.

4.3.9.1 Dénaturation chimique

Tout d'abord, le dépliement des différentes γ E cristallines par voie chimique a été suivi à 217 nm. La figure 4.18 montre le pourcentage de dénaturation de la protéine en

fonction de la concentration en guanidinium chloride. Cette perte de structure secondaire a été mise en pourcentage pour pouvoir comparer les différentes espèces.

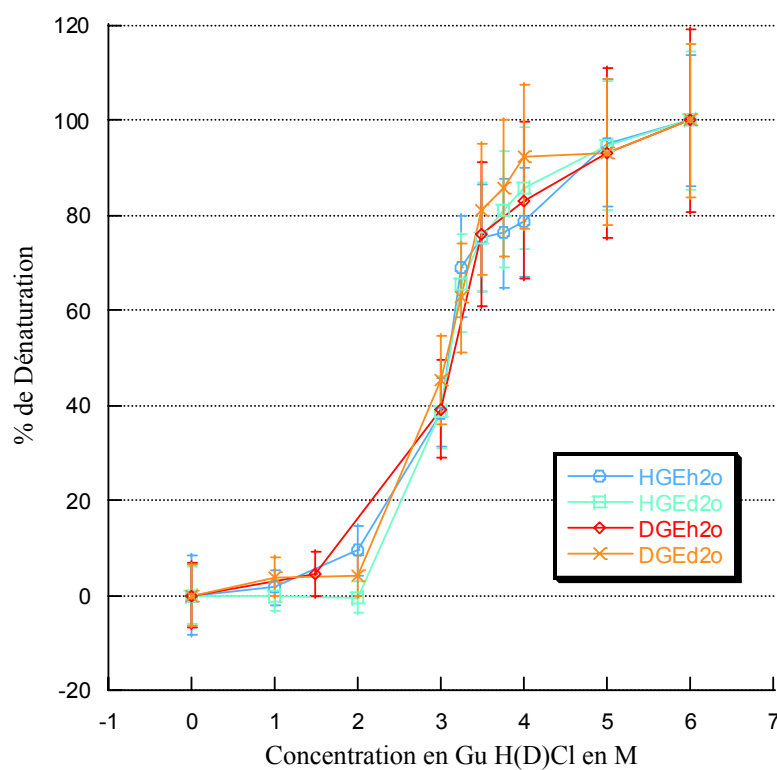


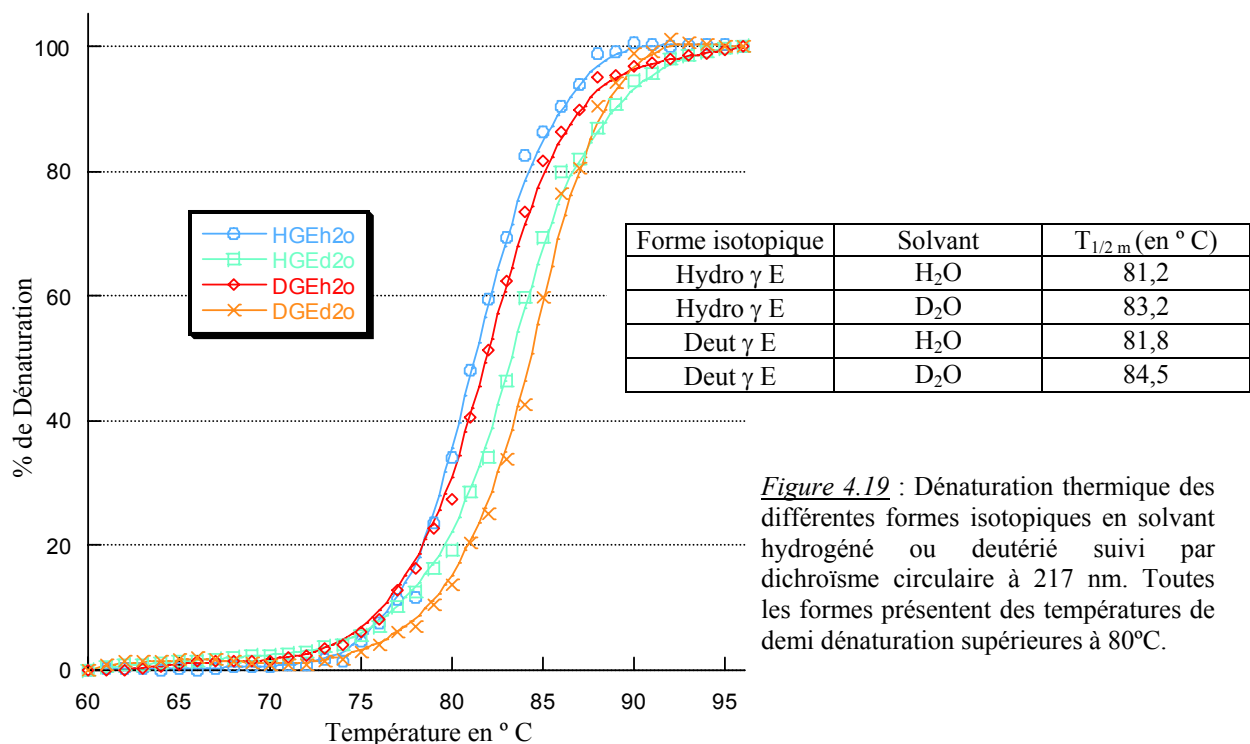
Figure 4.18 : Courbes dénaturation par le GuH(D)Cl des γ E cristallines sous forme hydrogénée et deutériées en H_2O et en D_2O . HGEh2o signifie la protéine hydrogénée dans l'eau, et HGEd2o, signifie la protéine hydrogénée en eau lourde et ainsi de suite.

Les courbes de dénaturation ont sensiblement la même forme et passent à peu près par les mêmes points. La concentration à laquelle se trouve la semi dénaturation est aux environs de 3 molaires en guanidinium H(D)Cl dans tous les cas. Mais, les marges d'erreurs sont trop importantes : de 4 à 9% de la valeur. A cause de ces erreurs, il n'est pas possible d'étudier si la perdeutériation a un effet sur la stabilité de la protéine vis à vis d'un dénaturant chimique. Il est certain que, si effet il y a, celui ci ne peut être que minime étant donné la forte similarité des courbes de dénaturation.

4.3.9.2 Dénaturation thermique

L'enregistrement de la dénaturation thermique devrait être beaucoup plus précis puisque, l'augmentation en température se fait très lentement ($20^\circ C/heure$) et la température est bien contrôlée par un Peletier ($0,1^\circ C$ de précision). Les dénaturations thermiques des différentes espèces ont été générées et les données se révèlent être reproductibles à quelques dixièmes de degrés près (voir figure 4.19). La première évidence est que la dénaturation de la

γ E cristalline ne s'effectue qu'à très haute température, aux alentours de 80°C (Sen et al. 1992). Les courbes de dénaturation ont une allure identique mais il s'avère que les températures de demi-dénaturation sont légèrement divergentes suivant les différentes formes isotopiques. En effet, les protéines dans le solvant hydrogéné présentent des points de demi dénaturation inférieurs à ceux des protéines dans un solvant deutérié. La protéine hydrogénée dans l'eau lourde possède un point de demi dénaturation plus élevé de 2,0°C dans l'eau légère ; et les protéines deutériées ont une différence de 1,7°C dans le même sens. La substitution isotopique démontre une augmentation de stabilité. La protéine deutériée présente un point de demi dénaturation plus élevé de 0,6°C dans l'eau légère et de 1,3°C dans l'eau lourde. L'eau lourde a un effet stabilisant de la γ E cristalline vis à vis de l'eau légère.



Cette stabilisation par l'eau lourde a été démontrée à de nombreuses reprises et même exploitée pour l'étude de protéine. Des structures de complexes composés de sous unités, telles que la malate déshydrogénase halophile (Bonnete et al. 1994) ou la F1-ATPase, ont montré une stabilisation en présence de D₂O. Cette propriété de l'eau lourde a été utilisée dans la cristallisation de la F1-ATPase mitochondriale du cœur bovin (Lutter et al. 1993).

4.3.10 La stabilité des protéines

La stabilité d'une protéine doit être considérée comme un ensemble de forces, qui s'additionnent ou qui s'opposent, agissant à différents niveaux. En effet, les forces stabilisantes ou déstabilisantes peuvent avoir pour origine la composition même en acides aminés de la protéine, mais aussi le solvant environnant cette dernière. Dans le cas d'une dénaturation thermique comme représentée sur le schéma suivant, la protéine peut être dans deux états : le premier est natif et structuré et le second est déplié et sans interaction intramoléculaire apparente.

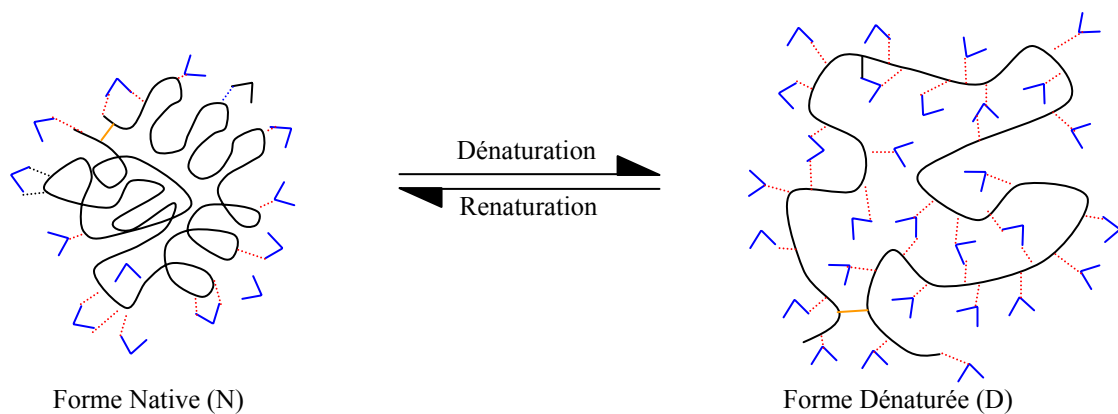


Figure 4.20 : Représentation schématique d'une dénaturation réversible d'une protéine. Les molécules d'eau interagissant avec la protéine sous les formes native et dénaturée.

Sous conditions physiologiques, la réaction de repliement d'une protéine est une réaction thermodynamiquement favorable suivant l'équation de Gibbs, si l'énergie libre est négative :

$$\Delta G_D^N = \Delta H_D^N - T\Delta S_D^N \quad (4.6)$$

où ΔG_D^N est l'énergie libre de la réaction de repliement, ΔH_D^N est l'enthalpie du système biologique (sous pression et température constante) lors de la réaction, T est la température du système et ΔS_D^N est l'entropie du système représentant une mesure du désordre du système.

Dans un état non replié, une protéine montre de nombreuses conformations possibles exprimant une flexibilité très importante comparée à l'état replié. Cette haute liberté de configurations confère une haute entropie S .

$$S = k \cdot \ln W \quad (4.7)$$

où S est l'entropie, k une constante et W , le nombre d'états possibles. Lors du repliement de cette protéine, le nombre de configurations va être restreint, donc l'entropie va être réduite. Le

fait que la protéine va prendre un état ordonné, il va avoir, de le même temps, un gain enthalpique et une perte entropique. Les variations sont grandes et de quantités négatives.

Cette réaction de repliement est favorable étant donné que l'énergie libre totale est négative. Pour qu'il y ait dénaturation à partir de l'état natif, il faudrait apporter l'équivalent énergétique de ΔG . Dans ce type de réaction, l'entropie et l'enthalpie du solvant (l'eau) sont à prendre en compte. Pendant le repliement, la protéine adopte une structure définie pour différentes raisons, mais généralement c'est une question d'interaction de van der Waals, de liaisons hydrogènes et surtout d'effet hydrophobe (voir figure 4.21 et tableau 4.4).

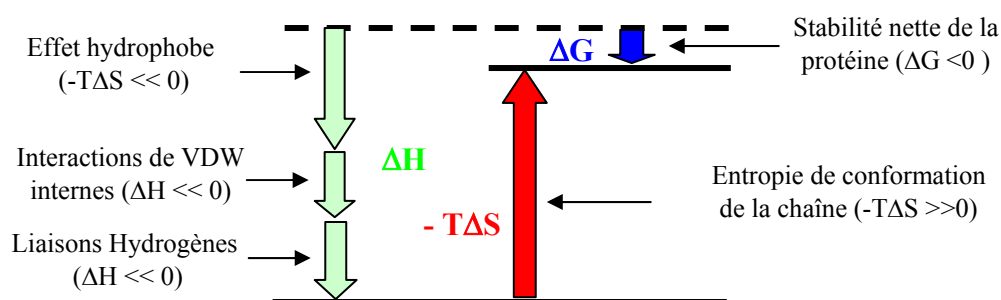


Figure 4.21 : Diagramme montrant les contributions relatives à l'énergie libre ΔG pour le repliement d'une protéine.

<i>Interactions</i>	<i>Energie (kJ/mole)</i>	<i>Localisation</i>	<i>Requis</i>
<i>van der Waals</i>	<1	Partout	---
<i>Interactions électrostatiques</i>	4 (surface) 20 (à l'intérieur)	Principalement en surface	Charge des chaînes latérales (dépendant du pH)
<i>Liaisons hydrogènes</i>	10-60?	Intramoléculaire et avec le solvant	Accepteur donneur
<i>Effet hydrophobe</i>	20-50	A l'intérieur	Acides aminés hydrophobes

Tableau 4.4 : Différentes interactions non covalentes intervenant dans la stabilisation de la protéine.

Dans un état dénaturé, les chaînes latérales hydrophobes des résidus apolaires sont exposées à l'eau. A cause de cela, l'eau va tendre à former des sortes de "cages" appelées clathrates, autour de ces résidus non polaires (sans interactions avec ces derniers), l'entropie de l'eau va diminuer. En revanche, beaucoup de liaisons hydrogène avec l'eau vont être formées à partir de la chaîne principale et des chaînes latérales des résidus polaires. Il y aura donc aussi, une diminution d'entropie de l'eau. Lors du repliement, l'incompatibilité de ces zones hydrophobes avec l'eau va tendre à minimiser cette surface en encourageant la formation de structures tertiaires. Une telle diminution de surface hydrophobe est nécessairement accompagnée par la formation de structures secondaires maintenues par des liaisons hydrogène, entre les différents acides aminés en parallèle de leur désolvation (Fernandez et al. 2003). Les molécules d'eau "relarguées" iront hydrater d'autres résidus, de surface ou dans des cavités protéiques, dont la polarité est plus grande. Ainsi, la grande majorité de l'entropie libérée pour le repliement de la protéine compensera la perte d'entropie de conformation de l'eau.

L'eau est donc intimement impliquée dans le repliement "guidé" de la protéine et est de plus en plus considérée, dans les prédictions de repliement structural (Papoian et al. 2004).

En fait, la stabilité des protéines, en condition physiologique, est le résultat d'un fin équilibre entre des changements d'enthalpie et d'entropie, stabilisant et déstabilisant la protéine et l'eau environnante. Pour des protéines globulaires, l'énergie libre de dénaturation ΔG_N^D a généralement un maximum entre 10 et 30°C, diminuant avec le froid ou la chaleur jusqu'au T_c et T_m . A ces températures, lorsque la variation d'énergie libre de dénaturation est de 0, il va arriver les phénomènes de dénaturation froide ou chaude, respectivement (Tsai et al. 2002). L'hydratation des groupements polaires est principalement responsable de la dénaturation froide car il est plus difficile d'hydrater et de faire des liaisons avec ces acides aminés quand la température est faible. La formation naturelle, à faible température, d'une eau "structurée", au voisinage de ces groupements polaires, va provoquer la déstabilisation de la protéine. Le coût entropique de la dénaturation est extrêmement réduit, car il n'y a plus les molécules d'eau environnantes structurées par les groupes polaires de surface : c'est une dénaturation dirigée par l'entropie (voir figure 4.22).

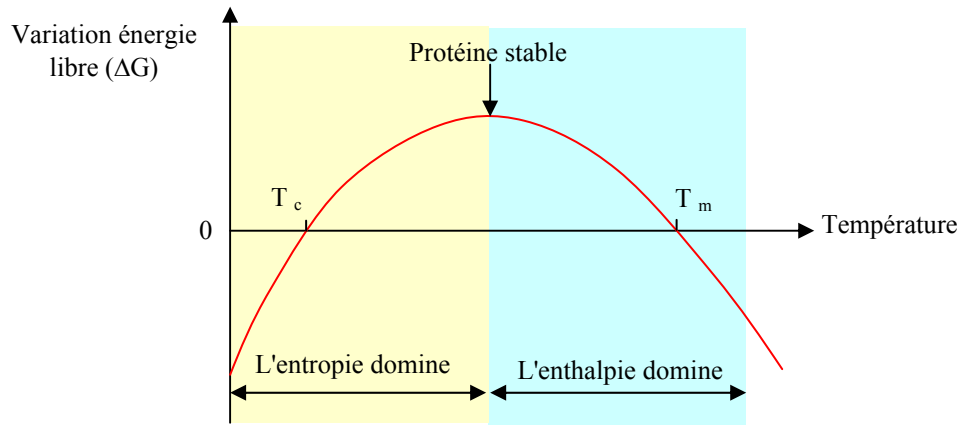


Figure 4.22 : Diagramme représentant les changements d'énergie libre ΔG lors de repliement et de dénaturation d'une protéine (adapté de (Klotz 1999)).

A des températures plus élevées, la déstabilisation de la protéine est dirigée principalement par des phénomènes enthalpiques. L'agitation thermique va conduire au dépliement de la protéine. Les zones hydrophobes seront exposées mais la chaîne principale dans son entier va pouvoir interagir avec les molécules d'eau extérieures en formant de nouvelles liaisons hydrogène. Autour de ces zones, vont se former des regroupements de molécules d'eau structurées pour ne pas interagir avec les groupements apolaires. C'est l'enthalpie qui est donc dominante quand la température est élevée.

4.3.11 Effets isotopiques du solvant sur la stabilité des protéines

Ainsi, l'énergie libre (ΔG_N^D) lors du repliement ou de la dénaturation d'une protéine est dirigée principalement par une combinaison de liaisons hydrogène et des interactions hydrophobes. L'influence du solvant sur la stabilité des protéines sera donc extrêmement liée à ces deux contributions. Lors de l'étude de dénaturation thermique par dichroïsme circulaire, l'eau légère et l'eau lourde ont été utilisées comme solvant de la γE cristalline et des différences de températures de dénaturation ont été déterminées. Pourtant, l' H_2O et le D_2O sont des molécules quasi identiques dans leur moment dipolaire, leur taille et leur longueur de liaison. Mais leur masse est différente : le deutérium a une masse deux fois plus grande que l'hydrogène. Il en résulte une fréquence de vibration plus faible dans les liaisons hydrogène avec le deutérium (correspondant à une énergie au point zéro de 1,3 kcal/mol). L'eau lourde a un plus grand degré de structure que l'eau légère à une température donnée et montre donc, une plus grande viscosité et de plus hautes chaleurs de vaporisation et de sublimation en comparaison avec l' H_2O (voir tableau 4.5).

	<i>Eau légère : H₂O</i>		<i>Eau lourde : D₂O</i>	
<i>Masse molaire (g.mol⁻¹)</i>	18,015268		20,027508	
<i>Point d'ébullition (en °C)</i>	100,0°C		101,42°C	
<i>Point de fusion (en °C)</i>	0,00°C		3,82°C	
<i>Viscosité à 25 °C (en mPa.s)</i>	0,8909		1,095	
<i>pK à 25 °C</i>	13,9991		14,87	
<i>Densité à 25 °C (kg.m⁻³)</i>	997,05		1104,36	
<i>Densité max. en kg.m⁻³ à T en °C</i>	999,972	3.984	1105.3	11.185

Tableau 4.5 : Caractéristiques physico-chimiques de l'eau légère et de l'eau lourde (Rukes 2001).

Pour les protéines, cette différence de propriétés entre l'eau légère et l'eau lourde conduit à une hydratation beaucoup plus forte dans le D₂O, donc à une plus grande affinité enthalpique D₂O-D₂O, et ainsi la solubilité des groupes apolaires devient plus faible, et les interactions hydrophobes en seront favorisées. Ainsi, la substitution de l'H₂O pour le D₂O conduit souvent à une faible augmentation de la température de transition (Antonino et al. 1991; Makhatadze et al. 1995; Tehei et al. 2001) et ceci est en accord avec nos valeurs.

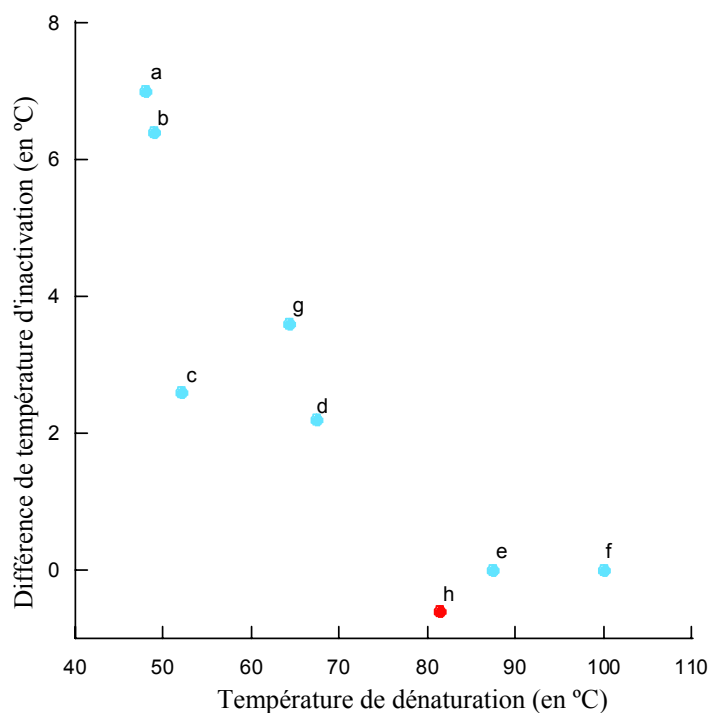
Nemethy et Sheraga (1964) ont précisé, que les liaisons deutériums intramoléculaires étaient vraisemblablement plus fortes et ont proposé que la plus grande stabilité des protéines était partiellement due à ces liaisons deutériums plus fortes. Mais, la seule évidence connue de nos jours, est que la substitution isotopique de l'hydrogène par le deutérium augmente la force de la liaison covalente (Benjamin et Benson 1962; Scheiner et Cuma 1996). Il n'est pas nécessairement évident que de plus fortes liaisons deutériums intramoléculaires vont conduire à une plus haute stabilité de la protéine en D₂O. Il existe beaucoup de possibilités pouvant expliquer cette légère augmentation de stabilité, parmi lesquelles il pourrait y avoir des changements d'énergie de désolvatation, des effets sur les liaisons hydrogène protéine protéine, des effets sur liaisons hydrogène protéine solvant et des effets sur les interactions solvant-solvant. Et comme il a été noté en figure 4.19, la stabilité nette d'une protéine est le résultat d'effets enthalpiques hydrophobes, d'interactions de van der Waals, et de liaisons hydrogène combinés aux effets entropiques de conformations de la chaîne polypeptidique. Donc ces possibilités ne sont pas à exclure.

Dans la bibliographie, quand, seul le solvant était considéré, il a souvent été pensé que l'effet hydrophobe était en majeure partie responsable de ce gain de stabilité (Nemethy et Sheraga 1964; Kresheck et al. 1965; Bonnete et al. 1994; Makhatadze et al. 1995). L'effet hydrophobe représente, à proprement dit, la faible solubilité d'un groupement apolaire dans l'eau. Donc la solubilité de ce groupement est encore plus faible sous l'effet du D₂O. Il est en effet possible que, la substitution isotopique de l'eau légère par l'eau lourde, augmente la force des liaisons deutériums dans les interactions solvant-solvant (Scheiner et Cuma 1996). Dans la suite logique, ces plus fortes interactions devraient alors, favoriser la formation du réseau de molécules d'eau que cela soit à basse ou à haute température, et défavoriser les interactions avec les groupements non polaires de la protéine. Il est aussi intéressant de remarquer que les températures de fusion (3,82°C) et d'ébullition (101,42°C) du D₂O sont plus hautes en comparaison du H₂O (voir tableau 4.5). La "déformation du réseau de glace" au niveau de la température de fusion, ainsi que le réseau à haute température, en seront donc plus difficiles car ceux-ci devraient être plus stables. Donc, à des températures identiques, ce gain entropique peut stabiliser le système de quelques kilo Joules. A un niveau macroscopique, sous l'effet hydrophobe, les polypeptides vont tendre à réduire leur surface de contact avec le solvant en adoptant une forme globulaire plus compacte ou en s'associant en de larges agrégats (observés lors de la concentration en D₂O de nos échantillons). Il est aussi intéressant de noter que le D₂O augmente aussi la rigidité de la structure native des protéines (Cioni et Strambini 2002), effet généralement augmenté avec la température, pouvant expliquer l'état replié plus stable d'une protéine dans le D₂O.

4.3.12 Effet de la perdeutériation sur la stabilité

Dans certains cas, une perdeutériation de la protéine n'influe pas non plus sur la stabilité de cette dernière (voir figure 4.23). Il semblerait que la stabilité des protéines dites stables, dont la température de demi dénaturation est élevée : l'alkaline phosphatase de *E. coli*, la ribonucléase de *F. diplosiphon* (Rokop et al. 1969) et γ E cristalline de *R. norvegicus*, ne soit pas ou que très faiblement affectée par la perdeutériation. En revanche les protéines avec une faible température de dénaturation : la phycocyanine de *P. calothricoides*, la phycocyanine de *P. luridum*, la Glutathione S transférase de *S. japonicum*, la phycocyanine de *S. lividus* (Hattori et al. 1965b) ou le cytochrome P₄₅₀+camphore de *Pseudomonas putida* (Meilleur et al. 2004), sont profondément déstabilisées par la perdeutériation.

Figure 4.23 : Effets de la deutériation sur la dénaturation de quelques protéines (Brockwell et al. 2001). La différence de température d'inactivation entre les formes hydrogénées et deutériées sont positionnées en fonction de la température d'inactivation de la forme hydrogénée de la protéine. (a) Phycocyanine de *P. calothricoides*, (b) phycocyanine de *P. luridum*, (c) GST de *S. japonicum*, (d) phycocyanine de *S. lividus* (Hattori et al. 1965b), (e) alcaline phosphatase de *E.coli*, (f) ribonucléase de *F. diplosiphon* (Rokop et al. 1969) et (g) cytochrome P₄₅₀+camphore (Meilleur et al. 2004) effectué en D₂O. La protéine (h) γ E cristalline de *R. norvegicus* est représentée suivant la perte de structure secondaire en DC. La forme deutériée étant plus stable, la différence est négative.



La première chose intéressante, est de se rendre compte que la perdeutériation a faiblement affecté la stabilité de la γ E cristalline. Mais cette perturbation, qui est très faible (0,6°C dans l'eau légère, et 1,3°C dans l'eau lourde) est toutefois à l'opposé des précédents résultats : la deutériation de la protéine n'a pas déstabilisé mais a augmenté la température de demi-dénaturation de quelques degrés, elle a rendu la protéine plus stable.

La perdeutériation signifie la substitution des hydrogènes liés aux carbones par des deutériums. Cet échange devrait donc concerner essentiellement les interactions de type hydrophobe entre les acides aminés. Comme il a été précisé auparavant, l'énergie au point zéro de la liaison C–D, qui est plus faible en comparaison de la liaison C–H, est accompagnée par une diminution de l'amplitude de vibration. Les contraintes stériques devraient être, de ce fait, moindres pour une protéine deutériée, entraînant une certaine forme de compaction. Dans le cas des γ cristallines, celles-ci semblent être légèrement plus stables avec cette compaction. Avec la perdeutériation, il est arrivé que la formation de complexe actif soit diminuée (Hattori et al. 1965b; Brockwell et al. 2001). Si la forme plus compacte de la protéine ne permet pas la formation de ce complexe, la protéine peut s'en trouver déstabilisée donnant un sens à la diminution d'énergie libre de dénaturation de ces protéines.

Dans le cas des γ cristallines, de nombreuses d'interactions maintiennent les deux domaines (voir chapitre 6) en formant un cœur hydrophobe entouré de charges, ainsi que le

peptide de liaison qui favorise le rapprochement de ces domaines (Sergeev et al. 1988). Si, comme Hattori et ses confrères l'ont décrit (Hattori et al. 1965a), la deutériation conduit à la diminution, d'un point de vue stérique, des chaînes latérales ; alors il est fort probable que cette réduction entraîne la formation d'interactions intramoléculaires de type hydrophobe ou/et électrostatique entre les deux domaines stabilisant la structure. Et ceci ne doit pas être un facteur défavorable puisque les γ cristallines semblent avoir besoin de cette compaction au sein même de la cellule. Bien que la protéine γ E cristalline soit un cas exceptionnel, du fait de son gain de stabilité par la perdeutériation, il ne faut pas, non plus tomber dans l'excès et oublier que cette stabilité n'est que très minime (que quelques 0,6°C), comparée à la déstabilisation de certaines protéines telles que la phycocyanine ou le cytochrome P₄₅₀ (jusqu'à 7°C).

Cette étude en dichroïsme circulaire a démontré que la perdeutériation n'affecte pas à température ambiante les structures secondaires γ cristallines, que la haute stabilité des cristallines est maintenue en solvant deutérié (et même légèrement augmentée), et que la perdeutériation n'affecte pas drastiquement la stabilité de la protéine. Mais tout système protéique est différent et la perdeutériation et le solvant peuvent agir à de nombreux niveaux ne serait ce que pour la stabilité de la protéine, il est donc important de ne pas généraliser les très faibles effets de la perdeutériation, observés pour la γ E cristalline, à toutes les protéines.

4.4 Bilans

Dans ce chapitre, la purification de la la γ E cristalline sous forme hydrogénée et deutériée a été décrite. De manière à avoir un échantillon homogène la méthionine amino peptidase a été surexprimée avec la cristalline. Le niveau d'incorporation, vérifié par spectrométrie de masse, du deutérium est de 98%.

Les différentes études biochimiques n'ont pas permis de visualiser de réelles différences entre les formes hydrogénées et deutériées. Les deux formes possèdent un comportement migratoire similaire en gel natif et un point isoélectrique identique. La protéolyse limitée a montré que la très haute résistance des γ cristallines vis à vis des protéases est toujours présente malgré la deutériation. Bien que cette dernière étude a pu montré une stabilité légèrement plus importante pour la protéine hydrogénée, ces observations

restent tout de même extrêmement faibles puisqu'elles ne montrent une différence de protéolyse uniquement après 20 heures d'incubation ou en conditions semi-dénaturantes.

Les deux analyses spectroscopiques ont fourni des résultats similaires sur la perdeutériation. La résonance infrarouge par transformée de Fourier a montré l'influence que peut avoir la perdeutériation sur les spectres infrarouges mais cette étude a permis de poser des hypothèses et non des affirmations sur d'éventuels changements de structures secondaires. En revanche, l'analyse par dichroïsme circulaire, a démontré que la stabilité de la γE cristalline n'est que très faiblement perturbée et que la structure secondaire ne l'est pas plus.

Ces études, en accord avec la littérature, montrent que la protéine recombinante : la γE cristalline de rat, ne subit pas de changement drastique, d'un point de vue biochimique, suite à la perdeutériation et/ou au changement de solvant. Les propriétés physico-chimiques peuvent être légèrement différentes, mais celles ci restent toujours très proches.

Bibliographie :

- Antonino, L.C., Kautz, R.A., Nakano, T., Fox, R.O., and Fink, A.L. 1991. Cold denaturation and 2H₂O stabilization of a staphylococcal nuclease mutant. *Proc Natl Acad Sci U S A* **88**: 7715-7718.
- Ben-Bassat, A., Bauer, K., Chang, S.Y., Myambo, K., Boosman, A., and Chang, S. 1987. Processing of the initiation methionine from proteins: properties of the *Escherichia coli* methionine aminopeptidase and its gene structure. *J Bacteriol* **169**: 751-757.
- Benjamin, L., and Benson, G. 1962. A deuterium isotope effect on excess enthalpy of methanol-water solutions. *J. Phys. Chem.*: 858-861.
- Blundell, T., Lindley, P., Miller, L., Moss, D., Slingsby, C., Tickle, I., Turnell, B., and Wistow, G. 1981. The molecular structure and stability of the eye lens: x-ray analysis of gamma-crystallin II. *Nature* **289**: 771-777.
- Bonnete, F., Madern, D., and Zaccari, G. 1994. Stability against denaturation mechanisms in halophilic malate dehydrogenase "adapt" to solvent conditions. *J Mol Biol* **244**: 436-447.
- Brockwell, D., Yu, L., Cooper, S., McClelland, S., Cooper, A., Attwood, D., Gaskell, S.J., and Barber, J. 2001. Physicochemical consequences of the perdeuteration of glutathione S-transferase from *S. japonicum*. *Protein Sci* **10**: 572-580.
- Cioni, P., and Strambini, G.B. 2002. Effect of heavy water on protein flexibility. *Biophys J* **82**: 3246-3253.
- Coligan, J., Dunn, B., Speicher, D., and Wingfield. 2003. *Current protocol in protein sciences*, John Wiley and Sons, Inc ed.
- David, L.L., and Shearer, T.R. 1986. Purification of calpain II from rat lens and determination of endogenous substrates. *Exp Eye Res* **42**: 227-238.
- Fernandez, A., Kardos, J., and Goto, Y. 2003. Protein folding: could hydrophobic collapse be coupled with hydrogen-bond formation? *FEBS Lett* **536**: 187-192.
- Fu, L., and Liang, J.J. 2002. Conformational change and destabilization of cataract gammaC-crystallin T5P mutant. *FEBS Lett* **513**: 213-216.
- Hallez, R. 1997. Etude comparative de différents méthodes de broyage cellulaire. In *Travail de fin d'études à la haute école provinciale du Hainaut Occidental*.
- Hattori, A., Crespi, H.L., and Katz, J.J. 1965a. Association and dissociation of Phycocyanin and the effects of deuterium substitution on the processes. *Biochemistry* **4**: 1225-1238.
- Hattori, A., Crespi, H.L., and Katz, J.J. 1965b. Effect of side-chain deuteration on protein stability. *Biochemistry* **4**: 1213-1225.
- Hellwig, P., Pfitzner, U., Behr, J., Rost, B., Pesavento, R.P., Donk, W.V., Gennis, R.B., Michel, H., Ludwig, B., and Mantele, W. 2002. Vibrational modes of tyrosines in cytochrome c oxidase from *Paracoccus denitrificans*: FTIR and electrochemical studies on Tyr-D4-labeled and on Tyr280His and Tyr35Phe mutant enzymes. *Biochemistry* **41**: 9116-9125.
- Hirel, P.H., Schmitter, M.J., Dessen, P., Fayat, G., and Blanquet, S. 1989. Extent of N-terminal methionine excision from *Escherichia coli* proteins is governed by the side-chain length of the penultimate amino acid. *Proc Natl Acad Sci U S A* **86**: 8247-8251.
- Jung, C. 2000. Insight into protein structure and protein-ligand recognition by Fourier transform infrared spectroscopy. *J Mol Recognit* **13**: 325-351.
- Klotz, I.M. 1999. Parallel Change with Temperature of Water Structure and Protein Behavior. *J. Phys. Chem. B* **103**: 5910-5916.
- Kresheck, G.C., Schneider, H., and Scheraga, H.A. 1965. The effect of D₂O on the thermal stability of proteins. Thermodynamic parameters for the transfer of model compounds from H₂O to D₂O. *J Phys Chem* **69**: 3132-3144.
- Lamba, O.P., Borchman, D., Sinha, S.K., Shah, J., Renugopalakrishnan, V., and Yappert, M.C. 1993. Estimation of the secondary structure and conformation of bovine lens crystallins by infrared spectroscopy: quantitative analysis and resolution by Fourier self-deconvolution and curve fit. *Biochim Biophys Acta* **1163**: 113-123.
- Lutter, R., Abrahams, J.P., van Raaij, M.J., Todd, R.J., Lundqvist, T., Buchanan, S.K., Leslie, A.G., and Walker, J.E. 1993. Crystallization of F1-ATPase from bovine heart mitochondria. *J Mol Biol* **229**: 787-790.
- Makhatadze, G.I., Clore, G.M., and Gronenborn, A.M. 1995. Solvent isotope effect and protein stability. *Nat Struct Biol* **2**: 852-855.
- Mandal, K., Bose, S.K., Chakrabarti, B., and Siezen, R.J. 1985. Structure and stability of gamma-crystallins. I. Spectroscopic

- evaluation of secondary and tertiary structure in solution. *Biochim Biophys Acta* **832**: 156-164.
- Mandal, K., Chakrabarti, B., Thomson, J., and Siezen, R.J. 1987. Structure and stability of gamma-crystallins. Denaturation and proteolysis behavior. *J Biol Chem* **262**: 8096-8102.
- Meilleur, F., Contzen, J., Myles, D.A., and Jung, C. 2004. Structural stability and dynamics of hydrogenated and perdeuterated cytochrome P450cam (CYP101). *Biochemistry* **43**: 8744-8753.
- Nemethy, G., and Sheraga, H. 1964. Structure of water and hydrophobic bonding in proteins: thermodynamic properties of liquid deuterium oxide. *J. Chem. Phys.* **41**: 680-689.
- Papoian, G.A., Ulander, J., Eastwood, M.P., Luthey-Schulten, Z., and Wolynes, P.G. 2004. Water in protein structure prediction. *Proc Natl Acad Sci U S A* **101**: 3352-3357.
- Rokop, S., Gajda, L., Parmerter, S., Crespi, H.L., and Katz, J.J. 1969. Purification and characterization of fully deuterated enzymes. *Biochim Biophys Acta* **191**: 707-715.
- Rukes, B. 2001. Guideline on the Use of Fundamental Physical Constants and Basic Constants of Water. *International Association for the Properties of Water and Steam*.
- Sambrook, and Russell. 1998. Molecular Cloning a laboratory manual. *CSHL Press* **1**: 1-105-110.
- Scheiner, S., and Cuma, M. 1996. Relative stability of hydrogen and deuterium bonds. *J. Am. Chem. Soc.* **118**: 1511-1521.
- Schowen, K.B., and Schowen, R.L. 1982. Solvent isotope effects of enzyme systems. *Methods Enzymol* **87**: 551-606.
- Sen, A.C., Walsh, M.T., and Chakrabarti, B. 1992. An insight into domain structures and thermal stability of gamma-crystallins. *J Biol Chem* **267**: 11898-11907.
- Sergeev, Y.V., Chirgadze, Y.N., Mylvaganam, S.E., Driessen, H., Slingsby, C., and Blundell, T.L. 1988. Surface interactions of gamma-crystallins in the crystal medium in relation to their association in the eye lens. *Proteins* **4**: 137-147.
- Siezen, R.J., and Argos, P. 1983. Structural homology of lens crystallins. III. Secondary structure estimation from circular dichroism and prediction from amino acid sequences. *Biochim Biophys Acta* **748**: 56-67.
- Siezen, R.J., Wu, E., Kaplan, E.D., Thomson, J.A., and Benedek, G.B. 1988. Rat lens gamma-crystallins. Characterization of the six gene products and their spatial and temporal distribution resulting from differential synthesis. *J Mol Biol* **199**: 475-490.
- Stuart, B. 1997. *Biological applications of infrared spectroscopy*, Greenwich.
- Tehei, M., Madern, D., Pfister, C., and Zaccai, G. 2001. Fast dynamics of halophilic malate dehydrogenase and BSA measured by neutron scattering under various solvent conditions influencing protein stability. *Proc Natl Acad Sci U S A* **98**: 14356-14361.
- Tsai, C.J., Maizel, J.V., Jr., and Nussinov, R. 2002. The hydrophobic effect: a new insight from cold denaturation and a two-state water structure. *Crit Rev Biochem Mol Biol* **37**: 55-69.
- Wade, D. 1999. Deuterium isotope effects on noncovalent interactions between molecules. *Chem Biol Interact* **117**: 191-217.
- Wistow, G., Turnell, B., Summers, L., Slingsby, C., Moss, D., Miller, L., Lindley, P., and Blundell, T. 1983. X-ray analysis of the eye lens protein gamma-II crystallin at 1.9 Å resolution. *J Mol Biol* **170**: 175-202.

Chapitre 5

La cristallographie aux rayons X : rappels, collectes et traitements

Abstract

X-ray crystallography is one of the rare techniques which allows detailed structural information of biological macromolecules to be obtained. It is the most common technique used to get atomic information even for large macromolecular assemblies. Other techniques to obtain molecular structures are NMR, which is generally limited to macromolecules smaller than 40 kDa, electron microscopy, which gives electron density with a resolution rarely exceeding 6 Å, and finally electron crystallography, which is limited to two dimensional crystals and gives a resolution of 3.5 Å in the best direction. X-ray crystallography at high resolution (1-2.5 Å) makes it possible to visualise atomic positions averaged in time and over the unit cells of a crystal.

After some reminders concerning X-ray crystallography, this chapter presents the different crystallisation conditions required to obtain crystals of γ E-crystallin. These conditions were obtained using the vapour diffusion technique and especially the hanging drop method. The cryo-conditions, necessary to protect the crystal against harmful effect of X-ray radiation, were determined to avoid the formation of crystalline ice in the crystal. Hydrogenated and deuterated crystals in hydrogenated and deuterated solvent conditions were used to collect diffraction patterns on the beam lines ID14-EH2 and EH3 of the ESRF. Recorded images were treated with software *Mosflm* and *XDS* for indexing and integrating intensities.

The molecular replacement technique was used to “transfer” phase information from the known structure at 2.3 Å to our crystalline forms. Then, the *O* and *Xtalview* programs were used to build the molecular model inside the electronic density map, and *Refmac* and *ShelX* programs served to refinement the structural models. They have been validated by the program *Procheck*.

5.1 La cristallographie aux rayons X

La cristallographie aux rayons X est une des rares techniques qui permettent d'obtenir une information structurale détaillée sur des macromolécules biologiques. Les modèles moléculaires obtenus à partir de la cristallographie aux rayons X permettent dans de nombreux cas de révéler certains détails moléculaires des processus de la vie. Les scientifiques utilisent ces modèles pour comprendre et apprendre comment les molécules fonctionnent : comment les enzymes catalysent des réactions métaboliques, comment certaines protéines lient l'ADN, comment les anticorps se fixent et activent les défenses immunitaires, comment certains cofacteurs (même une simple molécule d'eau) sont essentiels au bon déroulement catalytique d'une réaction. La cristallographie aux rayons X est la technique la plus utilisée pour acquérir une information atomique, même sur des assemblages macromoléculaires très grands. Les autres techniques pour obtenir des structures moléculaires sont la RMN, qui est limitée généralement par des tailles d'assemblages macromoléculaires de 150 résidus, la microscopie électronique qui donne des densités électroniques à des résolutions n'excédant pas 6 Å, et la cristallographie électronique qui est limitée à de cristaux en 2 dimensions et donne des résolutions de 3,5 Å dans la meilleure direction. La cristallographie à haute résolution (1 à 2,5 Å) aux rayons X permet de visualiser les sites atomiques moyennés dans le temps et sur l'ensemble des mailles d'un cristal.

Dans ce chapitre, la cristallisation, dans des conditions similaires, de la protéine recombinante γ E cristalline sous forme hydrogénée et deutériée en solvant H₂O et D₂O est développée. Au cours de ma thèse, j'ai obtenu différentes formes cristallines mais une seule a été retenue pour envisager des expériences de diffraction. Le choix de cette condition a été motivé par la reproductibilité de cristallisation et la taille des cristaux obtenus pour envisager de la cristallographie aux neutrons.

Pour réaliser de la diffraction aux rayons X qui cause des dommages radiatifs, ces cristaux ont été préparés et testés avec de nombreux cryoprotecteurs et de différentes manières. De façon à homogénéiser les résultats, un seul cryoprotecteur, le plus approprié, a été utilisé pour les quatre protéines pour congeler les cristaux.

Après quelques brefs rappels sur la cristallographie aux rayons X, ce chapitre décrit les différentes étapes et méthodes que j'ai utilisées pour résoudre les quatre structures de la γ E cristalline : de la collecte de données, à l'obtention des modèles, en passant par l'utilisation du remplacement moléculaire pour résoudre le problème de phase.

5.2 Rappels sur la cristallographie aux rayons X

5.2.1 Principe de la diffusion ou de la diffraction

La diffusion élastique (c'est à dire sans perte d'énergie) est produite lorsqu'un objet reçoit une onde électromagnétique (comme les rayons X), plane ou sphérique, et que cet objet ré-émet (ou diffuse) cette onde dans tout l'espace, sous la forme d'une onde sphérique sans changement de phase (préservation de la cohérence).

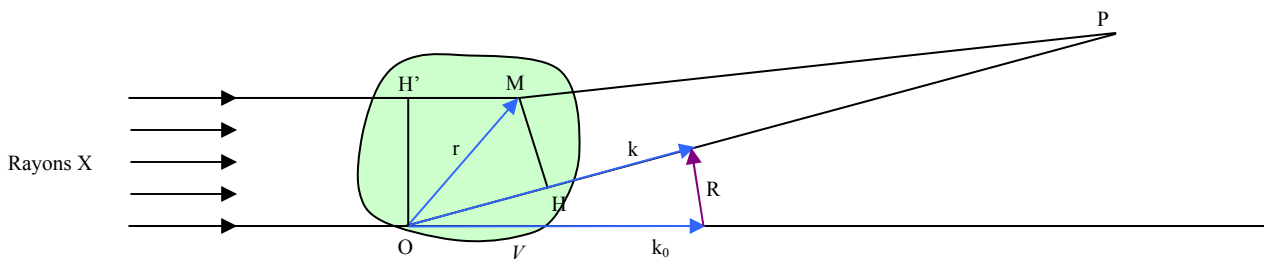


Figure 5.1 : Interaction d'une onde électromagnétique comme les rayons X avec un objet de volume V. L'objet va ré-émettre cette onde dans toutes les directions (dont celle du point P).

Si un petit volume δv de cet objet est irradié par l'onde incidente de vecteur d'onde k_0 , ce volume va diffuser l'onde incidente dans toutes les directions (dont celle du point P) sous la forme d'une onde sphérique de vecteur d'onde k . Pour simplifier, il y aura $s = k / 2\pi$ et $s = 1/\lambda$.

$$s = \frac{k}{2\pi} = \frac{1}{\lambda} \quad (5.1)$$

Dans le cadre de diffusion élastique $k_0 = k$ et le vecteur R de diffusion de l'onde, décrivant la direction d'observation, est défini de la manière suivante :

$$\vec{R} = \vec{s} - \vec{s}_0 \quad \text{soit} \quad R = \frac{2 \sin \theta}{\lambda} \quad (5.2)$$

La phase ϕ de l'onde émise par le volume δv en P peut alors se calculer de façon simple. La différence δ de chemin est calculée entre l'onde passante par l'origine O et celle passant par un point M (au centre du volume δv) de l'objet et définie par $r = OM$:

$$\delta = OH - H'M = r \cdot s \cdot \lambda - r \cdot s_0 \cdot \lambda_0 = r \cdot R \cdot \lambda \quad \text{où la différence entre MP et HP est négligée.}$$

La phase ϕ vaut donc

$$\phi = \frac{2\pi\delta}{\lambda} = 2\pi \cdot r \cdot R \quad (5.3)$$

Si $f(r)$ est le pouvoir diffusant de cet objet au point r , l'onde émise par le volume dv peut donc s'écrire :

$$dF(R) = \frac{F_0}{r'} f(r) \exp(i\phi) \cdot dv_r = \frac{F_0}{r'} f(r) \exp(2i\pi \cdot rR) \cdot dv_r \quad (5.4)$$

L'onde émise par l'objet est la somme des ondes émises par tous les petits volumes δv de l'objet :

$$F(R) = \int_v \frac{F_0}{r'} f(r) \exp(2i\pi \cdot rR) \cdot dv \quad (5.5)$$

L'onde diffusée $F(R)$ est donc la transformée de Fourier du pouvoir diffusant $f(r)$ de l'objet étudié :

$$F(R) = \int_v \rho(r) \exp(2i\pi \cdot rR) \cdot dv_r \quad (5.6)$$

Dans le cas des rayons X, le pouvoir diffusant est proportionnel à la densité électronique $\rho(r)$ des atomes, et dans le cas des neutrons, ce pouvoir diffusant est proportionnel à la densité de longueur d'onde de diffusion des noyaux. Les densités électroniques résultantes pour les rayons X seront donc :

$$\rho(r) = \int_v F(R) \exp(-2i\pi \cdot rR) \cdot dv_R \quad (5.7)$$

5.2.2 Principe de diffraction des rayons X par un cristal

5.2.2.1 Un cristal

Un cristal est un solide ordonné où toutes les molécules qui le composent sont disposées de façon régulière dans les trois directions. Il peut être décrit par un volume élémentaire appelé maille élémentaire, de forme parallélépipédique, qui s'empile dans les trois dimensions. Les positions de tous les atomes du cristal se déduisent des positions des atomes de la maille élémentaire. Un cristal est décrit par un réseau direct qui définit la périodicité tridimensionnelle. Le réseau est repéré par une origine O et trois vecteurs \vec{a} , \vec{b} et \vec{c} , qui définissent la maille élémentaire. Il existe sept systèmes cristallins capables de quadriller l'espace en trois dimensions de façon périodique (voir tableau 2.1)

Système cristallin	Type de Bravais	Caractéristiques	
Triclinique	P	$a \neq b \neq c$	$\alpha \neq \beta \neq \gamma \neq 90^\circ$
Monoclinique	P, C	$a \neq b \neq c$	$\alpha = \gamma = 90^\circ \quad \beta \neq 90^\circ$
Orthorhombique	P, C, I, F	$a \neq b \neq c$	$\alpha = \beta = \gamma = 90^\circ$
Quadratique (tétragonal)	P, I	$a = b \neq c$	$\alpha = \beta = \gamma = 90^\circ$
Rhomboédrique	R	$a = b = c$	$\alpha = \beta = \gamma \neq 90^\circ$
Hexagonal	P	$a = b \neq c$	$\alpha = \beta = 90^\circ \quad \gamma = 120^\circ$
Cubique	P, I, F	$a = b = c$	$\alpha = \beta = \gamma = 90^\circ$

Tableau 2.1 : Les 7 systèmes cristallins capables de paver l'espace en 3 dimensions.

Un cristal peut être décrit par une maille simple qui contient un nœud du réseau par maille. Elle est appelée maille primitive P. Une maille multiple contient plusieurs nœuds. La maille multiple est choisie pour maille élémentaire quand celle-ci possède une plus grande symétrie. Trois types de mailles multiples existent et sont décrits par les symétries suivantes :

- Deux faces centrées avec deux nœuds par mailles, appelée C, B ou A ;
- Un corps centré avec deux nœuds par maille, appelée maille I
- Toutes les faces centrées avec quatre nœuds par maille, appelée F.

Par les différentes mailles, 14 réseaux de Bravais ont été représentés (voir tableau 2.1).

Un cristal et sa maille peuvent être soumis à une ou à plusieurs symétries. Cette symétrie est décrite par le groupe d'espace. Il existe 230 groupes d'espaces possibles pour décrire la symétrie d'un cristal. Seules, une soixantaine sont possibles pour les protéines car ce sont des molécules chirales (L acides aminés). Les seules symétries possibles pour les protéines sont les axes de rotation et les axes hélicoïdaux. Le groupe d'espace trouvé pour l'ensemble des structures résolues dans cette thèse est un orthorhombique en $P2_12_12_1$, ce qui définit trois axes d'ordre 2 avec une translation de $1/2$ le long de ces axes.

5.2.2.2 Principe de la diffraction par un cristal

La diffraction est considérée lorsqu'il y a diffusion par un solide périodique comme un cristal. Elle n'a lieu que si la longueur d'onde est significativement inférieure à la période spatiale de l'objet.

Ce phénomène de diffraction est un phénomène d'interférence entre les ondes électromagnétiques (rayons X) diffusées par tous les atomes du cristal. Du fait de la périodicité 3D du cristal, la somme des ondes diffusées par chacune des mailles formera des interférences constructives que dans certaines directions : c'est un pic de diffraction. Dans les autres directions,

le rayonnement diffusé sera très faible (interférence destructive due à la somme d'un très grand nombre d'ondes diffusées, avec des déphasages aléatoires).

5.2.2.3 La loi de Bragg

Dans ce cas, la diffraction d'un cristal dans une direction donnée peut s'interpréter comme la réflexion du faisceau incident sur une famille de plans réticulaires du cristal (voir figure 5.2).

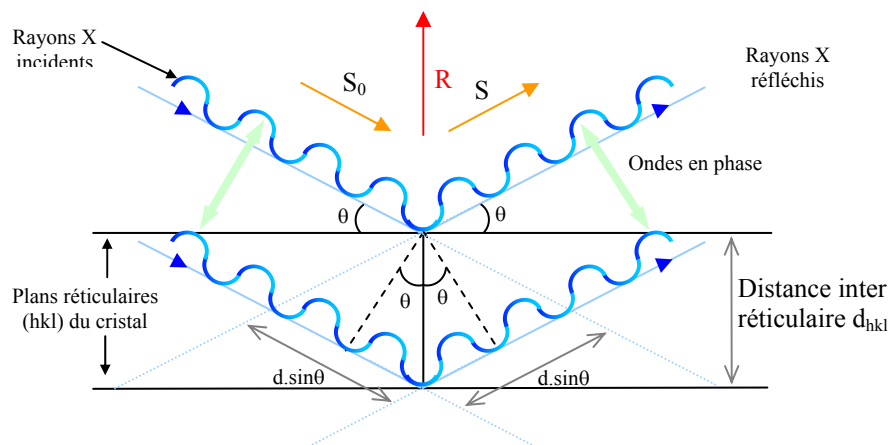


Figure 5.2 : Représentation graphique de la loi de Bragg. Les rayons incidents sont réfléchis suivant les plans réticulaires du cristal. Lorsqu'il y a une tâche de diffraction, c'est qu'il y a interférences constructives et donc que les ondes sont en phase à la sortie du cristal malgré une différence de marche de $2d\sin\theta$.

La différence de marche entre les deux faisceaux est $2 d \sin \theta$, d étant la distance entre deux plans réticulaires de la même famille, θ l'angle d'incidence du faisceau sur la famille de plans. La condition pour avoir une interférence constructive est que la différence de marche entre les deux faisceaux soit égale à un nombre entier n de longueur d'onde, d'où la loi de Bragg :

$$2.d.\sin \theta = n.\lambda \quad (5.8)$$

5.2.2.4 Notion d'espace réciproque et de réseau réciproque

Dans les équations suivantes, les densités électroniques et l'onde diffractée sont reliées par les relations de transformée de Fourier.

$$\vec{F}(R) = \int_{\mathcal{V}} \rho(r) \exp(2i\pi.r.R).dv_r \quad \text{et} \quad \rho(r) = \int_{\mathcal{V}} \vec{F}(R) \exp(-2i\pi.r.R).dv_R \quad (5.9)$$

Dans ce cas, l'espace direct des positions (définis par r) est associé à un espace réciproque de vecteur d'onde diffusée défini par R . Cet espace réciproque est l'espace dans lequel sont repérées les raies de diffraction.

Le réseau réciproque est associé au réseau direct, il est repéré par trois vecteurs de base a^* , b^* et c^* reliés à a , b et c : les vecteurs de bases du réseau direct (voir figure 5.3).

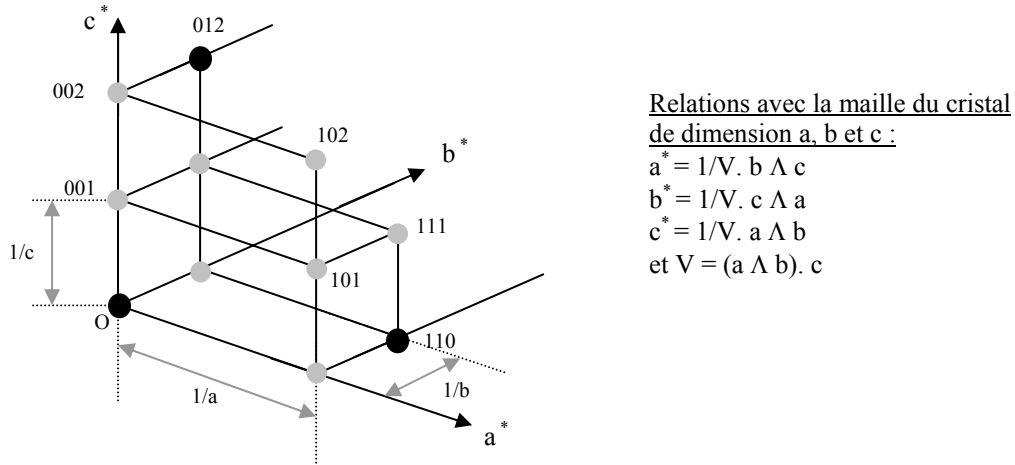


Figure 5.3 : Représentation de l'espace réciproque et des relations avec l'espace réel. Une petite maille cristalline va engendrer un grand espace réciproque.

Comme la périodicité du cristal va induire la présence d'une onde diffractée (interférence constructive) uniquement dans certaines conditions privilégiées, ces directions privilégiées seront définies par :

$$R = ha^* + kb^* + lc^* \text{ avec } h, k \text{ et } l \text{ entiers} \tag{5.10}$$

Ainsi, à chaque nœud du réseau réciproque repéré par les indices hkl correspondra une tâche de diffraction.

Quand on reconsidère la loi de Bragg avec une onde diffractée par la famille de plan réticulaire (hkl), le vecteur de diffusion $R = s - s_0$ est perpendiculaire au plan réticulaire (voir figure 5.2). La distance d_{hkl} entre deux plans de la famille de plans (hkl) est donnée par :

$$\frac{1}{d_{hkl}} = |ha^* + kb^* + lc^*| \tag{5.11}$$

et la loi de Bragg devient : $2.d_{hkl} \cdot \sin \theta = n.\lambda$ ou encore $d_{hkl} = \frac{\lambda}{2 \cdot \sin \theta}$ (5.12)

où d_{hkl} est la résolution de la réflexion hkl .

5.2.2.5 La sphère d'Ewald

Un modèle géométrique permet la compréhension des conditions nécessaires lorsqu'il y a réflexion sur un détecteur plan tel que ceux utilisés sur les lignes de lumière de l'ESRF.

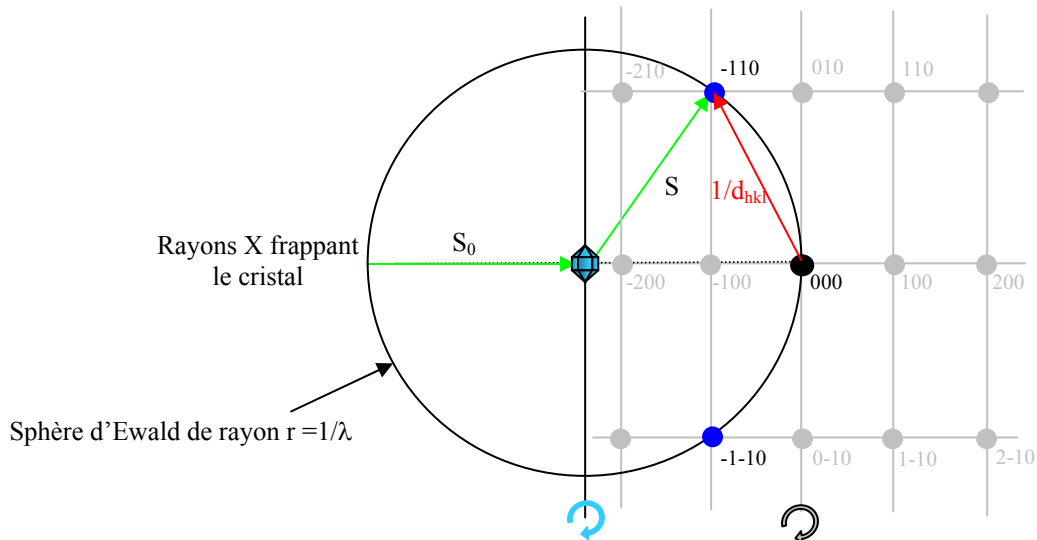


Figure 5.4 : Construction d'Ewald. La sphère d'Ewald a le cristal pour centre et $1/\lambda$ pour rayon. Les rayons X (S_0) vont être diffusés suivant le vecteur S . Le vecteur $1/d_{hkl}$ décrit la position d'une réflexion dans l'espace réciproque. Quand le cristal va être mis en rotation, il n'y aura que là où la sphère coupe un nœud du réseau réciproque (représenté en gris clair et ayant pour origine, l'intersection avec le faisceau non diffracté avec la sphère). Les nœuds 101 et 10-1 sont des tâches de réflexion qui apparaissent sur le détecteur plan.

La sphère a pour origine le cristal et pour rayon $1/\lambda$. L'origine de l'espace réciproque est positionnée à l'intersection du faisceau non diffracté et la sphère d'Ewald. La construction est la représentation géométrique des conditions de diffraction en accord avec la loi de Bragg. A chaque nœud du réseau est associé une tâche de diffraction que si ce dernier se trouve sur la sphère d'Ewald. Pendant une collecte de données, le cristal va tourner sur un axe perpendiculaire aux rayons de façon à ce que chaque nœud coupe la sphère et donne une réflexion.

v) Propriétés des spectres de diffraction :

- Sur un spectre de diffraction, seules l'intensité de l'onde diffractée et la position des réflexions sont mesurées. Et l'intensité de l'onde correspond au carré de l'amplitude de l'onde :

$$I(\mathbf{R}) = F(\mathbf{R}) \cdot F^*(\mathbf{R}) = F(\mathbf{R})^2 \quad (5.14)$$

Pour calculer les densités électroniques, c'est à dire calculer la transformée de Fourier des ondes diffractées, il faudra retrouver les phases qui correspondent à ces amplitudes.

- L'origine du spectre est le faisceau non diffracté.

- La position des tâches de diffraction ne dépend que des dimensions de la maille et de l'orientation du cristal.
- L'intensité diffractée dépend du contenu de la maille (nombre de maille et leurs composants).

5.2.3 Traitements des données

5.2.3.1 Des images de diffraction aux données de diffractions

Le traitement des images de diffraction consiste en trois étapes distinctes :

- la détermination précise des paramètres du cristal, du groupe d'espace et l'orientation du cristal.
- l'intégration des images avec les précédents paramètres du cristal, ceux de la ligne et du détecteur.
- la mise à l'échelle des données.

Quelques étapes sont nécessaires pour réaliser ceci :

i) L'indexation :

Les images numérisées des clichés de diffraction vont être utilisées dans un premier temps pour déterminer l'orientation et la maille du cristal. En fonction des paramètres de la machine (la longueur d'onde λ , la distance cristal détecteur, l'orientation du détecteur,...) et de la position des tâches de diffractions (en x , y sur le détecteur et en Φ , axe de rotation du cristal), les logiciels déterminent l'orientation et la maille primitive du cristal. Cette procédure permet d'associer à chaque tâche de diffraction ses coordonnées (indices de Miller hkl) dans l'espace réciproque (voir figure 5.4).

ii) L'intégration :

Cette étape consiste à déterminer l'intensité de chaque tâche de diffraction, en sommant l'intensité reçue par chaque pixel du détecteur appartenant à une même tâche. A la fin de l'intégration, M mesures d'intensité de tâches de diffraction vont être obtenues. Ces mesures M vont correspondre au volume de l'espace réciproque balayé pendant l'enregistrement. Parmi ces mesures, certaines sont redondantes, dues à la loi de Friedel (centro symétrie) et aux symétries éventuelles du cristal (d'où la nécessité de déterminer la classe de Laue du cristal, en chapitre 6). La prise en compte de ces symétries permet de réduire les données, en calculant pour chaque hkl unique, la valeur d'intensité moyenne. Ainsi, un certain nombre N de réflexions uniques, seront acquises avec une redondance donnée par le rapport M/N . Le nombre de réflexions mesurées va dépendre de la rotation effectuée pendant l'enregistrement, de la maille, de la résolution et de la

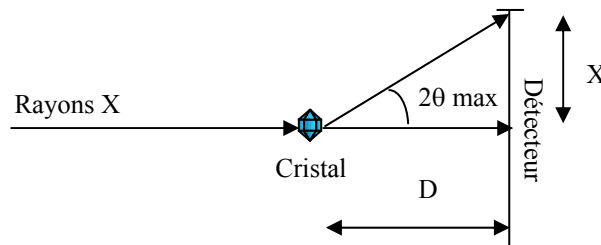
symétrie du cristal. Le nombre de réflexions théoriquement possibles à une résolution d donnée, est le nombre de nœuds du réseau réciproque compris dans une sphère de rayon $1/d$, c'est à dire :

$$M = \frac{4\pi}{3d^3V^*} = \frac{4\pi V}{3d^3} \quad \text{et} \quad N = \frac{M}{\text{nombre d'équivalents}} \quad (5.15)$$

où V est le volume de la maille, V^* est le volume de la maille réciproque.

En revanche, la résolution est limitée par plusieurs facteurs :

- la surface du détecteur, qui détermine, en fonction de la longueur d'onde et de la distance cristal détecteur, l'angle 2θ maximum (voir figure 5.5).



Où θ est relié à l'équation 5.12 par :

$$\tan \theta = \frac{X}{D}$$

Figure 5.5 : Détermination de la résolution par la surface de détecteur.

- la qualité du cristal : présence de défauts, mosaïcité, désordre...
 - l'agitation thermique des atomes car l'intensité I est proportionnelle au facteur de Debye-Waller B et à la résolution d :

$$\langle I \rangle \propto \exp\left(\frac{-2B}{d^2}\right) \quad \text{et} \quad B \propto \langle \Delta r^2 \rangle \quad (5.16)$$

où Δr est le déplacement quadratique moyen de l'atome.

La structure tridimensionnelle est la transformée de Fourier des diagrammes de diffraction. La densité électronique ρ à un point (x,y,z) dans la maille peut être décrite par :

$$\rho(x, y, z) = V^{-1} \sum \sum \sum |F(h, k, l)| \exp[i\alpha(h, k, l)] \exp[-2\pi i(hx + ky + lz)] \quad (5.17)$$

où x , y et z sont les fractions de la maille de dimensions a , b , et c ; $|F(h, k, l)| = \sqrt{I(h, k, l)}$ est l'amplitude de la réflexion (h, k, l) ; et $\alpha(h, k, l)$ est la phase de la réflexion (h, k, l) .

La densité électronique peut être alors calculée pour tout point (x, y, z) dans la maille si le facteur de structure $F(h, k, l)$ est connu pour chaque réflexion. Cette densité électronique va servir plus tard à mettre en place le modèle de la protéine.

Pour vérifier la qualité des données à ce niveau de l'analyse, il existe certains critères tels qu'une bonne complétude (entre 90 et 100 %), une redondance élevée des données, un rapport signal sur bruit $I/\sigma I$ important, et une bonne cohérence interne des données exprimée par le R_{sym} .

$$R_{sym} = \frac{\sum_{hkl} \sum_j |I_j(h, k, l) - \overline{I(h, k, l)}|}{\sum_{hkl} \sum_j I_j(h, k, l)} \quad (5.18)$$

où pour n réflexions indépendantes et j observations d'une réflexion donnée, $\overline{I(h, k, l)}$ est l'intensité moyenne pour les j observations.

5.2.3.2 Résolution du problème de phase

Lorsqu'un spectre de diffraction est enregistré, cela signifie que l'intensité des ondes diffractées par le cristal est mesurée. Cette intensité permet de remonter au facteur de structure, donc à l'amplitude de l'onde diffusée pour chacune des directions de diffractions avec les équations 5.5 et 5.6. Mais pour pouvoir calculer la densité électronique, c'est à dire calculer la transformée de Fourier des ondes diffractées, il faut connaître leur phase. Actuellement, trois méthodes sont les plus couramment utilisées pour fournir une estimation des phases des cristaux des protéines :

- la méthode du remplacement isomorphe multiple (MIR). Par diffusion d'un atome lourd sous forme de sel (Au, Pt, U, Hg) dans le cristal de protéine, ce dernier va se fixer sur la protéine de façon spécifique sans modifier l'empilement cristallin. Si l'on retrouve le facteur de structure de la protéine avec le sel, ainsi que la position des atomes lourds, il est possible de déterminer la phase dans le cristal natif correspondant.
- la méthode de diffusion anormale multiple (MAD). La diffusion anormale est la diffusion de rayons X par un atome (Se, Hg, Fe, Pt, U) lorsque la longueur est proche du seuil d'absorption. Il en découle un déphasage entre l'onde incidente et l'onde ré-émise par l'atome. Cela se traduit dans la forme du facteur atomique qui devient plus complexe : $f = f_0 + f'(\lambda) + if''(\lambda)$. Les termes f' et f'' , dépendants, de la longueur d'onde, sont négligeables pour les atomes C, N, O, S mais deviennent significatifs pour ceux précédemment cités. En enregistrant à plusieurs longueurs d'onde (minimum trois : deux proches du seuil et une loin du seuil), les jeux de données différeront au niveau de la contribution anormale. L'analyse des différents jeux permet de localiser les

atomes diffuseurs anomaux et de déterminer les phases des facteurs de structures de la protéine.

- la méthode du remplacement moléculaire (RM), qui consiste à placer dans la nouvelle maille, une molécule proche en structure de celle que l'on cherche à déterminer. L'affinement permet ensuite de faire converger le modèle de départ vers la nouvelle structure. Le but est de placer une de ces structures analogues dans la nouvelle maille : six paramètres sont donc à définir, trois pour trouver la position du barycentre du modèle et trois pour déterminer son orientation. Ce nombre de six paramètres est à multiplier par le nombre de molécules contenues dans la maille ou l'unité asymétrique. La meilleure position permet d'obtenir des facteurs de structures calculés les plus proches des facteurs de structures mesurés.

Dans notre cas, les structures de plusieurs γ cristallines, dont la γ E cristalline de rat, sont connues à haute résolution, la méthode de remplacement moléculaire a donc été utilisée.

La première étape est d'orienter le modèle grâce à la fonction de Patterson. La fonction de Patterson $P(R)$ est la fonction d'auto-corrélation de la densité électronique, elle se calcule simplement à partir des intensités mesurées :

$$P(r) = \sum_R (F(R))^2 \exp(2i\pi r \cdot R) = \int \rho(u) \cdot \rho(u+r) du \quad (5.19)$$

Cette fonction présente un maximum quand r correspond à un vecteur inter atomique. Schématiquement, la fonction de Patterson représente l'ensemble des vecteurs inter atomiques. Elle dépend donc de l'orientation de la molécule. Pour orienter l'orientation de la molécule, il suffit de superposer la fonction de Patterson expérimentale et celle correspondant au modèle ; c'est la fonction de rotation.

La seconde étape correspond à trouver la position du modèle qui donne le meilleur accord entre les facteurs de structures mesurés et calculés ; c'est la fonction de translation.

Note : la méthode de remplacement moléculaire n'apporte pas de phases expérimentales et les phases initiales vont être biaisées par le modèle de départ choisi.

5.2.3.3 Reconstruction et affinement

L'affinement est l'étape qui permet de transformer le modèle initial (construit dans la carte expérimentale ou issu du remplacement moléculaire) et d'obtenir un modèle final optimum.

Il s'agit en fait d'une succession d'étapes de reconstruction sur station graphique et d'étapes d'affinement numérique. Le modèle final doit offrir un accord optimal entre les facteurs de structure mesurés et ceux calculés à partir du modèle.

5.2.3.3.1 La reconstruction

Le modèle doit être modifié pour mieux expliquer la carte de densité électronique. Pour cela des cartes différentielles vont être étudiées sur des stations graphiques, telles que les cartes $\{F_{obs} - F_{calc}\}$ (équation 5.20) ou $\{2F_{obs} - F_{calc}\}$ (équation 5.21) :

$$\rho_{obs}(xyz) - \rho_{calc}(xyz) \approx \sum_{hkl} (F_{obs} - F_{calc}) \exp(i\phi_{calc}) \exp(2\pi i(hx + ky + lz)) \quad (5.20)$$

Aussi appelée carte résiduelle, elle illustre le désaccord entre le modèle et les facteurs de structures observés. Un pic positif indique la possibilité d'atomes manquant dans le modèle. Un pic négatif sur un atome du modèle indique que celui-ci est sans doute mal placé. En fin d'affinement, cette carte doit être aussi « plate » que possible.

$$2\rho_{obs}(xyz) - \rho_{calc}(xyz) \approx \sum_{hkl} (2F_{obs} - F_{calc}) \exp(i\phi_{calc}) \exp(2\pi i(hx + ky + lz)) \quad (5.21)$$

C'est une combinaison linéaire de la carte de densité électronique et de la carte résiduelle. Elle permet de visualiser les imperfections du modèle.

5.2.3.3.2 Les méthodes numériques d'affinement

L'affinement consiste à compléter le modèle et optimiser les paramètres que sont les coordonnées des atomes, les facteurs de températures individuels, l'occupation pour obtenir le meilleur accord entre les facteurs de structures mesurés et calculés tout en maintenant une stéréochimie acceptable. Pour maintenir cette stéréochimie, il est indispensable d'introduire des contraintes rigides (liaisons inter atomiques figées, blocs rigides...) ou flexibles (angles de valence par exemple). Ces informations stéréochimiques, provenant d'autres structures connues à très haute résolution, constituent autant d'observations supplémentaires pour l'affinement. D'un point de vue pratique, ces programmes d'affinement, *Shelx* basé sur la méthode des moindres carrés et *Refmac*, utilisant la fonction de probabilité d'avoir le même type de structure, cherchent à minimiser des fonctions où sont mentionnées les contraintes du type suivant :

$$\begin{aligned}
E(r, b, q) = & \sum_h 1/\sigma_{F(h)^2} (F_o(h) - F_C(h))^2 + \sum_{liaisons} 1/\sigma_{D^2} (d_j^{idéal} - d_j^{model})^2 + \sum_{angles} 1/\sigma_{A^2} (\alpha_j^{idéal} - \alpha_j^{model})^2 \\
& + \sum_{chiral} 1/\sigma_{Chi^2} (V_j^{idéal} - V_j^{model})^2 + \sum_{plans} \sum_k 1/\sigma_{PI^2} (m_k \cdot r_{ik} - d_k)^2 + \sum_{ltorsions} 1/\sigma_{\sigma T^2} (\chi_j^{idéal} - \chi_j^{model})^2 \\
& + \sum_{liaisons} 1/\sigma_{n^4} (d_{ij}^{idéal} - d_{ij}^{model})^4 \tag{5.22}
\end{aligned}$$

Note : L'affinement par deux programmes n'utilisant pas la même méthode, ne conduit pas à des facteurs R de même valeur. C'est le cas pour l'affinement de la protéase K résolue à 0,98 Å (Betz et al. 2001), où les facteurs R_{fact} , R_{libre} et B de *ShelX* sont toujours plus importants que ceux de *Refmac*.

5.2.3.3 Critères de validité d'un affinement

Pour valider un affinement, il faut que :

- le coefficient de corrélation entre F_{obs} et F_{calc} soit supérieur à 0,9 :

$$cc(F_{obs}, F_{calc}) = \frac{\langle F_{obs} F_{calc} - \langle F_{obs} \rangle \langle F_{calc} \rangle \rangle}{\sqrt{\langle F_{obs}^2 - \langle F_{obs} \rangle^2 \rangle \langle F_{calc}^2 - \langle F_{calc} \rangle^2 \rangle}} \tag{5.23}$$

- le facteur R ou facteur de désaccord soit aux environs de 0,2 en fin d'affinement :

$$R_{fact.} = \frac{\sum |F_{obs} - F_{calc}|}{\sum F_{obs}} \tag{5.24}$$

- le facteur R libre, se situe entre 0,25 et 0,30. 5 à 10 % des données sont choisies aléatoirement et sont exclues de l'affinement. Si, au cours de l'affinement, le facteur R libre (facteur calculé sur ce sous ensemble de réflexions) diminue, il y a une amélioration objective du modèle.

$$R_{libre} = \frac{\sum_{libre} |F_{obs} - F_{calc}|}{\sum F_{obs}} \tag{5.25}$$

Ce facteur permet d'éviter un sur-affinement, c'est à dire des critères statistiques trop optimistes ne traduisant que le manque de données, comparé au nombre de paramètres inclus dans l'affinement. La stéréochimie des résidus est aussi un critère à vérifier. L'écart type sur les liaisons covalentes doit être de l'ordre de 0,01Å et celui sur les angles de valence de l'ordre de 1,5°. De plus les contraintes exercées sur la chaîne principale, vérifiées par le Diagramme de Ramachandran, sont des critères de validité du modèle.

5.3 Matériels et méthodes

5.3.1 Cristallisation des γ cristallines

Différentes techniques de cristallisation, pour obtenir des cristaux hydrogénés et deutériés des γ E cristallines de rat, dans des tampons de cristallisation eux même hydrogénés et deutériés ont été utilisées. Ces techniques se regroupent principalement en :

- des techniques par diffusion de vapeur en goutte assise ou goutte suspendue avec micro, macro et simple ensemencements;
- des techniques de dialyse utilisant des boutons de dialyse ou des capillaires;
- des techniques d'induction par la température.

Toutes ces techniques seront décrites et discutées dans le chapitre 7 pour l'obtention de cristaux de taille suffisante pour la diffraction neutronique. La technique de diffusion de vapeur en goutte suspendue est la technique ayant donné le plus grand nombre de résultats probants pour la cristallographie aux rayons X.

Les conditions de cristallisation de la γ E cristalline de rat, ayant été déjà définies pour l'obtention d'une structure à 2,3 Å (Norledge et al. 1997), deux approches ont été effectuées en goutte suspendue. La première consistait à améliorer les conditions connues de cristallisation. La seconde consistait à cribler de nouvelles conditions pour ensuite optimiser la qualité des nouvelles formes cristallines obtenues. Pour cette technique, le réservoir, dans une boîte de 24 puits (boîte Linbro de chez Hampton Research) est rempli avec une solution aqueuse de précipitant, cette solution est mélangée dans différentes proportions (généralement 2 μ l et 2 μ l de protéines) avec la solution protéique en une petite goutte sur une lamelle de verre (Hampton Research). Finalement, la lamelle est retournée et recouvre hermétiquement un puit grâce à la graisse à vide. Les échanges gazeux entre le réservoir et la goutte peuvent avoir lieu, conduisant à une variation du volume de la goutte jusqu'à ce que l'équilibre soit atteint entre la solution du réservoir et la goutte.

La protéine γ E cristalline extraite de lentille de rat (30mg/ml), avait été cristallisée par la méthode de la goutte suspendue dans 13 % (w/v) PEG 6000 dissout dans un tampon 0,05 M Σ NaPO₄, pH 6,86 à température ambiante (Norledge et al. 1997). Au laboratoire de deutériation de l'ILL, j'ai cristallisé la même protéine native, exprimée chez *E. coli* (5-20 mg/ml) par différentes méthodes et dans de nombreuses conditions (voit tableau 5.2) en utilisant les conditions d'Hampton Research et de Molecular Dimension Limited. Les cristaux natifs, des formes

hydrogénées et deutériées, utilisés pour résoudre les structures, en H₂O et en D₂O, ont été formés en 3-4 jours, par la méthode de la goutte suspendue. Les solutions aqueuses de précipitations étaient composées de Magnésium Acétate 0,2 M (voir tableau 5.2 rangées 6 à 9).

Cristal	Concentration en γE cristalline	Tampon	pH	Précipitant	Additif
1	8-20 mg/ml	0,05-0,2 M Σ Sodium Phosphate	5,9-7,1	10-15% PEG 6000	+/- Manganese Chloride.4H ₂ O +/- Strontium Chloride.6H ₂ O +/- Hexaminecobalt Trichloride +/- Sodium Thiocyanate +/- Glycérol 13% +/- Polypropylène glycol P400 +/- γ Butyrolactone
2	10-15 mg/ml	0,1-0,2 M Sodium Acétate	5,7-6,5	17-19% PEG 6000	+/- Glycérol +/- Imidazole +/- N-octyl- β -D-glucoside
3	9 -12 mg/ml	0,2 M Magnésium Formate	5,9	20 % PEG 3350	-
4	10 mg/ml	0,2-1,4 M Sodium Acétate 3H ₂ O	6,5	0 % 20% PEG 3350	0,1 M Cacodylate -
5	10-15 mg/ml	0,1M Hepes Sodium	7-7,5	0,8 M NaH ₂ PO ₄ 0,8M KH ₂ PO ₄	-
6	8-12 mg/ml hydrogénée (H γ E)	0,2 M Magnésium Acétate 100 % H ₂ O	7,7	18 % PEG 3350	-
7	8-12 mg/ml H γ E	0,2 M Magnésium Acétate 100 % D ₂ O	7,4	20 % PEG 3350	-
8	10 mg/ml deutériée (D γ E)	0,2 M Magnésium Acétate 100 % H ₂ O	7,7	19 % PEG 3350	-
9	10 mg/ml D γ E	0,2 M Magnésium Acétate 100 % D ₂ O	7,4	20 % PEG 3350	-

Tableau 5.2 : Conditions de cristallisation ayant donné des cristaux. PEG signifie Poly Ethylène Glycol.

5.3.2 Optimisation des solutions cryogéniques

Depuis de nombreuses années, les cristallographes se sont rendus compte des avantages de collecter des données sur des cristaux à très faibles températures, telles que celle de l'azote liquide. En plus d'une meilleure diffraction, il existe d'autres avantages de la cryocristallographie : la diminution des dommages dus aux radiations des rayons X et donc la possibilité de collecter plus de données à partir d'un cristal ; la réduction de la diffusion des rayons X provenant de l'eau qui entoure le cristal, ce qui produit un plus faible bruit de fond. En théorie, le fait d'abaisser la température doit augmenter l'ordre moléculaire dans le cristal et donc augmenter la diffraction. Cependant, en pratique, la congélation de cristaux peut résulter en des dommages causés par la formation de cristaux de glace. Pour palier à ce problème, les

cristallographes ont développé des techniques de congélation rapide des cristaux en présence d'agents comme le glycérol, ce sont les *cryoprotecteurs*.

Pour protéger le cristal des effets néfastes de la cryogénie, il est possible de co-cristalliser ou de transférer le cristal de protéine dans une solution d'agent cryoprotectant. Le cristal dans sa boucle est rapidement placé dans un flux gazeux d'azote à 100K, ou alternativement congelé dans de l'azote liquide. La présence de solution cryoprotectrice et une congélation très rapide préviennent de la formation de glace cristalline et doivent préserver les propriétés de diffraction du cristal, bien que la mosaïcité va être légèrement augmentée.

Au niveau du cristal, la cryogénie consiste à introduire un cryoprotecteur dans les réseaux du solvant ou un surfactant qui va entourer les protéines. Certains composés sont de bons candidats pour agir comme des cryoprotecteurs : MPD (2-méthyl-2,4-pentanediol) dans une gamme de 25 à 50 %; certains PEG (PolyEthylène Glycol) déjà présents dans certaines conditions de cristallisation; l'éthylène glycol dans une fourchette de 10 à 40 % ; le glycérol et le sucrose dans la même gamme de 10 à 50 % ; $(\text{NH}_4\text{SO}_4$ ou $(\text{Li}_4)_2\text{SO}_4$ dans un éventail de 60 à 90 %...Un bon cryoprotecteur va se combiner avec les composants de la liqueur mère pour remplacer une partie voire tout le solvant autour des protéines sans dissoudre le cristal.

5.3.3 Collecte des données cristallographiques

5.3.3.1 Le synchrotron : source de rayons X

La source de rayons X utilisée pendant cette étude est le rayonnement synchrotron de l'E.S.R.F. de Grenoble (voir figure 5.6).

Les synchrotrons sont des grands « outils » employés pour faire circuler à très haute vitesse (proche de la vitesse de la lumière) des particules chargées électriquement (des électrons chargés négativement ou des positrons chargés positivement).

Les particules sont injectées dans l'anneau de stockage via l'anneau d'accélération (le booster) qui a été alimenté par l'accélérateur linéaire. Quand le faisceau de particules change de direction, les électrons ou les positrons sont accélérés à travers le centre de l'anneau et vont émettre des radiations électromagnétiques et surtout des rayons X, et donc une perte d'énergie. Cette perte d'énergie est compensée par de nouvelles injections de radiofréquences à chaque cycle.

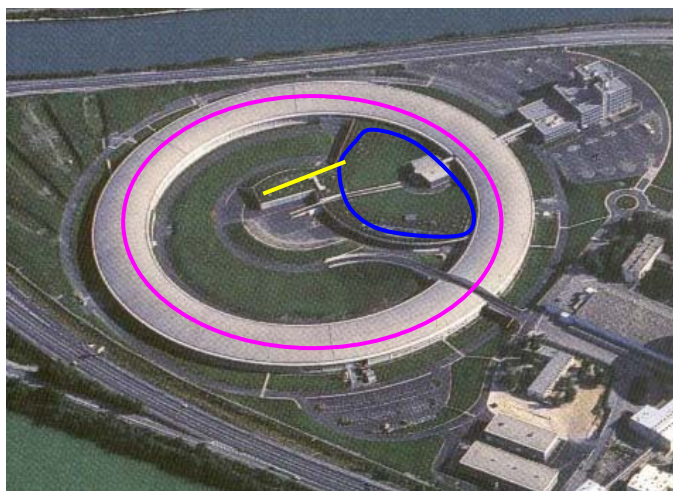


Figure 5.6 : L'E.S.R.F. à Grenoble a une circonférence de 844,39m. Il est composé d'un accélérateur linéaire (en jaune), d'un « booster » (en bleu) et d'un anneau de stockage (en violet).

La trajectoire des particules est déterminée par leur énergie et par le champ magnétique, qui permettent leur changement de direction. Il y a 4 types d'éléments magnétiques dans l'anneau de stockage :

- les aimants de courbures nécessaires pour guider les électrons ;
- les décaleurs de longueur d'onde, avec un champ magnétique intense et une courbure plus précise ;
- le « wiggler », aimant faisant osciller la trajectoire des électrons ;
- et l'onduleur, similaire au « wiggler » mais avec la différence que les effets d'interférences produisent l'émission de radiation à des longueurs d'ondes plus spécifiques. Les ondulations dues au « wiggler » et à l'onduleur vont aussi augmenter le flux du faisceau.

Toutes les collectes ont été effectuées à l'ESRF sur les lignes ID14-EH2 et EH3. Dans les deux cas, les rayons X proviennent d'un onduleur donnant une haute brillance et un spectre énergétique très précis. Sur ID14, le faisceau est dédoublé par les monochromateurs semi-réfléchissants de diamant pour conduire le rayon X dans quatre cabines expérimentales différentes.

5.3.3.2 Diffraction des rayons X

Tous les jeux de données ont été collectés par la technique d'oscillation où le cristal tourne autour d'un axe perpendiculaire au faisceau incident de rayons X par pas de 0,5 à 1°, sur une plage angulaire dépendante de la symétrie du cristal. La diffraction résultante est enregistrée sur un détecteur ADSC (Area Detector Systems Corporation) quantum-4 CCD bidimensionnel (voir figure 5.7).

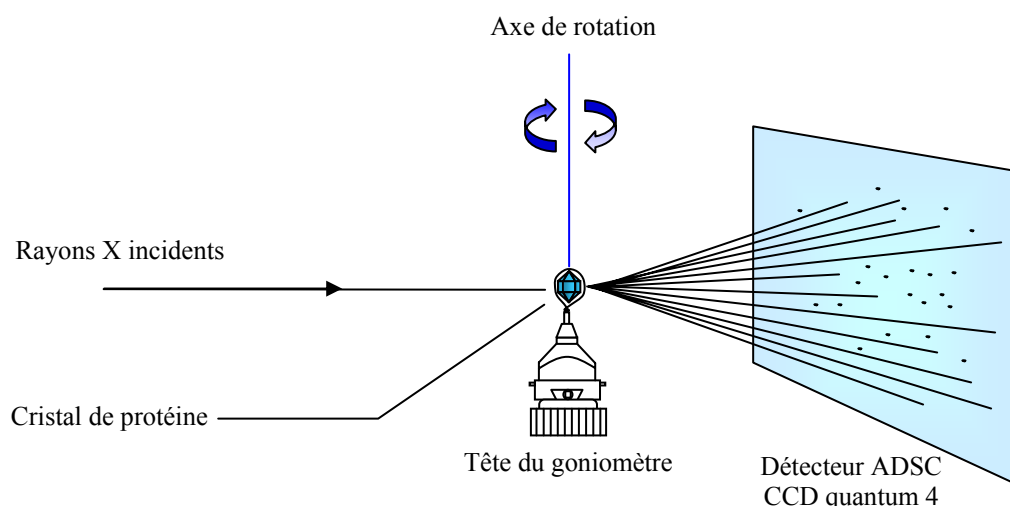


Figure 5.7 : Représentation schématique d'une expérience de diffraction de rayons X par un cristal. Le faisceau de rayons X va frapper à intervalle de temps régulier le cristal, qui aura oscillé de quelques degrés autour d'un axe (en bleu) perpendiculaire au faisceau incident. Les faisceaux diffractés du cristal seront collectés sur le détecteur ADSC à 2 dimensions.

5.3.4 Traitement des données

5.3.4.1 Traitement des images

Le traitement des images a été effectué par les programmes *Mosflm* (Powell 1999) et/ou *Xds* (Kabsch 1993). Le but de cette étape est d'obtenir avec précision les différents paramètres expérimentaux, tels que les dimensions de la maille cristalline, le groupe d'espace du cristal ; et les paramètres de l'instrument tels que la distance cristal détecteur. Ces éléments vont permettre de fournir à chaque spot ses propres indices de Miller et d'intégrer par la suite, leur intensité (Kabsch 1988). La mise à l'échelle des données a été effectuée par le programme *Scala* de CCP4.

5.3.4.2 Résolution du problème de phase

Etant donné que plusieurs structures de γ cristallines ont été résolues, dont celle de la γ E cristalline de rat, il n'y a pas de problème de phase, il suffit de placer notre molécule dans les nouvelles mailles qui ont été obtenues. Le programme *Molrep* (Vagin et Teplyakov 2000) de CCP4 a été utilisé pour définir les paramètres de rotation et de translation. Le modèle de départ est le modèle de la γ E cristalline de rat résolue à 2,3 Å de résolution (Norledge et al. 1997). Une fois la position déterminée, une optimisation de corps rigide (réalisé avec le programme *Rigid Body* de CCP4) en utilisant les réflexions de 10 Å à 3 Å permet d'affiner les six paramètres d'orientation et de translation.

5.3.4.3 Reconstruction et affinement

Par la suite, les programmes *Refmac* (Murshudov et al. 1997) et *ShelXL* et l'interface *ShelXPro* (Sheldrick et Schneider 1997) ont été utilisés pour affiner la position du modèle en respectant avec plus ou moins de poids les contraintes géométriques (longueur de liaison, angle etc...) et pour calculer les cartes de densité électroniques $2|F_o|-|F_c|$ et $|F_o|-|F_c|$. Les programmes *O* (Jones et al. 1991) et *Xtalview* (McRee 1992) ont permis de reconstruire manuellement, l'architecture des différentes formes isotopiques de γ cristallines en plaçant le modèle dans les densités électroniques expérimentales. Plusieurs structures de γ cristallines ont été utilisées pour affiner les modèles où la densité était peu ou mal définie. Enfin, les molécules d'eau ont été positionnées à l'aide du programme *Arp-warp* (Perrakis et al. 2001) en suivant certains critères : (a) la molécule d'eau doit faire au minimum une liaison hydrogène dont la longueur est comprise entre 2,4 et 3,5 Å , (b) la densité électronique doit être bien définie au dessus de σ dans la carte $2|F_o|-|F_c|$, (c) la densité ne présente pas de désaccord à 3σ dans la carte $|F_o|-|F_c|$. Chaque étape de reconstruction a été suivie par une étape d'affinement pendant laquelle les positions atomiques du modèle étaient ajustées à la fois avec les facteurs de structures expérimentales et les contraintes stéréochimiques. Les critères de validité de l'affinement seront le facteur R ($R_{\text{fact.}}$), ou facteur de désaccord, et le facteur R libre, qui exprime l'amélioration objective de l'affinement, sur 5% des données prises aléatoirement.

5.3.4.4 Validation des structures

Tous les modèles protéiques des γ cristallines ont été vérifiés et validés par le programme *Procheck* (Laskowski et al. 1993) en utilisant des critères et des contraintes stéréochimiques. Le programme *Moleman* a permis de calculer les centres de symétrie pour déterminer les axes pseudo dyades qui ont été visualisés par le programme de visualisation *Pymol*. *Contact* a permis de déterminer les différents contacts protéine-protéine, protéine-solvant et solvant-solvant. *Baverage* est un programme qui permet d'analyser les facteurs thermiques B par résidus et par atomes d'oxygène pour les molécules d'eau. Enfin, *Superpose* permet d'analyser les différences topologiques entre différentes structures et le programme *Lsqkab* en a calculé les déviations.

5.4 Résultats et discussions

5.4.1 Conditions de cristallisation

La cristallisation de la γE cristalline de rat produite dans *E. coli* a été réalisée dans différentes conditions. Quatre formes cristallines de la protéine hydrogénée, ont été obtenues, à température ambiante (20°C), par les méthodes de diffusion de vapeur (voir figure 5.8) : des aiguilles regroupées ou non, des plaques individuelles ou non et des cristaux en 3 dimensions.

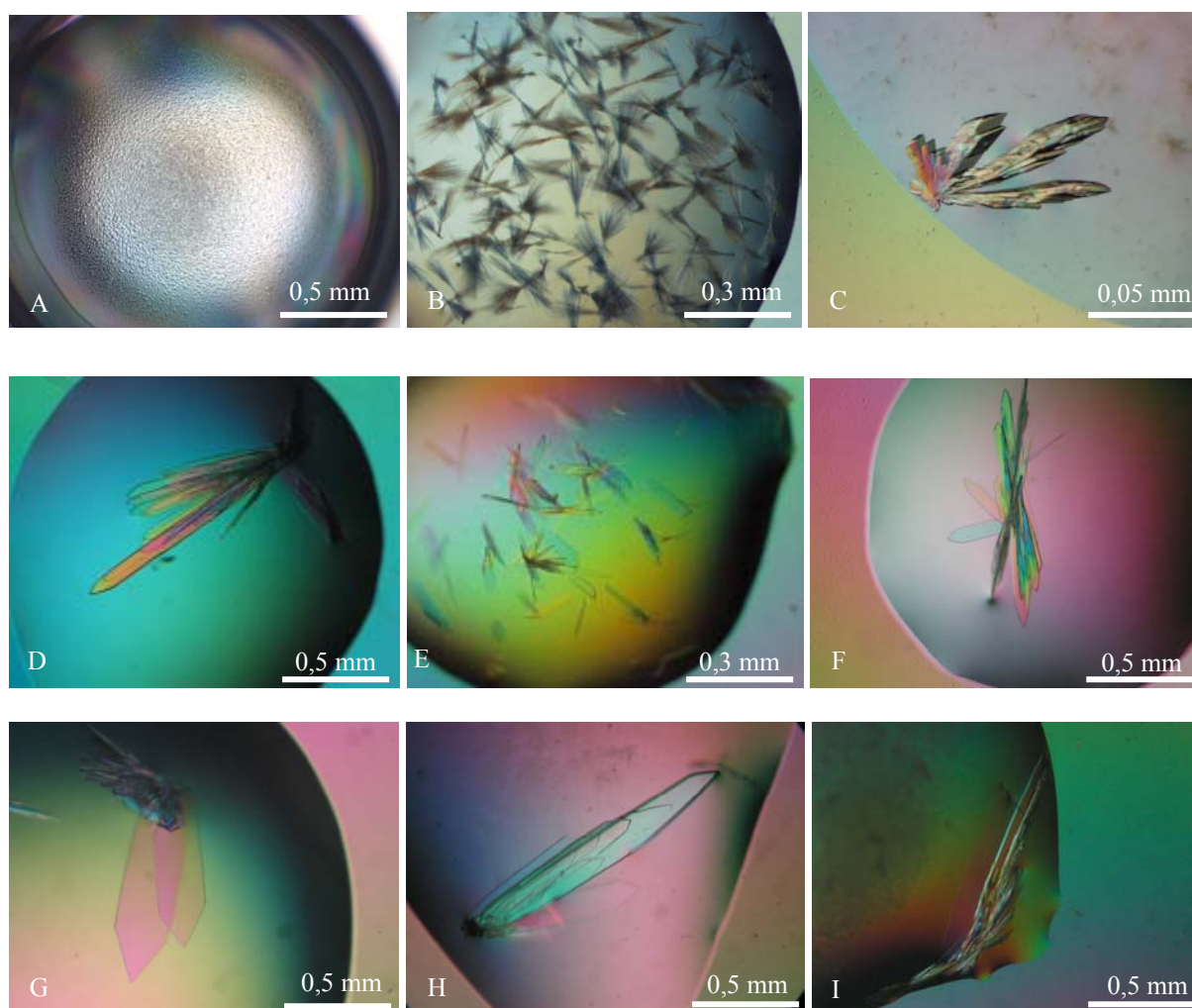


Figure 5.8 : Morphologie des différentes formes cristallines obtenues pour la γE cristalline. La barre représente l'échelle des cristaux. **A)** Quasi cristaux 0,05 M NaPO₄, pH 6,8, PEG 6K 13 %; **B)** Ensemble d'aiguilles 0,1M Hepes pH7,3, Na, K PO₄ 0,8M; **C)** Cristal en 3D 0,1 NaAc pH 6,5, PEG 6K 18 %, N-octyl- β -D-glucoside, **D)** ensemble de cristaux en plaques 0,05M NaAc pH6,2, PEG 17%; **E)** Cristaux individuels 0,05M NaPO₄ pH6,5, PEG 6K12 %; **F)** protéine hydrogénée HyE en tampon h₂O (HyEh₂O) 0,2 M MgAc pH 7,7, 18% PEG 3,35K; **G)** HyEd₂O 0,2 M MgAc pH 7,4, 20% PEG 3,35K; **H)** DyEh₂O 0,2 M MgAc pH 7,7, 19% PEG 3,35K; **I)** DyEd₂O 0,2 M MgAc pH 7,4, 20 % PEG 3,35K.

Les deux problèmes majeurs, liés à la cristallogénèse et rencontrés pendant cette thèse, ont été :

- la reproductibilité des cristaux. Sur une solution protéique, dont la pureté a été vérifiée sur gel en conditions dénaturantes, par chromatographie d'exclusion et par spectrométrie de masse, seulement 40 à 45 % des essais fournissaient des cristaux ou des formes cristallines en condition identique. En revanche, ces essais ont toujours permis des cristaux sous forme de petits assemblages de plaques cristallines, après ensemencement.
- une très petite troisième dimension du cristal. La plupart des cristaux a une morphologie de plaques relativement fines qui peuvent être associées entre elles. Seul, l'hépes réuni au phosphate de sodium et au phosphate de potassium, a montré la formation d'aiguilles toujours associées sous la forme de hérisson. Bien qu'un grand nombre de conditions a été testé, grâce aux kits de chez Hampton Research et de Molecular Dimension Limited, la morphologie n'a guère évolué. Après optimisation, les cristaux obtenus par la méthode de diffusion de vapeur ont toujours été de taille importante dans deux dimensions et la troisième n'a rarement dépassé quelques dizaines de micromètres (voir figure 5.9). Seulement deux conditions ont donné une troisième dimension visible, mais celles-ci n'ont jamais pu être reproduites, malgré de nombreuses tentatives. La diffraction de ces cristaux n'a abouti qu'à de la diffraction supérieure à 2,1 Å avec des tâches de diffraction dédoublées, et aucun jeu de données complet n'a été collecté.

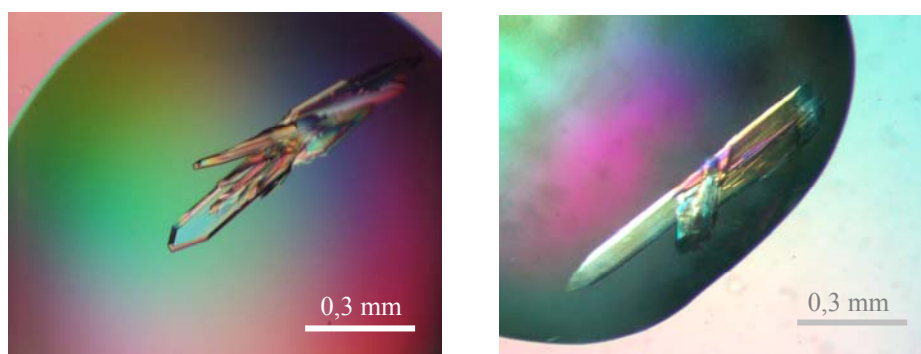


Figure 5.9 : Cristaux en 3D de la protéine γ E cristalline hydrogénée en solution hydrogénée. La barre représente l'échelle des cristaux. A) 0,1 M NaPO₄ pH 6,9, 19 % PEG 6K + 15 % sucrose ; 0,2 M Mg Ac pH 7,6, 13 % PEG 3,35K.

Les conditions de cristallisation se trouvent dans une gamme de pH assez large mais centrée aux alentours de la neutralité. Le pH ne semble pas être une condition limitante puisque des conditions de cristallisation ont été observées dans une gamme allant de pH 5,7 à 7,7 (voir tableau 5.1). Le polyéthylène glycol (PEG), de différente taille, a souvent été utilisé comme précipitant, mais ceux qui ont démontré une plus grande reproductibilité, sont les PEG 3350 et 6000.

5.4.2 Collectes des données

Pour collecter des jeux de données à partir des cristaux sous forme de plaque, il a souvent été nécessaire de les séparer. Cette opération a généralement conduit à la destruction de l'assemblage en de nombreuses plaques de différentes tailles mais aucun cristal n'a été dissout suite à cette séparation.

Une fois, les cristaux séparés les uns des autres, ces derniers sont préparés pour la cryogénie comme décrite dans la partie Matériels et Méthodes. A partir des cristaux de γ E cristallines, il n'a pas été possible de collecter des données sans ajout de cryoprotecteur (test réalisé sur 8 cristaux), malgré le fait que la solution mère contienne une quantité importante de polyéthylène glycol (~ 20 %). En effet, la congélation rapide dans ces conditions, montrait la formation de glace autour du cristal. Pour palier à ce problème, plusieurs cryoprotecteurs ont été testés avec plus ou moins de succès (critère fondé sur la formation de glace autour du cristal). Le meilleur cryoprotecteur utilisé, dans notre cas comme pour de nombreuses protéines (Garman 2003), est le glycérol supplémentant la solution mère de cristallisation. Des essais avec un trempage direct dans la solution mère supplémentée de 20 % glycérol ont permis d'obtenir des collectes de données uniquement jusqu'aux environs de 1,7 Å. En transférant un cristal provenant de la même goutte, dans plusieurs solutions similaires à la solution mère avec des concentrations croissantes en glycérol (jusqu'à 20 %) sur une période de cinq minutes, des collectes de données à de résolution de 1,4 Å ont été obtenues.

Après les avoir protégés de la cryogénie, les cristaux sont montés sur une cryo boucle (Hampton Reseach) et congelés dans un bain d'azote liquide ou directement placés dans le flux gazeux d'azote.

La plupart des cristaux testés ont une limite de résolution aux alentours ou supérieure à 1,7 Å, mais beaucoup d'entre eux présentent des tâches de diffraction doubles voire triples. Ceci

est expliqué par le fait que les cristaux obtenus ont poussé sous forme de plaques difficilement dissociables au microscope. Ces plaques peuvent générer le même type de diffraction mais avec une orientation du cristal différente : les tâches de diffraction en seront alors dédoublées voire triplées. Ce dédoublement des tâches de diffraction est dû au fait que les ondes ne sont pas réfléchies par le même réseau cristallin, ce qui ne constitue pas un empilement idéal. Ces cristaux sont appelés cristaux jumeaux. Cependant, certaines plaques dites "uniques" peuvent être isolées, et les tâches de diffraction sont alors uniques (voir figure 5.10). Etant donné que les cristaux ont une longueur souvent très importante ($> 0,3$ mm), il arrive que des cristaux présentent une diffraction double d'un côté du cristal, mais ce n'est pas le cas à l'autre bord. A partir de cristaux "uniques" de protéines hydrogénée ou deutériée en solvant H_2O ou en D_2O , des jeux de données, dont le groupe d'espace était toujours un système orthorhombique en $P2_12_12_1$, ont été collectés à l'ESRF à la ligne ID14 entre de résolution de 1,72 et 1,36 Å (voir tableau 5.3). Ces données, une fois traitées, ont permis d'effectuer une étude comparative sur des cristaux hydrogénés ou deutériés en H_2O et D_2O , dont les conditions de cristallisation étaient similaires (~ 0.2 M MgAc pH ~ 7.7 , $\sim 19\%$ PEG 3350).

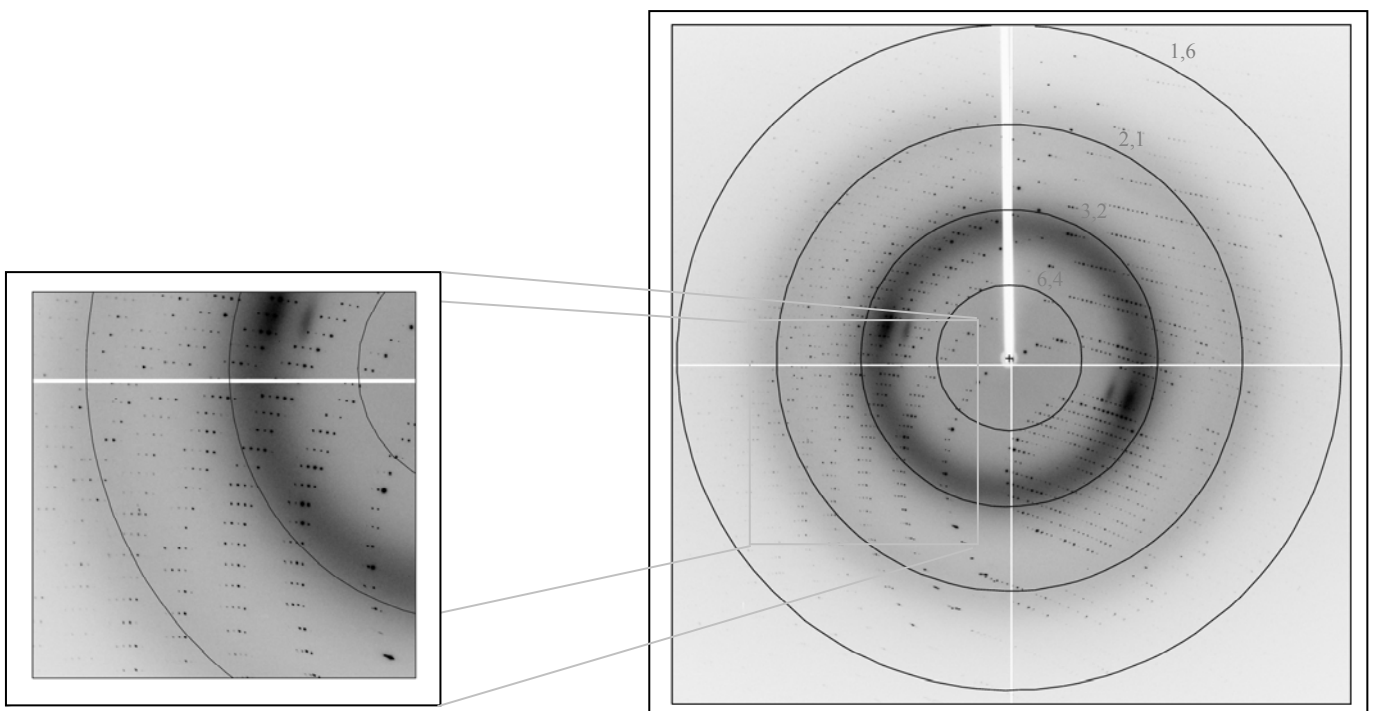


Figure 5.10 : Clichés de diffraction issus du jeu de données collecté à partir des cristaux du HyEd₂₀. Les tâches de diffraction ne sont pas doubles ce qui exprime la présence d'un cristal unique.

Cristal	Groupe d'espace	a	b	c	α	β	γ	Molécules par unité asymétrique	Contenu en solvant en %	Volume de la maille en Å ³	Résolution en Å
<i>Rat γE^*</i>	P12 ₁ 1	37,76	43,71	107,44	90,00	108,54	90,00	2	38,20	168125,59	2,30
<i>HγEh2o</i>	P2 ₁ 2 ₁ 2 ₁	37,20	43,52	102,35	90,00	90,00	90,00	1	37,30	165698,92	1,45
<i>HγEd2o</i>	P2 ₁ 2 ₁ 2 ₁	37,63	43,42	101,86	90,00	90,00	90,00	1	37,57	166428,50	1,44
<i>DγEh2o</i>	P2 ₁ 2 ₁ 2 ₁	37,53	43,32	101,61	90,00	90,00	90,00	1	37,11	165197,50	1,36
<i>DγEd2o</i>	P2 ₁ 2 ₁ 2 ₁	37,39	43,33	101,53	90,00	90,00	90,00	1	36,84	164489,64	1,72

Tableau 5.3 : Paramètres de différentes collectes de données à partir des protéines hydrogénées ou deutériées en solvant H₂O ou D₂O. * sont les données cristallographiques de la structure connue 1A5D (Norledge et al. 1997).

5.4.3 Traitements des données

Tous les jeux de données ont été traités de façon individuelle, comme si les protéines cristallisées étaient complètement différentes les unes des autres. Les programmes *Mosflm* et *Xds* ont servi à déterminer l'orientation du cristal, les paramètres de la maille et les possibles groupes d'espaces. Ils vont servir aussi à générer les listes de réflexions et à intégrer les tâches de diffractions des images collectées. Le programme *XDS* a été utilisé car la plupart des données présentaient une mosaïcité élevée. Ce programme permet de produire des profils, en trois dimensions, des intensités des tâches de diffraction, et utilise ces profils pour estimer les amplitudes des facteurs de structure à partir de l'intensité de chaque réflexion.

5.4.4 Remplacement moléculaire

Afin de replacer les différents modèles dans les mailles expérimentales, la méthode par remplacement moléculaire a été adoptée. Le modèle utilisé est en fait la structure de la protéine γE cristalline du rat référencée en 1A5D dans la PDB (Protein Data Bank : <http://www.rcsb.org/pdb/index.html>). Ce modèle présente deux molécules par unité asymétrique. En revanche, toutes les protéines recombinantes de cette thèse, cristallisées dans de l'acétate de magnésium et du PEG 3350, présentent une seule molécule par unité asymétrique. Cette valeur, ainsi que la proportion de solvant dans la maille, ont été calculés par le programme *Matthews_coef*

de CCP4 (Matthews 1968). Le remplacement moléculaire a tout de même été effectué dans un premier temps avec le modèle original (2 molécules par unité asymétrique). Dans un autre temps, un nouveau remplacement moléculaire a été effectué avec pour référence, une seule molécule (la chaîne A) du modèle 1A5D. La recherche des six paramètres (trois pour le barycentre et trois pour l'orientation) ne se fait pas simultanément, car trop coûteuse en calcul. Tout d'abord, le programme recherche l'orientation du modèle, grâce à la fonction de Patterson comme il a été précédemment expliqué : c'est la rotation du modèle ; puis il recherche la position de ce dernier : c'est l'opération de translation. Les possibilités du remplacement moléculaire pour le jeu de données de la $D\gamma EH_2O$ sont montrées en tableau 5.4, dans lequel, les solutions de rotation sont traitées en premier, suivies par les solutions de translation. Les premières solutions de rotation et de translation sont clairement les meilleures et seront donc choisies pour la suite des traitements de données.

Solution de rotation	θ	ϕ	χ	α	β	γ	R_f/σ
Sol_RF	0.00	0.00	0.00	0.00	0.00	0.00	9.18
Sol_RF	21.72	176.95	46.08	108.51	16.66	294.61	4.40
Sol_RF	164.71	-88.80	19.02	172.02	5.00	169.62	3.54
Sol_RF	41.18	142.01	176.58	139.74	82.32	35.71	3.07

Solution de translation	α	β	γ	X_{fact}	Y_{fact}	Z_{fact}	TF/σ	$Corr.$
Sol_TF_1	0.00	0.00	0.00	0.263	0.076	0.310	24.15	0.633
Sol_TF_1	0.00	0.00	0.00	0.261	0.076	0.225	11.24	0.359
Sol_TF_1	0.00	0.00	0.00	0.262	0.074	0.127	11.21	0.368
Sol_TF_1	0.00	0.00	0.00	0.263	0.254	0.310	10.92	0.371

Tableaux 5.4 : Solutions proposées (pour $D\gamma Eh_2o$) par le programme *Molrep* pour le remplacement moléculaire avec 1 seule molécule. Les solutions pour la fonction de rotation sont dans la partie supérieure du tableau et les solutions pour la translation du modèle pour la 1^{ère} rotation sont dans la partie inférieure. Dans les 2 cas, les premières solutions seront choisies.

Après ces opérations, le programme *Refmac* est utilisé pour effectuer le premier affinement sur chacun des deux nouveaux modèles. Les facteurs R et R_{libre} obtenus après 10 cycles, sont représentés dans le tableau 5.5.

<i>Remplacement moléculaire</i>	Facteur R après le 1 ^{er} cycle d'affinement	Facteur R_{libre} après le 1 ^{er} cycle d'affinement
<i>avec 1 molécule</i>	40,2 %	41,9 %
<i>avec 2 molécules</i>	56,6 %	53,7 %

Tableau 5.5 : Facteur R et R_{libre} après le 1^{er} affinement numérique réalisé par *Refmac* pour les modèles avec 1 ou 2 molécules par maille.

L'affinement établi avec une seule molécule lors du remplacement moléculaire présente des statistiques meilleures de ~ 16 % pour les facteurs R et R_{libre} . Il est certain qu'il n'existe pas la même relation entre les molécules des modèles cristallisés pendant cette thèse et celui obtenu par Norledge et al. (1997). Mais cette observation sera évoquée dans le chapitre suivant, traitant de la structure de la γ E cristalline de rat à haute résolution.

5.4.5 Reconstructions et affinements

L'affinement des structures a été réalisé en de nombreuses étapes qui consistaient en une vérification de tous les acides aminés avec leur chaîne latérale complète (les modèles obtenus possédaient de nombreuses absences de chaînes latérales), puis en un affinement manuel grâce à l'inspection des cartes de densités électroniques $2|F_o|-|F_c|$ et $|F_o|-|F_c|$, et en l'addition des molécules d'eau. Les statistiques finales des quatre structures résolues sont présentées dans le tableau 5.6.

L'affinement a été réalisé par les deux programmes *Refmac* et *ShelX* de façon à, dans un premier temps, mettre en place et/ou corriger certains résidus sous un environnement informatique plus aisé, et par la suite, pour calculer les occupations partielles des résidus en double conformation. Les facteurs B , R et R_{free} sont différents et plus importants avec l'affinement avec *ShelX*. Ceci a, également, été observé dans le cas de protéinase K (Betzel et al. 2001).

Données	H γ Eh ₂ o		H γ Ed ₂ o		D γ Eh ₂ o		D γ Ed ₂ o	
Résolution (Å)	1,45-51,3		1,44-50,64		1,36-50,64		1,72-50,63	
λ utilisée (Å)	0,9393		0,9393		0,9393		0,9393	
Réflexions utilisées	29137		29159		36431		18135	
Complétude (%)	95,6 (89,8)		97,1 (90,3)		99,9 (93,7)		99,2 (89,6)	
Multiplicité	3,5 (2,1)		3,7 (2,5)		7,4 (6,1)		6,9 (5,8)	
I / σ (I)	7,0 (3,5)		6,1 (2,7)		5,6 (2,1)		4,7 (2,3)	
R sym (%)	7,6 (20,8)		8,8 (25,9)		6,7 (36,6)		8,1 (29,6)	
R _{fact.} (%)	Refmac	ShelX	Refmac	ShelX	Refmac	ShelX	Refmac	ShelX
	17,9	21,0	17,2	17,2	15,8	18,7	16,4	19,1
R _{libre} (%)	Refmac	ShelX	Refmac	ShelX	Refmac	ShelX	Refmac	ShelX
	20,0	25,4	19,0	21,8	19,9	21,3	19,2	22,3
Nombre d'atomes	1738		1768		1767		1732	
B moyen (Å ²)	21,1		18,1		16,9		22,2	
B moyen Atomes Protéine (Å ²)	20,3		16,9		15,9		21,5	
B moyen Molécules d'Eau (Å ²)	29,7		28,2		26,3		30,6	
rmsd Liaisons (Å)	0,016		0,017		0,017		0,018	
rmsd Angle (°)	1,689		1,766		1,768		1,761	

Tableau 5.6 : Statistiques des collectes des données et des affinements finaux calculés par *Refmac* et *ShelX*. Les données entre parenthèses sont les statistiques de la coquille de plus haute résolution. Les facteurs thermiques B sont été obtenus après affinement final par *ShelX*. rmsd signifie root mean square deviation.

Les modèles de départ obtenus après le premier affinement du corps rigide ($R_{\text{fact.}} \sim 40\%$ et $R_{\text{libre}} \sim 42,5\%$), présentent une chaîne principale bien positionnée excepté pour les résidus 116 à 119 (la glycine N terminale étant l'acide aminé numéro 1) et la partie C terminale comprenant les résidus 171 à 173. Les modèles n'ont pas de réelles divergences les uns par rapport aux autres mais certains détails de la structure sont plus visibles à haute résolution (1,36Å pour la D γ Eh2o) qu'à plus faible résolution (1,7Å pour la D γ Ed2o).

L'examen des cartes de densités électroniques et la reconstruction des structures après quelques cycles d'affinement ont donné la preuve que certaines chaînes latérales étaient désordonnées, mais la plupart des résidus sont bien définis et il est relativement aisé de positionner le modèle dans la densité électronique (voir figure 5.11).

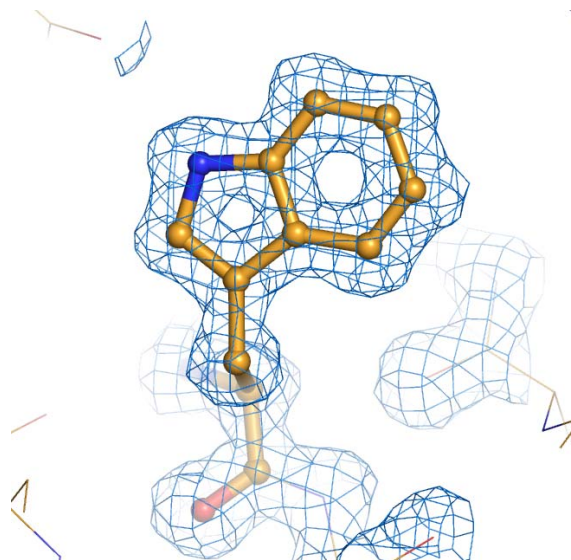


Figure 5.11 : Résidu tryptophane 42 où les deux cycles aromatiques sont bien positionnés dans la densité électronique $2|F_o| - |F_c|$ à 1σ (en bleu).

Pour toutes les structures, la totalité des chaînes latérales dont la densité électronique $2|F_o| - |F_c|$ est peu, voire mal définie, appartient à des résidus se trouvant en surface de la protéine. Dans certains cas, comme pour l'arginine 36, la densité ne montre qu'un léger désordre ou une forme étrange traduisant la double conformation du résidu (voir figure 5.12A). La densité électronique de ces groupes est pourtant améliorée tout au long de l'affinement avec l'ajout de double conformation dont l'occupation était de moitié. Dans d'autre cas, comme pour l'arginine 59, cette densité électronique est faible voir totalement absente suggérant alors la présence de conformères multiples (voir figure 5.12B). Cette absence de densité est la conséquence de deux raisons physiques importantes : le *mouvement thermique* et le *désordre*. Le *mouvement thermique* fait référence à la vibration d'un atome autour de sa position au repos. Le *désordre* lui se réfère à l'atome ou au groupe d'atomes qui n'occupe pas la même position dans toutes les mailles, ou dans

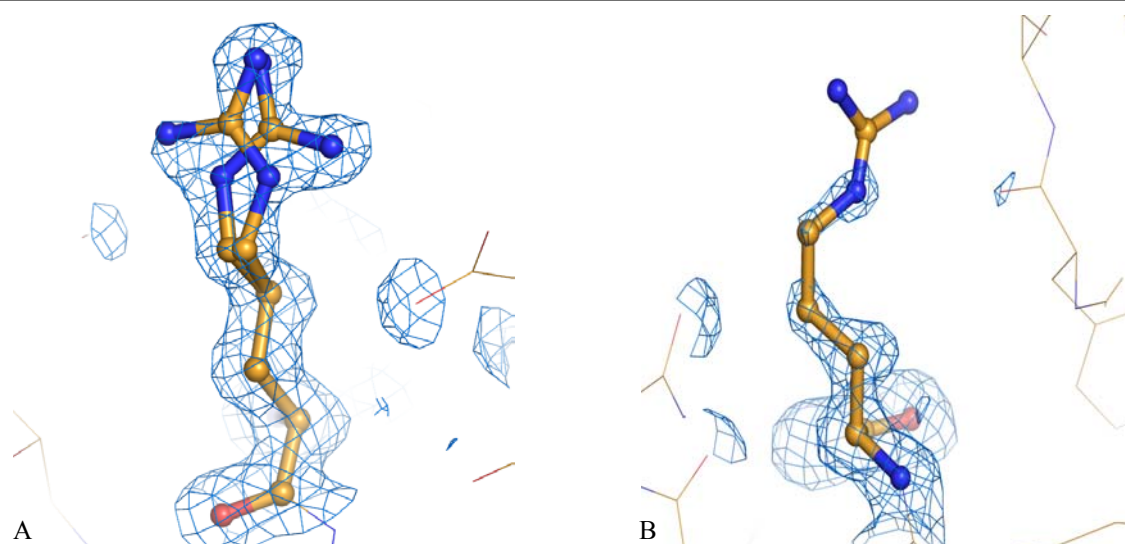


Figure 5.12 : Qualité de la densité électronique $2|F_o|-|F_c|$ à 1σ . (A) Le résidu arginine 36 a une densité électronique en forme de trèfle à trois feuilles traduisant la double conformation de ce résidu. (B) La densité électronique de l'arginine 59 est très mal définie ce qui démontre une forte flexibilité de cette chaîne latérale et donc des conformations multiples.

toutes les molécules de la maille. Le facteur de température B_j obtenu au cours de l'affinement reflète, à la fois, le mouvement thermique et le désordre d'un atome j . Dans certains cas, des résidus présentent des valeurs du facteur B_j s'éloignant complètement de la valeur moyenne de la molécule. Il arrive souvent que ces valeurs expriment le désordre de ces groupements qui peut être visible dans les cartes $2|F_o|-|F_c|$ et $|F_o|-|F_c|$. Si B_j est purement une mesure du déplacement moyen de l'atome j , alors pour le plus simple des mouvements thermiques avec des amplitudes équivalentes dans toutes les directions (appelé aussi vibration isotrope), B_j est relié à la grandeur de la vibration qui suit :

$$B_j = 8\pi^2 \{u_j^2\} = 79\{u_j^2\}$$

où $\{u_j^2\}$ est le déplacement moyen de l'atome à partir de sa position de repos.

Pour les petites molécules, il est courant d'obtenir des facteurs de température anisotropes pendant l'affinement, donnant une image des directions préférentielles de vibration pour chaque atome. Mais une description anisotrope requière six paramètres supplémentaires par atome. Dans toutes les données traitées, le nombre de paramètres à chercher, incluant les trois coordonnées atomiques, et les six paramètres thermiques par atomes, se rapproche ou dépasse le nombre de réflexions mesurées (Merritt 1999). Or, pour un bon affinement, le nombre de réflexions doit dépasser le nombre de paramètres désirés. Pour ces raisons, un affinement anisotrope n'a pu être réalisé. Les paramètres thermiques isotropes de température des γ E cristallines sont, en général, bas. Les valeurs de B sont variables le long de la séquence. Les zones où ce dernier est le plus élevé sont des zones exposées, en surface de la protéine. Le désordre, exprimé par ces hautes

valeurs, surtout des chaînes latérales, montre la flexibilité et la possibilité d'interaction de ces dernières avec le solvant ou d'autres protéines.

5.4.6 Qualité des modèles

Les données utilisées pour la détermination de la structure ont été résumées dans le tableau 5.6 et des exemples de densité électronique d'un point de vue générale sont montrés en figures 5.11 et 5.12. Les facteurs B moyens des chaînes protéiques et des molécules d'eau sont représentés dans le tableau 5.6. Les diagrammes de Ramachandran des différentes structures sont représentés en annexe E.1 et sont résumés dans le tableau 5.7 :

<i>Résidus placés dans :</i>	<i>HγEh₂O</i>	<i>HγEd₂O</i>	<i>DγEh₂O</i>	<i>DγEd₂O</i>
- une région favorable	90,6%	89,9%	91,3%	89,9%
- une région probable	9,4%	10,1%	8,7%	10,1%
- une région de façon gênéreuse	0%	0%	0%	0%
- une région défavorable	0%	0%	0%	0%

Tableau 5.7 : Statistiques issues des diagrammes de Ramachandran.

Tous les résidus sont placés de façon convenable, même les résidus C terminaux, dont la densité électronique n'est pas définie, et possèdent une géométrie correcte malgré la haute flexibilité de cette région.

5.5 Bilans

Dans ce chapitre, l'obtention de différentes formes de cristaux de γ E cristalline hydrogénée a été montrée. Une condition de cristallisation, formant des plaques bidimensionnelles, a été retenue pour sa reproductibilité et l'obtention de tailles importantes. A partir de ce moment, la cristallisation de la γ E cristalline a été entreprise sous forme hydrogénée ou deutériée et dissoute dans un solvant hydrogéné ou deutérié. Les collectes de données de diffraction aux rayons X ont permis d'obtenir des structures comparables à des résolutions allant de 1,71 à 1,36 Å. Les différents modèles sont de bonne qualité et sont discutés et comparés dans le chapitre 6, ainsi que l'influence de la deutériation sur les données cristallographiques.

Bibliographie :

- Betzl, C., Gourinath, S., Kumar, P., Kaur, P., Perbandt, M., Eschenburg, S., and Singh, T.P. 2001. Structure of a serine protease proteinase K from *Tritirachium album limber* at 0.98 Å resolution. *Biochemistry* **40**: 3080-3088.
- Garman, E. 2003. 'Cool' crystals: macromolecular cryocrystallography and radiation damage. *Curr Opin Struct Biol* **13**: 545-551.
- Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard. 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* **47 (Pt 2)**: 110-119.
- Kabsch, W. 1988. *J Appl. Cryst* **21**: 916-924.
- Kabsch, W. 1993. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J Appl Cryst*, **26**: 795-800.
- Laskowski, R., MacArthur, M., Moss, D., and Thornton, D. 1993. PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Cryst*, **26**: 283-291.
- Matthews, B.W. 1968. Solvent content of protein crystals. *J Mol Biol* **33**: 491-497.
- McRee, D. 1992. XtalView: A visual protein crystallographic software system for X11/XView. *J Mol Graph* **10**: 44-47.
- Merritt, E.A. 1999. Expanding the model: anisotropic displacement parameters in protein structure refinement. *Acta Crystallogr D Biol Crystallogr* **55 (Pt 6)**: 1109-1117.
- Murshudov, G.N., Vagin, A., and Dodson, E.J. 1997. Refinement of Macromolecular Structures by the maximum likelihood method. *Acta Cryst D* **53**: 240-255.
- Norledge, B.V., Hay, R.E., Bateman, O.A., Slingsby, C., and Driessen, H.P. 1997. Towards a molecular understanding of phase separation in the lens: a comparison of the X-ray structures of two high Tc gamma-crystallins, gammaE and gammaF, with two low Tc gamma-crystallins, gammaB and gammaD. *Exp Eye Res* **65**: 609-630.
- Perrakis, A., Harkiolaki, M., Wilson, K.S., and Lamzin, V.S. 2001. ARP/wARP and molecular replacement. *Acta Crystallogr D Biol Crystallogr* **57**: 1445-1450.
- Powell, H.R. 1999. The Rossmann Fourier autoindexing algorithm in MOSFLM. *Acta Crystallogr D Biol Crystallogr* **55 (Pt 10)**: 1690-1695.
- Sheldrick, G.M., and Schneider, T.R. 1997. SHELXL : High Resolution Refinement. *Methods Enzymol* **277**: 319-343.
- Vagin, A., and Teplyakov, A. 2000. An approach to multi-copy search in molecular replacement. *Acta Crystallogr D Biol Crystallogr* **56 Pt 12**: 1622-1624.

Chapitre 6

Structure de la γ E cristalline à 1,32 Å et comparaison des formes hydrogénées et deutériées en H₂O et en D₂O

Abstract

The molecular models of the hydrogenated γ E crystallin in H₂O (H γ Eh₂O) and in D₂O (H γ Ed₂O) and of the deuterated γ E crystallin in H₂O (D γ Eh₂O) and in D₂O (D γ Ed₂O), resolved at 1.45, 1.44, 1.36 and 1.72 Å resolution respectively are compared in this chapter.

The comparison is based on the root mean square (RMS) deviation of the atomic coordinates and on the thermal factor (B). Comparison of the backbone atoms of the different isotopic forms of the protein yields an rms deviation of 0.156 Å. The largest concerted differences between backbone atoms are found at residues 171-173. These residues constitute the flexible C terminal part which previous investigators have found to be disordered. The maximum deviation on the side chains with the H γ Eh₂O model reaches 0.158 Å. The temperature factor variation of the structures at identical resolution (H γ Eh₂O H γ Ed₂O and D γ Eh₂O) depends on the isotope labelling. The deuterated structure or with D₂O solvent have a lower temperature factor. The same comparison is made with similar water molecules and the conclusions are identical.

The molecular model of D γ Eh₂O crystalline was studied at 1.36 Å and some new relevant details are described in this chapter. The *cd* loop in motif 3, which was considered as a flexible part in the previous structure at 2.3 Å, presents two conformations, different from the two conformations existing between the cryo- and non cryo-crystallin proteins. The protein also contains an arginine-glycine-aspartate (RGD) peptide which is characteristic of cell adhesion motif found in cell adhesion pathway. And some other details are discussed like disulfide component, the stacking of tyrosine on the protein surface. A detailed comparative analysis of the solvent around the γ E crystallin concludes this chapter.

6.1 Introduction

A ce jour, malgré le grand nombre d'études sur les γ cristallines, seules deux structures ont été étudiées à haute résolution. La γ_B cristalline de bœuf a été résolue en 1996 par Kumaraswamy à 1,2 Å, et la γ_D humaine, ainsi qu'un de ses mutants, ont été plus récemment décrites à 1,25 et à 1,15 Å respectivement (Basak et al. 2003). Au cours de cette thèse, la structure de γ_E cristalline de rat sous différentes formes isotopiques, a été résolue entre 1,36 et 1,72 Å.

Après avoir traité les différents jeux de données obtenus par diffraction aux rayons X, les structures qui en résultent vont être examinées en différents points, dans ce chapitre. De façon à démontrer que la perdeutériation et l'échange isotopique au niveau du solvant (obligatoire pour l'étude neutronique) n'ont pas eu d'influence d'un point de vue structural, il sera établi une comparaison, à des résolutions quasi équivalentes, des structures des protéines hydrogénées et deutériées, en solvant H_2O et en D_2O , déterminées par diffraction des rayons X. En premier lieu, les structures vont être comparées au niveau de leur forme (surfaces et volumes). Puis cette étude sera basée, à la fois, sur la différence de localisation dans l'espace des chaînes principales et latérales des modèles hydrogénés et deutériés en H_2O et en D_2O . Enfin, un comparatif des facteurs B thermiques complètera cette analyse.

Par la suite, la structure de la γ_E cristalline de rat à 1,36 Å de résolution va être étudiée et discutée. Certains détails et certaines différences avec la structure déjà publiée à 2,3 Å (Norledge et al. 1997) sont désormais plus visibles et plus captivants à cette résolution. Par exemple, au cours de cette thèse l'existence d'une double conformation de la chaîne principale au niveau de la boucle cd du motif 3, sera examinée. Enfin, le solvant, l'eau dans le cas de notre étude, représente environ 37 % du volume de la maille cristalline (voir tableau 5.3), correspondant au niveau d'hydratation des protéines dans les cellules qui varie de 30 à 60 % (Matthews 1975). Sans perdre l'idée qu'un des objectifs de cette thèse, est de positionner les molécules d'eau structurées autour de ces mêmes protéines et de comprendre leur rôle au niveau du cristallin, une étude du réseau d'hydratation présent autour de ces protéines terminera ce chapitre. Cette étude démontrera que malgré une haute résolution, qui permet de localiser avec précision, quasiment l'ensemble (95,57 %) des atomes de carbone, d'oxygène, d'azote et de soufre, il est impossible de visualiser la totalité des atomes d'hydrogène (soit 47,70% des atomes) qui jouent un rôle déterminant pour la stabilité et les interactions inter- et intra-moléculaires de la protéine. Ce qui appuie la nécessité d'une étude par diffraction neutronique.

6.2 Comparaison des différentes structures résolues

Pour déterminer si la perdeutériation ou le solvant ne produisent pas de changements structuraux au niveau atomique, une étude comparative des quatre structures a été entreprise.

6.2.1 Surfaces et Volumes

Les structures semblent avoir des formes compactes identiques. La comparaison des surfaces de contact, des surfaces accessibles au solvant et du volume moléculaire (calculé par *Grasp*) a révélé une très grande similitude entre les structures (voir tableau 6.1). La surface accessible au solvant de tous les atomes de la protéine a été calculée en utilisant le programme *Surface* (CCP4) qui utilise l'algorithme de Lee et Richards (1971) qui définit une surface moléculaire comme la surface résultant du centre d'une sphère type (de 1,4 Å de rayon) roulant sur la surface de van der Waals d'une molécule. La surface de contact est la surface de van der Waals de la molécule en contact avec cette même sphère de rayon 1,4 Å. Le volume moléculaire correspond au volume occupé par la molécule suivant les paramètres de van der Waals.

Paramètres	<i>HγE h₂O</i>	<i>HγE d₂O</i>	<i>DγE h₂O</i>	<i>DγE d₂O</i>	<i>1A5D</i>
<i>Surface accessible au solvant (en Å²) (Surface)</i>	8565,9 (-)	8539,7 (- 0,3 %)	8539,8 (- 0,3 %)	8566,4 (0 %)	8226,2 (- 4,0 %)
<i>Surface de contact (en Å²) (Surface)</i>	2563,1 (-)	2553,9 (- 0,4 %)	2555,5 (- 0,3)	2564,2 (0 %)	2467,4 (- 3,7 %)
<i>Volume moléculaire (en Å³) (Grasp)</i>	24893,1 (-)	24933,6 (+0,16%)	24688,6 (-0,82%)	24611,8 (-1.13%)	24078,9 (-3,27%)

Tableau 6.1 : Comparatif des surfaces et volumes des structures étudiés et de la structure déjà résolue à 2,3 Å. Le nombre entre parenthèse est le % de différence par rapport à la protéine *H γ E h₂O*. Pour la structure *1A5D*, les résultats ont été moyennés pour une seule chaîne.

Les surfaces accessibles au solvant, les surfaces de contact et les surfaces moléculaires sont très proches entre les structures résolues pendant cette thèse. Il existe moins de 0,4 % de différence entre elles. Les volumes moléculaires sont très voisins aussi. La structure de la *D γ E d₂O* montre la plus grande variation quand les quatre structures sont comparées mais il est important de préciser que cette structure possède une moins bonne résolution et des statistiques plus mauvaises (voir tableau 5.6).

En revanche, les différences sont plus importantes ($\sim 4\%$) avec la structure précédemment résolue (code PDB : 1A5D) par Norledge et al (1997). Elles sont certainement la conséquence de résolutions différentes et de conditions de cristallisation légèrement différentes (voir tableau 5.1). Le modèle déjà publié a été cristallisé dans un groupe d'espace $P12_11$ alors que toutes les structures de cette thèse sont en $P2_12_12_1$. Il est donc compréhensible que les surfaces de contact, ainsi que les surfaces accessibles au solvant, soient différentes.

6.2.2 Déviation des chaînes

A partir des données cristallographiques, il est possible de comparer la déviation standard (rmsd) des atomes dans l'espace par rapport à un des modèles (calculée par le programme *Lsqkab* de CCP4). Le modèle, choisi arbitrairement, avec lequel ont été comparées toutes les protéines, est la structure de la protéine hydrogénée en solvant hydrogéné ($H\gamma E h_2O$). Le modèle de la γE cristalline à $2,3\text{\AA}$ (Norledge et al. 1997) a été aussi comparé à titre indicatif mais les données ne sont pas tellement révélatrices étant donné que le groupe d'espace et, surtout, la résolution ne sont pas les mêmes. Les déviations des chaînes principales et des chaînes latérales sont représentées en figures 6.1 et 6.2, respectivement. Pour la chaîne principale, la déviation le long de la séquence peptidique est assez uniforme excepté pour deux positions : la partie C terminale qui est très flexible, et la boucle variable entre les brins βc et d dans le motif 3 (résidus 114 à 120). Il est important de préciser que, dans les structures étudiées dans cette thèse et la résolution à laquelle elles ont été résolues, les acides aminés forment deux boucles de conformation distincte éloignées l'une de l'autre de $4,47\text{\AA}$ pour les $C\alpha$ de l'histidine 116. Sans la double conformation de cette boucle (la conformation A est conservée pour la comparaison), la déviation standard moyenne des chaînes principales ne dépasse pas $0,36\text{\AA}$ (par rapport à 1A5D) dans toutes les comparaisons (voir figure 6.1E et 6.2E). En revanche, les structures résolues pendant la thèse, entre $1,36\text{\AA}$ et $1,72\text{\AA}$, ne présentent qu'une déviation maximum de $0,15\text{\AA}$ observée sur les chaînes latérales. Les zones les plus divergentes sont représentées en couleur sur les figures 6.1.C et 6.2.C. Les parties, où la déviation de la chaîne principale est la plus importante, ne font pas partie des zones formant des structures secondaires (voir figure 6.1B) : le peptide connecteur entre les deux domaines et des boucles entre les motifs en clés grecques.

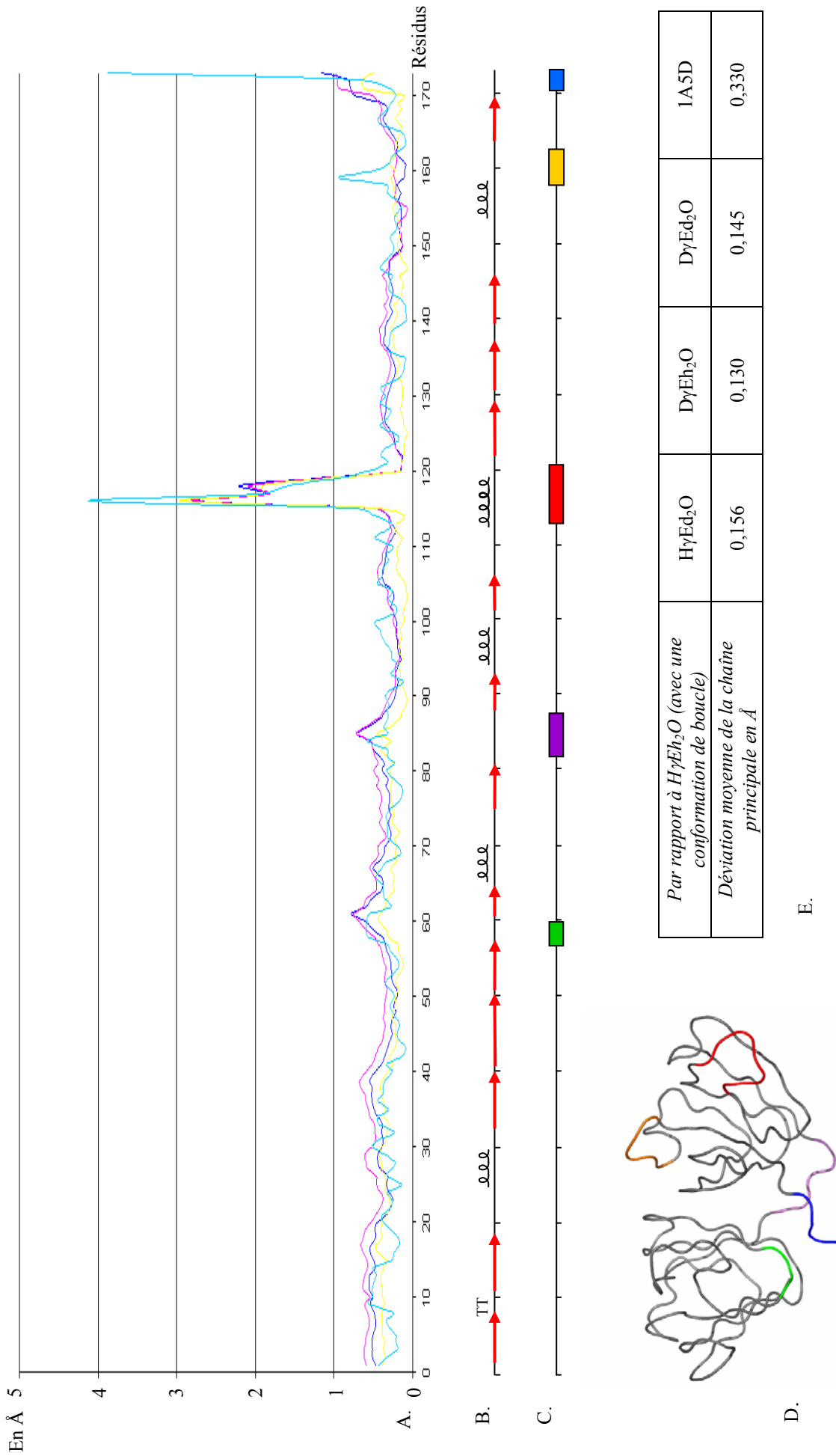


Figure 6.1: RMS déviation de la chaîne principale des γE cristalline de rat étudiées. A) Graphique représentant la déviation standard de la chaîne principale des protéines HyED₂O (—), DγEH₂O (—), DγED₂O (—) et le modèle de la PDB 1A5D (—) par rapport à la protéine HyED₂O (calculs exécutés par Lsqkab). B) Représentation linéaire de la structure secondaire de la γE cristalline (→, feuillets β et hélice ↱). C) Définition des domaines pouvant varier de façon importante. D) Localisation des domaines flexibles sur la structure 3D de la γE cristalline. E) Tableau indiquant la déviation standard de la chaîne principale des différentes structures par rapport à HyED₂O.

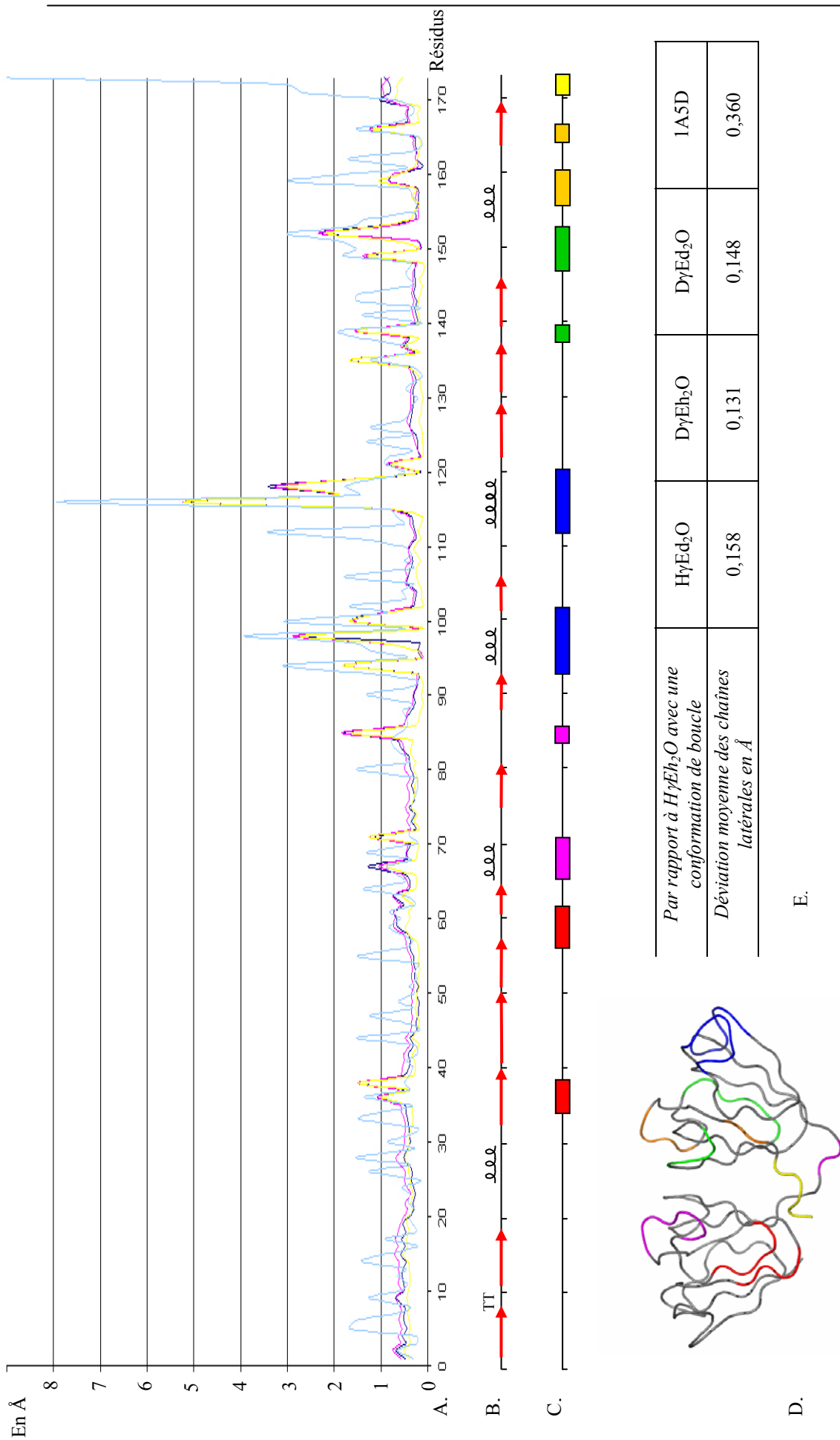


Figure 6.2 : RMS déviation des chaînes latérales des γE cristalline de rat étudiées. A) Graphique représentant la déviation standard de la chaîne latérale des protéines HyED₂O (—), DγEH₂O (---), DγED₂O (---) par rapport à la protéine HyEH₂O (calculs exécutés par Lsqkab). B) Représentation linéaire de la structure secondaire de la γE cristalline (→, feuillets β et hélice). C) Définition du domaine pouvant varier de façon importante. D) Localisation des domaines sur la structure 3D vue de face (a) et de l'arrière (b). E) Tableaux indiquant la déviation standard de la chaîne principale des différentes structures par rapport à HyED₂O.

Les dissemblances des chaînes latérales sont beaucoup plus difficiles à analyser mais il est aussi possible de remarquer que celles-ci sont principalement dues à des grandes chaînes latérales qui se trouvent en surface de la protéine. Les positions des résidus enfouis sont très semblables du fait de la spécificité de former des zones hydrophobes permettant une forte stabilité de la structure. Mais la divergence des chaînes latérales de surface, du modèle proposé par Norledge et al. (1997) peut s'expliquer, d'une part, par une résolution différente, et d'autre part, par le fait que la protéine a été cristallisée dans un groupe d'espace différent des modèles proposés dans cette thèse. Les contacts cristallins ne vont pas forcément être les mêmes, ce qui va éventuellement modifier la conformation des chaînes latérales faisant ces interactions.

Toutes les protéines, qu'elles soient natives ou recombinantes, hydrogénées ou deutériées, en solvant H₂O ou D₂O, peuvent être considérées comme similaires si l'on se réfère aux précédentes études sur la perdeutériation. Les pyrophosphatases inorganiques de levure (Y-PPase), hydrogénée et deutériée ont présenté une déviation standard de 0,15 Å pour les C^α (Tuominen et al. 2004). Mais cette conservation de structure a été aussi observée pour des protéines dont le solvant hydrogéné a été échangé avec un deutérié. Par exemple, le r.m.s.d des C^α pour la ribonucléase était de 0,2 Å (Wlodawer et al. 1986) ; celui de l'insuline pour tous les atomes était de 0,36 Å (Wlodawer et al. 1989).

6.2.3. Facteur thermique

Les données cristallographiques apportent aussi des informations sur l'agitation thermique de chaque atome et donc sur chacun des résidus protéiques. Le facteur thermique *B*, précédemment expliqué, a été étudié pour chaque résidu des différentes structures (voir figures 6.3 et 6.4). Ce facteur *B* est disponible pour les chaînes principales, les chaînes latérales et donc, pour chaque résidu dans son entier.

Les facteurs *B* après affinement par *ShelX*, sont plus importants d'environ 5,5 Å² que ceux après l'affinement par *Refmac*. Il a été énoncé, dans le chapitre 5 et pour la protéinase K (Betz et al. 2001), le même phénomène avec les facteurs *R* et *R libre*.

Les courbes sont similaires et suivent les mêmes fluctuations tout au long de la séquence mais à des niveaux différents : les facteurs *B* des deux systèmes homogènes (protéine hydrogénée en H₂O et protéine deutériée en D₂O) sont plus élevés que ceux des systèmes hétérogènes (protéine hydrogénée en D₂O et protéine deutériée en H₂O).

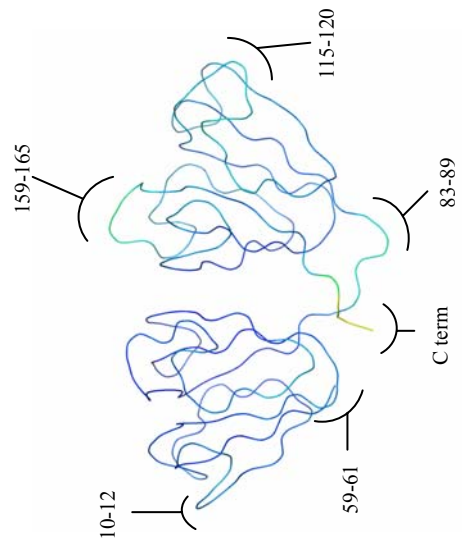
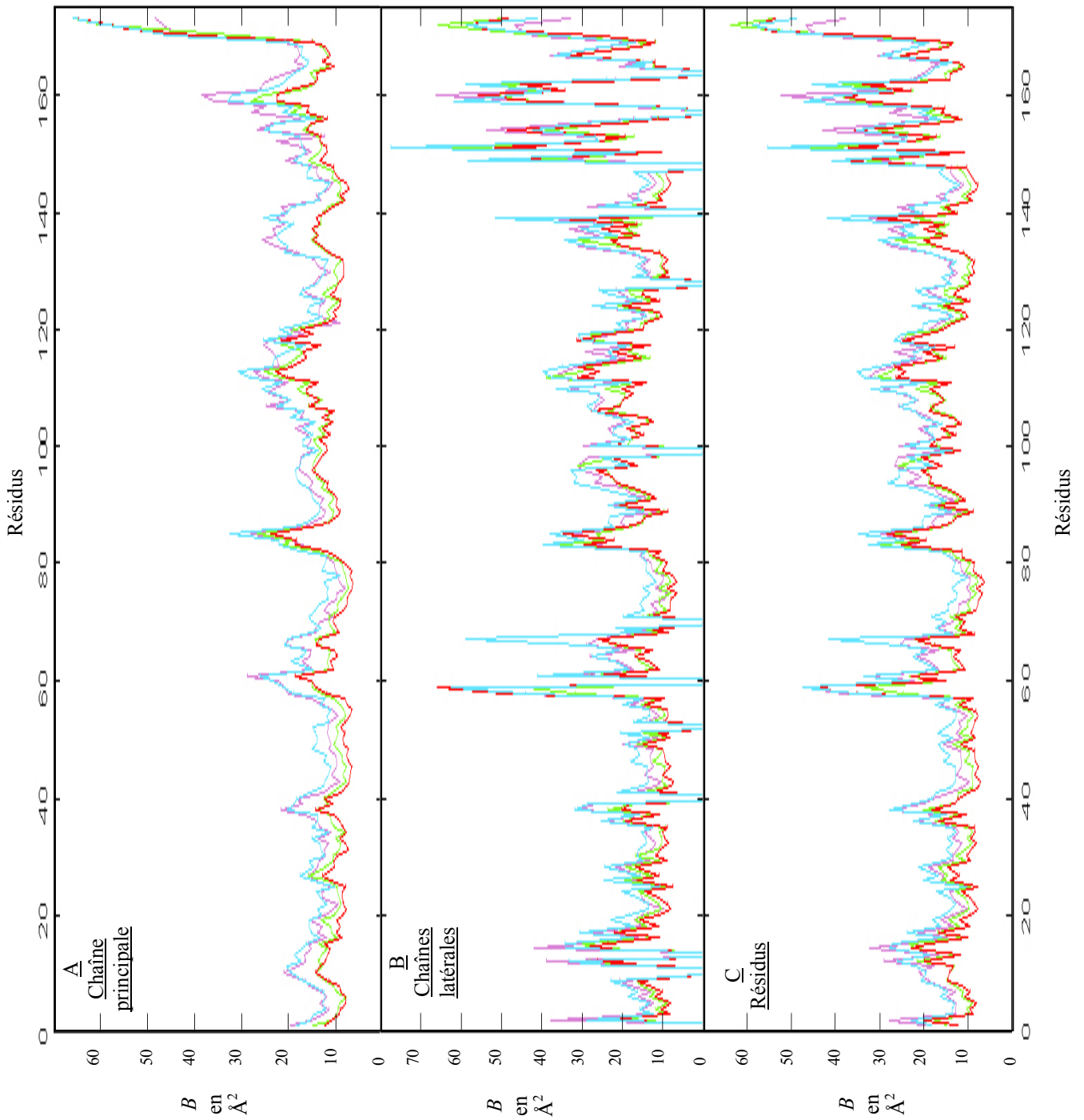


Figure 6.3 : Représentation de la chaîne principale suivant le facteur B . La couleur bleue signifie un facteur isotropique faible puis la couleur vire au bleu ciel, puis vert et enfin jaune et rouge quand celui-ci augmente.

Figure 6.4 : Représentations des facteurs B des différentes structures tout au long de la chaîne principale (A), au niveau des chaînes latérales (B), et suivant les résidus (C). Sont représentés : en violet la protéine hydrogénée γ E en solvant hydrogéné H_2O (HyEh_2O), en vert la $\text{H}\gamma\text{Ed}_2\text{O}$, en rouge la $\text{D}\gamma\text{Eh}_2\text{O}$, et en bleu la $\text{D}\gamma\text{Ed}_2\text{O}$.

Les paramètres thermiques isotropes de température des γE cristallines sont, en général, bas. Les valeurs de B sont variables le long de la séquence. Les zones où ce dernier est le plus élevé sont des zones exposées, en surface de la protéine. Le désordre, exprimé par ces hautes valeurs, surtout des chaînes latérales, montre la flexibilité et la possibilité d'interaction de ces dernières avec le solvant ou d'autres protéines. Les peptides exposés sont (voir figure 6.3) :

- les régions 10-12 et 159-165 dont les résidus ont des chaînes latérales longues telles que les arginines.
- les environs de la zone 59-61 qui est un tripeptide RGD (Arginine Glycine Aspartate) connu pour être impliqué dans des processus d'adhésion cellulaire (Pierschbacher et Ruoslahti 1987) où les chaînes latérales de l'arginine et de l'aspartate jouent un rôle prépondérant dans l'interaction peptidique.
- la région 83-89, correspondant au polypeptide de liaison entre les deux domaines, possède une légère flexibilité pour que les deux domaines s'interconnectent correctement.
- les acides aminés 115-120 qui ont été supposés responsables des différences de séparation de phase entre les groupes de basses et hautes températures de transition (Norledge et al. 1997).
- et la partie C terminale, de 170-173, connue comme une région extrêmement flexible (Kumaraswamy 1996), il était donc attendu que le facteur B de cette partie soit élevé.

Quand les deux systèmes, où la protéine hydrogénée est présente, sont pris en compte, un effet de solvant est observable. En effet, le système en D_2O présente un facteur B moyen moins important ($18,1 \text{ \AA}^2$) que celui en H_2O ($21,1 \text{ \AA}^2$). Sachant que le deutérium possède une masse plus grande que celle de l'hydrogène, les énergies de liaisons impliquant un deutérium avec l'hydrogène seront plus fortes, et la fréquence de vibration de ces liaisons sera plus faible (Benjamin et Benson 1962). Une liaison mettant en jeu un deutérium en sera moins flexible, donc plus rigide. Donc le D_2O , par son échange des hydrogènes avec les deutériums du solvant, augmente la rigidité de la structure native des protéines (Cioni et Strambini 2002). Il est donc normal que les structures cristallographiques aient des facteurs B moins importants en D_2O démontrant une flexibilité moindre de la protéine.

Si la comparaison des facteurs B des protéines hydrogénées (en violet) et deutériées (en rouge) en H_2O est faite, un effet sur la flexibilité de la protéine généré par la perdeutériation est visible. En effet, le facteur moyen B de la protéine hydrogénée ($21,1 \text{ \AA}^2$) en H_2O est plus élevé que celui de la protéine deutériée en H_2O ($16,9 \text{ \AA}^2$). La perdeutériation permet le changement

isotopique des hydrogènes non échangeables avec des atomes de deutérium. Ces nouvelles liaisons, pour les mêmes raisons qu'énoncées précédemment, vont alors aussi "rigidifier" la structure et donc vont abaisser les facteurs B : la structure en sera alors moins flexible.

Mais quand les deux protéines deutériées sont considérées, il apparaît aussi un effet de solvant mais celui ci est à l'opposé de ce qui est attendu : le facteur B de la protéine deutériée en D_2O ($22,2 \text{ \AA}^2$) est plus important que celui de la protéine deutériée en H_2O ($16,9 \text{ \AA}^2$). Ce résultat est surprenant mais il est nécessaire d'émettre certaines réserves. En effet, les résolutions des structures ne sont pas identiques ($1,36 \text{ \AA}$ pour la $D\gamma E h_2o$ et $1,72 \text{ \AA}$ pour la $D\gamma E d_2o$), et dans l'équation 5.16, l'intensité des diffractions est proportionnelle à B et à la résolution d . De plus, les statistiques des données de la protéine deutériée en D_2O sont de moins bonne qualité (voir tableau 5.6). Une comparaison et une discussion sur ces différences ne seraient que pures hypothèses voire spéculations mais si les données de la $D\gamma E d_2o$ étaient limitées à $1,4 \text{ \AA}$, il est fort possible que le facteur thermique B soit inférieur à toutes les précédentes.

Le désordre conformationnel exprime, comme il a été précisé auparavant, souvent les multiples conformations d'un acide aminé. Les différentes structures qui ont été résolues au cours de cette thèse, présentent de nombreuses doubles conformations. Une vingtaine d'acides aminés ont été trouvés dans une double conformation pendant la reconstruction du dernier affinement (voir tableau 6.2).

<i>Structure</i>	<i>HγE h₂o</i>	<i>HγE d₂o</i>	<i>DγE h₂o</i>	<i>DγE d₂o</i>
<i>Nbre d'ac. aminés avec une d. conf.</i>	21	19 (+R139 -E149 -M159 -L166)	21 (+Y97 +R139 +M170 -E149 -R153 -M159)	19 (+M170 -E149 -R153 -M159)

Tableau 6.2 : Nombre d'acides aminés ayant une double conformation dans les différentes structures. Entre parenthèses se trouvent les résidus différents de la structure $H\gamma E h_2o$.

La totalité de ses résidus est en surface de la protéine. Dans chacun des cas, le désordre est apparemment limité à deux positions distinctes. L'occupation de ces sites a été déterminée par le programme *ShelX*. Toute nouvelle liaison hydrogène possible avec le solvant a été vérifiée. L'occupation de ces molécules d'eau a été mis en accord avec l'occupation du résidu avec lequel elle était liée. Les différences d'occupations des conformations entre les modèles, quand elles sont comparables, sont montrées dans le tableau 6.3.

Occupation des différentes conformations				
Résidus ou groupes de résidus	H γ Eh ₂ O	H γ Ed ₂ O	D γ Eh ₂ O	D γ Ed ₂ O
R36	1 ^{ere} : 0,4 2 ^{nde} : 0,6	1 ^{ere} : 0,4 2 ^{nde} : 0,6	1 ^{ere} : 0,4 2 ^{nde} : 0,6	1 ^{ere} : 0,4 2 ^{nde} : 0,6
D38	1 ^{ere} : 0,4 2 ^{nde} : 0,6	1 ^{ere} : 0,5 2 ^{nde} : 0,5	1 ^{ere} : 0,4 2 ^{nde} : 0,6	1 ^{ere} : 0,5 2 ^{nde} : 0,5
F71	<u>1^{ere} : 0,7</u> <u>2^{nde} : 0,3</u>	1 ^{ere} : 0,5 2 ^{nde} : 0,5	1 ^{ere} : 0,6 2 ^{nde} : 0,4	1 ^{ere} : 0,5 2 ^{nde} : 0,4
S85	1 ^{ere} : 0,3 2 ^{nde} : 0,7	1 ^{ere} : 0,3 2 ^{nde} : 0,7	<u>1^{ere} : 0,5</u> <u>2^{nde} : 0,5</u>	1 ^{ere} : 0,3 2 ^{nde} : 0,7
R94	1 ^{ere} : 0,5 2 ^{nde} : 0,5	1 ^{ere} : 0,5 2 ^{nde} : 0,5	1 ^{ere} : 0,5 2 ^{nde} : 0,5	1 ^{ere} : 0,5 2 ^{nde} : 0,5
R98	1 ^{ere} : 0,5 2 ^{nde} : 0,5	<u>1^{ere} : 0,8</u> <u>2^{nde} : 0,2</u>	1 ^{ere} : 0,5 2 ^{nde} : 0,5	1 ^{ere} : 0,5 2 ^{nde} : 0,5
Q100	1 ^{ere} : 0,5 2 ^{nde} : 0,5	<u>1^{ere} : 0,8</u> <u>2^{nde} : 0,2</u>	1 ^{ere} : 0,5 2 ^{nde} : 0,5	1 ^{ere} : 0,5 2 ^{nde} : 0,5
M101	1 ^{ere} : 0,6 2 ^{nde} : 0,4	1 ^{ere} : 0,7 2 ^{nde} : 0,3	1 ^{ere} : 0,7 2 ^{nde} : 0,3	1 ^{ere} : 0,6 2 ^{nde} : 0,4
115-119 + 121 (boucle <i>cd</i>)	1 ^{ere} : 0,6 2 ^{nde} : 0,4	1 ^{ere} : 0,6 2 ^{nde} : 0,4	1 ^{ere} : 0,6 2 ^{nde} : 0,4	1 ^{ere} : 0,5 2 ^{nde} : 0,5
M135	<u>1^{ere} : 0,7</u> <u>2^{nde} : 0,3</u>	1 ^{ere} : 0,4 2 ^{nde} : 0,6	1 ^{ere} : 0,6 2 ^{nde} : 0,4	1 ^{ere} : 0,4 2 ^{nde} : 0,6
R152	1 ^{ere} : 0,5 2 ^{nde} : 0,5	1 ^{ere} : 0,5 2 ^{nde} : 0,5	1 ^{ere} : 0,6 2 ^{nde} : 0,4	1 ^{ere} : 0,4 2 ^{nde} : 0,6

Tableau 6.3 : Occupation des différentes conformations des résidus ayant 2 conformations dans les différents modèles.

Ce tableau de données montre qu'il existe des différences d'occupation arrondie, entre les cristaux de protéine hydrogénée et deutériée, en H₂O et en D₂O. Il est indispensable de rappeler que l'existence de double conformation signifie que la barrière énergétique séparant les deux conformations est très faible (voir figure 6.5). Cette limite est tellement petite que des variations de 0,1 d'occupation n'ont pas tellement de signification. Dans ce cas, seulement cinq résidus peuvent présenter de réelles différences d'occupation entre les structures : la phénylalanine 71, la sérine 85, l'arginine 98, la glutamine 100 et la méthionine 135 (en gras dans le tableau 6.3). Il semblerait qu'au niveau de ces acides aminés, un apport énergétique a modifié l'équilibre de forces qui existe au niveau des autres structures. Cet apport peut être sous forme d'une liaison hydrogène supplémentaire, d'une interaction hydrophobe plus grande. Mais l'origine de cet apport

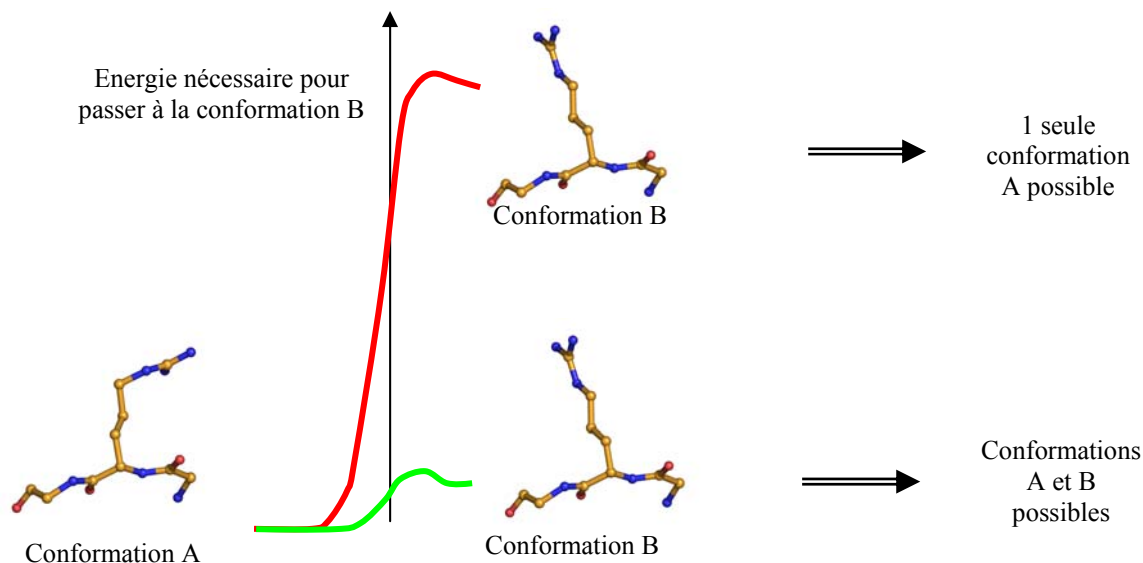


Figure 6.5 : Représentation schématique de l'existence des doubles conformations d'acides aminés comme cette arginine. Pour passer de la conformation A à la conformation B, le résidu a besoin d'énergie. Si cette barrière énergétique est faible (courbe verte), il y aura deux conformations, en revanche si cette énergie nécessaire est trop grande (courbe rouge), une seule conformation sera possible.

énergétique est difficilement identifiable et il serait imprudent de le justifier par la deutériation ou le solvant. Il ne faut pas oublier que ces structures ne sont que des modèles issus de cristaux non identiques, même s'ils poussent dans des conditions similaires. Au cours de ma thèse, il est arrivé que des cristaux de protéine hydrogénée en H_2O fournissent également des occupations différentes pour les mêmes acides aminés. Il serait plus sage de dire que le solvant deutérié ou la perdeutériation ne sont pas, uniquement, la cause de la formation ou la disparition de double conformation, ou encore l'adoption d'une conformation aux dépend d'une autre.

6.3 Structure de la γ E cristalline

La totalité de l'étude sur la structure a été effectuée à partir de la structure obtenue à la plus haute résolution et dont les statistiques sont les meilleures. La protéine deutériée en solvant hydrogéné ($D\gamma E$ H_2O) qui est similaire aux autres structures, est choisie.

6.3.1 Quatre motifs en clef grecque organisés symétriquement

L'unité asymétrique trouvée pour chaque structure cristalline comprend une molécule contrairement à ce qui avait été trouvé auparavant (Norledge et al. 1997). En revanche, la forme globulaire générale, les structures secondaires et tertiaires des protéines sont très similaires avec les γ cristallines déjà résolues. Chaque molécule de γ E cristalline de cette étude est

composée d'un domaine N terminal, compris entre les résidus Glycine 1 et Serine 82, d'un peptide connecteur (résidus Histidine 83 à Histidine 87) un domaine C terminal, formé entre les résidus Arginine 88 et la Méthionine 170, et une courte extension C terminal comprenant les résidus Acétate 171, Phénylalanine 172 et Tyrosine 173. Chacun des domaines N et C terminaux possède deux motifs en « clés grecques », stabilisés par des liaisons hydrogène (voir figure 6.7) et reliés par un axe en pseudo dyade, tout comme les deux domaines sont reliés par un autre axe en pseudo dyade (voir figure 6.6). Ces axes sont identiques pour les quatre structures étudiées.

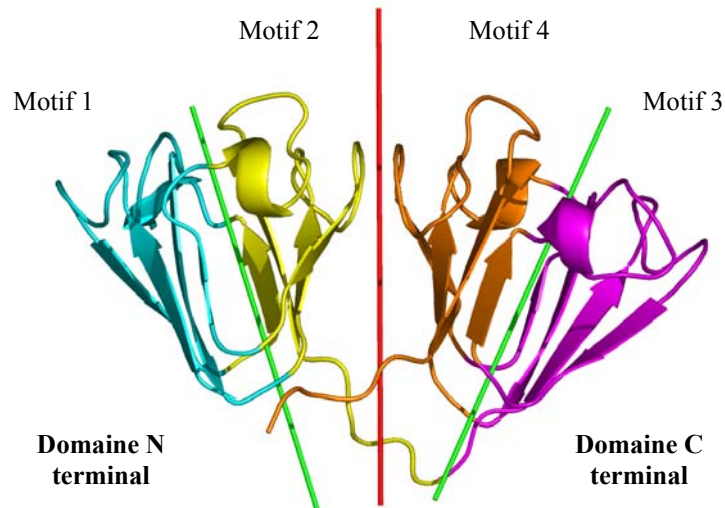


Figure 6.6 : Structure générale des γ cristallines résolues et de leurs symétries internes. Les motifs en clé grecque sont décrits de différentes couleurs : les motifs 1 et 2 du domaine N terminal sont en cyan et jaune respectivement; et les domaines 3 et 4 du côté C terminal sont représentés en magenta et orange. Les deux domaines sont reliés entre eux par un axe « pseudo dyade » en rouge de façon à ce que les motifs 2 et 4, et les motifs 1 et 3 se superposent les uns les autres. Deux autres axes du même type (en vert) relient le motif 1 au motif 2 et le motif 3 au motif 4 (White et al. 1989).

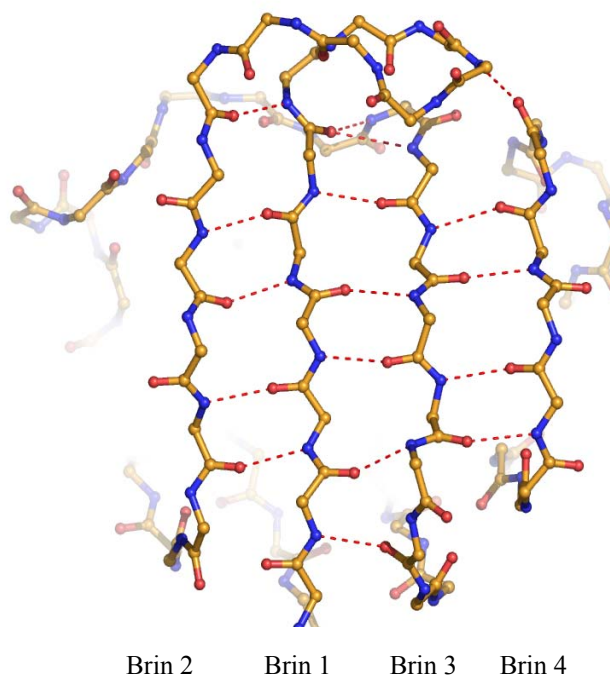


Figure 6.7 : Représentation du motif 1 en clé grecque stabilisé par les liaisons hydrogène (en pointillés rouge) inter brins β de façon antiparallèle.

Comme les précédentes γ cristallines déjà cristallisées, le solvant interagit essentiellement avec les motifs 1 et 3 alors que les motifs 2 et 4 stabilisent la structure globulaire en enfouissant des zones hydrophobes entourées de résidus polaires qui effectuent des interactions électrostatiques.

6.3.2 Stabilisation de la structure par interactions entre les deux domaines

Les domaines N et C terminaux sont associés de façon covalente par le peptide connecteur entre les motifs 2 et 3. Ils sont aussi liés ensemble par des interactions non covalentes par des résidus hydrophobes des motifs 2 et 4 et d'acides aminés polaires (voir figure 6.8). Ces résidus forment un cœur hydrophobe entouré d'interactions de type électrostatique et des liaisons hydrogène, comme il a déjà été observé dans toutes les structures des γ cristallines (Slingsby et Clout 1999). Il existe aussi des interactions hydrophobes intramoléculaires permettant la stabilisation de ces deux domaines (exemple la tyrosine 55 qui interagit avec la glutamine 142), qui ne sont pas au niveau de la zone hydrophobe.

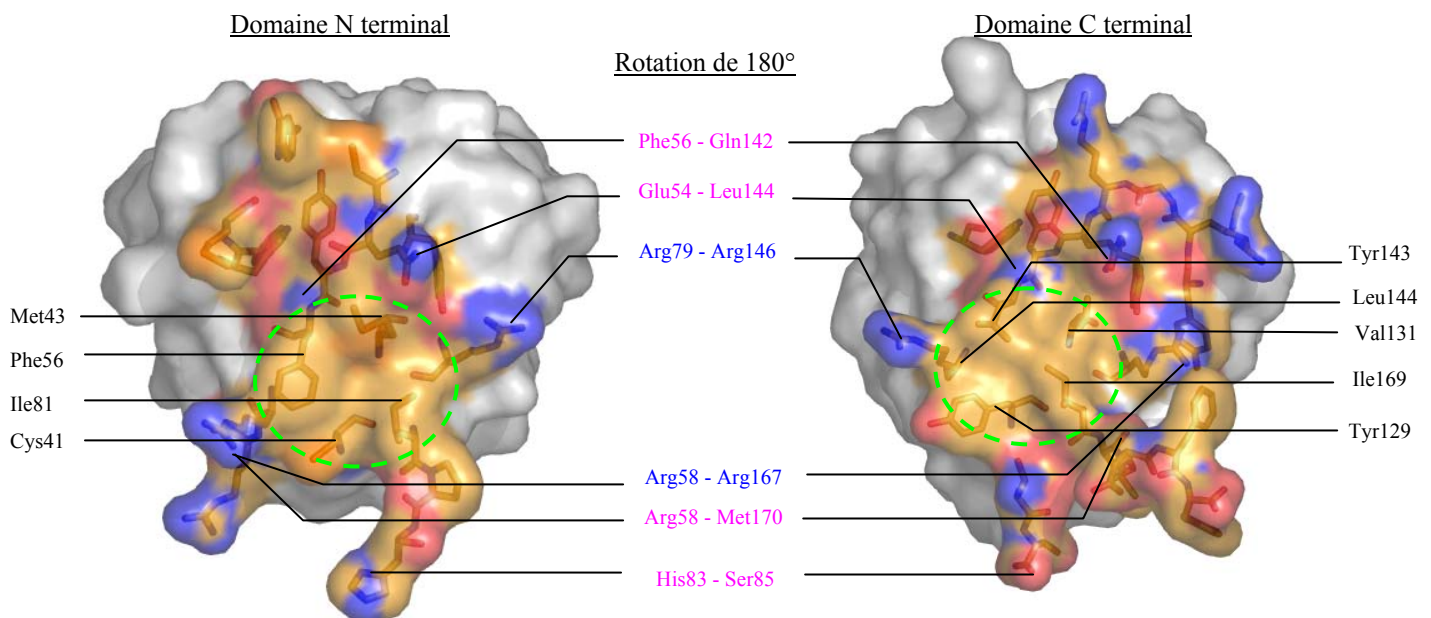


Figure 6.8 : Surfaces d'interaction entre les deux domaines N et C terminaux. Les acides aminés hydrophobes (texte noir) forme une zone hydrophobe (délimitée par des pointillés verts) qui est entourée par des résidus faisant interactions électrostatiques (texte en bleu) ou des liaisons hydrogènes (texte rose). Les surfaces sont colorées suivant les éléments qui se trouvent en surface (C: jaune; O : rouge; N : bleu et S : orange)

6.3.3 Contacts cristallins

L'empilement cristallin de la protéine étudiée est quasi identique à celui de la protéine native du rat : la superposition en couche des molécules va se faire de façon antiparallèle (voir figure 6.9). La protéine native cristallisait de façon à avoir deux molécules par unité asymétrique avec un axe 2_1 non cristallographique le long de l'axe z. Dans la $D\gamma E_{H_2O}$, le long de l'axe z est positionné un axe cristallographique ne permettant d'obtenir qu'une molécule dans l'unité asymétrique

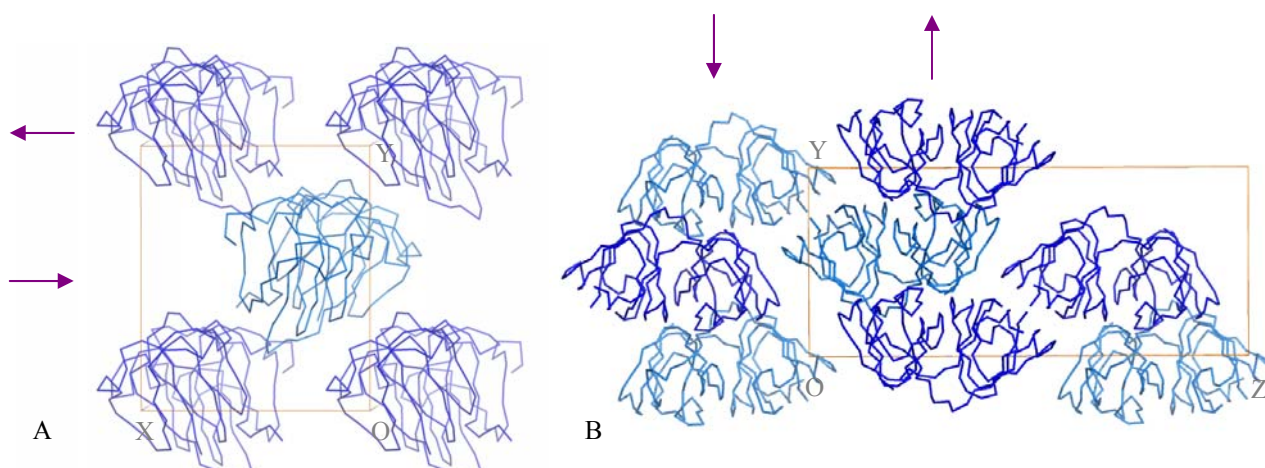


Figure 6.9 : Empilement cristallin de la protéine recombinante deutériée en H_2O . En A, la projection (001) de l'unité asymétrique représente l'empilement antiparallèle tout comme, dans la projection (100) de l'unité asymétrique (en B). Les flèches indiquent le sens antiparallèle des différentes couches de molécules.

Dans la protéine recombinante γE cristalline de rat, il y a 42 résidus qui interagissent de façon spécifique à moins de 4\AA avec un équivalent protéique mettant en jeu un grand nombre d'atomes. La majorité des résidus formant des contacts est regroupée en de petites surfaces de contact. Ces résidus vont participer à l'échafaudage de la maille cristalline en mettant en place des interactions spécifiques avec des résidus d'une autre molécule telles que des liaisons hydrogènes, des interactions électrostatiques (voir exemples en figure 6.11) ou des contacts entre atomes de types hydrophobes. D'autres interactions non spécifiques de type van der Waals vont aussi maintenir l'ensemble cristallin. Chaque interaction permet de caractériser un cristal d'un autre, de part la spécificité de ces interactions. Sur une molécule, nous avons identifié huit surfaces différentes qui interagissent chacune d'elles avec un équivalent cristallin différent (voir figure 6.10).

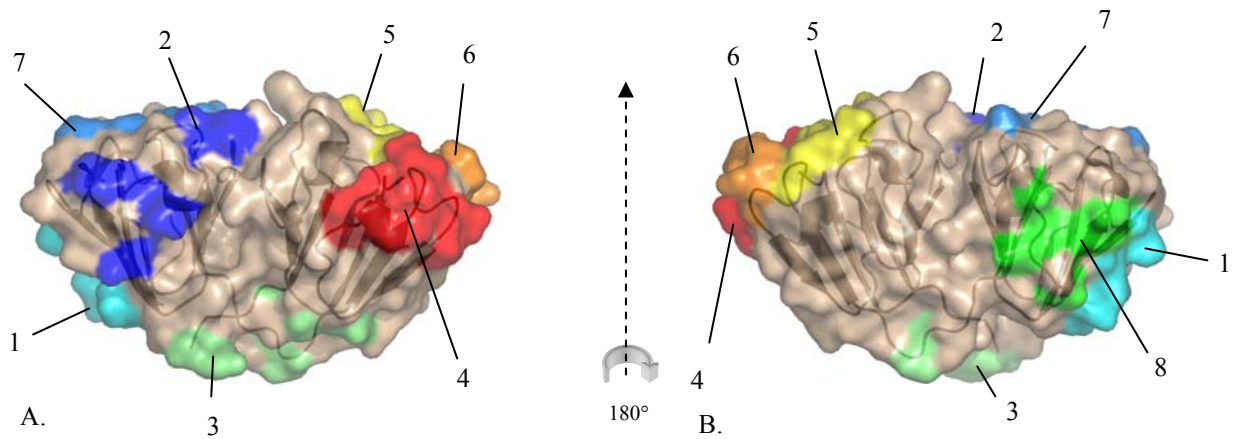


Figure 6.10 : Représentation des zones de contact cristallin (en couleur) sur la surface vue de face (A) et de l'arrière (B) de la protéine D γ E h₂O. Dans le cristal, ces zones vont s'apparier avec différents équivalents cristallins par des interactions spécifiques ou non et vont ainsi générer l'empilement cristallin.

Ces surfaces vont interagir avec un équivalent cristallin, par paire, de la façon suivante : 1 & 6, 2 & 3, 4 & 5 et 7 & 8. Les résidus impliqués de façon spécifique ont été classifiés dans le tableau 6.4 où ils sont classés en fonction de leur interaction avec un équivalent.

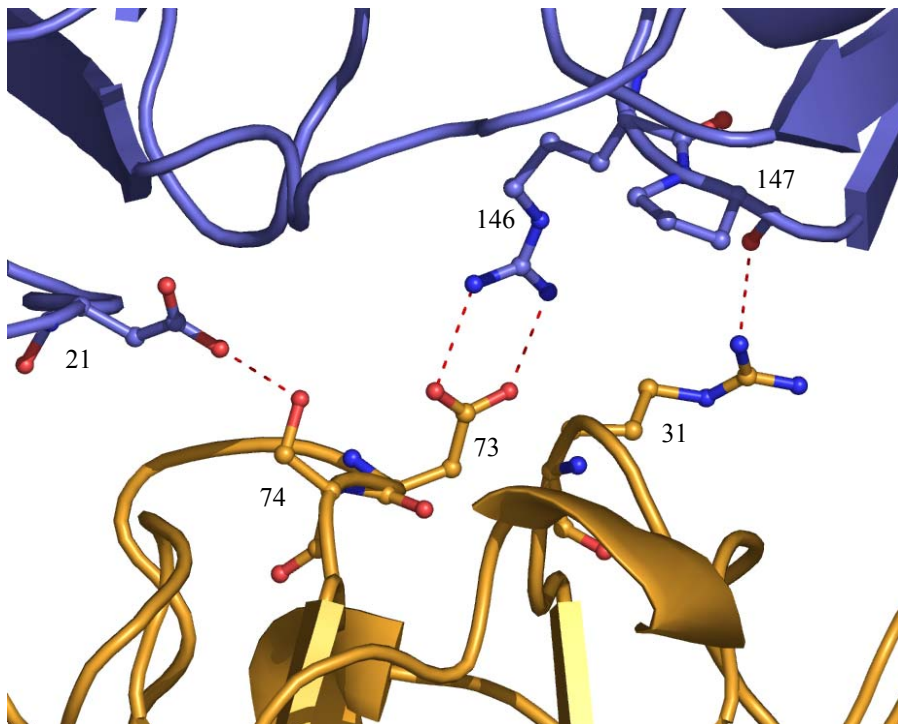


Figure 6.11 : Exemple de contact cristallin entre deux équivalents cristallins. Ces contacts sont réalisés entre les résidus 31, 73 et 74 de la zone 7 (en jaune) et les résidus 47, 146 et 21 de la zone 8 (en mauve) d'un équivalent symétrique.

Ces huit surfaces se décomposent en 4 zones de contact localisées du côté N terminal (1, 2, 7, 8) et 4 autres en C terminal (3, 4, 5, 6). Il n'y pas de surface purement hydrophobe

mais beaucoup de ces contacts mettent en jeu ce type d'interaction. Par exemple, la zone de contact 2 contient des résidus Gly70 et Phe71 qui interfèrent de façon hydrophobe avec les acides aminés Tyr153 et Phe50 de la zone 3. Il existe trois principaux domaines d'interactions : le domaine regroupant les zones 4, 5 et 6, le domaine comprenant les zones 1, 2, 7 et 8, et la zone 3.

<i>Molécule donnée</i>		<i>Molécule adjacente</i>		<i>Type de liaison spécifique</i>
<i>Zone de contact</i>	<i>Résidus</i>	<i>Résidus</i>	<i>Zone de contact</i>	
1 (x,y,z)	Lys2	Arg94	6 (-x+1/2,-y,z+1/2)	Liaison Hydrogène (L.H.)
	His15	Gly99		L. H.
	Glu17	Arg94		Ionique
	Asp38	Arg98		Ionique
2 (x,y,z)	Arg9	Thr105	3 (-x,y+1/2,-z+1/2)	L. H.
	Asp64	Arg168		Ionique
	Gln67	Ser84 Ser85		L. H.
	Gly70	Tyr173		L. H.
3 (x,y,z)	Phe71	Ser23	2 (-x,y+1/2,-z+1/2)	L. H.
	Ser23	Phe71		L. H.
	Ser84	Gln67		L. H.
	Ser85	Gln67		L. H.
	Thr105	Arg9		L. H.
	Arg168	Asp64		Ionique
4 (x,y,z)	Tyr173	Gly70	5 (x+1/2,-y+1/2,-z)	L.H.
	Gln100	Tyr153		L. H.
	Gln112	Glu95 Asp96		L. H.
	Asp113	Arg98		Ionique
	His116	His121 Glu95 Tyr153		L. H.
	Ser118	Glu95 Arg152		L. H.
5 (x,y,z)	Asp119	Arg152	4 (x+1/2,y+1/2,-z)	Ionique
	Glu95	Gln112 His116 Ser118		L. H.
	Asp96	Gln112		L. H.
	Arg98	Asp113 (Asp38)		Ionique
	His121	His116		L. H.
	Arg152	Ser118 Asp119		Ionique
6 (x,y,z)	Tyr153	Gln100 His116	1 (-x+1/2,-y,z+1/2)	L. H.
	Arg94	Lys2 Glu17		Ionique
	Arg98	Asp38 (Asp113)		Ionique
	Gly99	His15		L. H.
7 (x,y,z)	Arg31	Pro147	8 (-x,y+1/2,-z+1/2)	L. H.
	Glu46	Asp21		L. H.
	Glu47	Ser19		L. H.
	Ser72	Asp21		L. H.
	Asp73	Arg146		Ionique
	Asp74	Asp21		L. H.
8 (x,y,z)	Ser19	Gln47	7 (-x,y+1/2,-z+1/2)	L. H.
	Asp21	Glu46 Ser72 Ser74		L. H.
	His22	Arg141		L. H.
	Ser23	Phe71		L. H.
	Arg146	Asp73		Ionique

Tableau 6.4 : Appariements des résidus de contacts dans l'empilement cristallin de la γ E cristalline deutériée dans l'eau légère. La désignation des symétriques en contact avec la molécule de référence se fait par les opérations de symétrie décrites entre parenthèses. Les acides aminés indiqués entre parenthèses interagissent avec le même résidu mais proviennent d'un équivalent cristallin différent.

Les deux premiers domaines, situés aux extrémités de la protéine, mettent en place linéairement les molécules de façon "tête bêche" et forment une sorte de spirale positionnée le long de l'axe z (voir figure 6.12), ce qui peut rendre plus favorable alors un empilement bidimensionnel et donc peut expliquer l'obtention de cristaux en forme de plaque.

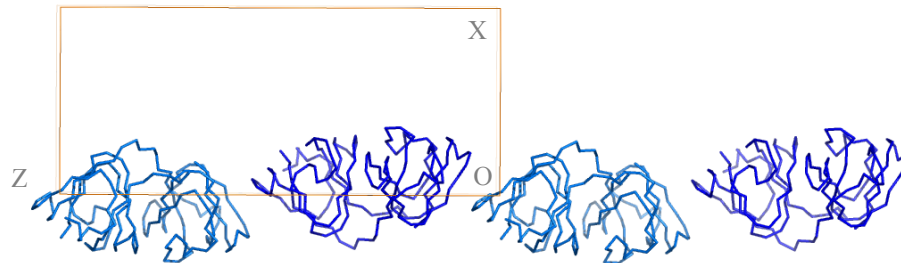


Figure 6.12: Empilement cristallin de façon "tête bêche" de la $D\gamma E h_2O$ en forme de spirale le long de l'axe z .

De plus la zone 3 qui devrait permettre la troisième dimension cristalline, contient les résidus N terminaux dont l'Asp171, le Phe172 et la Tyr173. Ces acides aminés ont été précédemment décrits comme ayant des facteurs thermiques élevés et sont donc, certainement des résidus considérés comme flexibles. Toutes les interactions survenant de ces résidus ne seront pas forcément présentes dans toutes les molécules du cristal. L'empilement cristallin dans cette direction risque d'être mal effectué. Ceci peut être une des explications à la formation de cristaux à deux dimensions.

6.3.4 Boucle entre les motifs c et d dans le motif en clef grecque 3

Il est intéressant de noter que l'empilement cristallin est maintenu aussi, par une boucle caractérisée comme flexible et variable (Norledge et al. 1997). En effet, la zone 4 montre de nombreuses interactions avec la surface 5 d'un autre symétrique. Cette boucle, située entre les brins c et d du motif en clé grecque 3, comporte les acides aminés Phe115, His116, Phe117, Ser118 et Asp119, et crée de nombreuses liaisons hydrogène et ioniques avec la zone 5. Dans cette étude, cette partie de la protéine a été observée sous deux conformations différentes. En effet, les chaînes principales et latérales de cette partie, divergent de quelques Å, sur certains résidus mais essentiellement pour l'histidine 116 : 4,47Å sur les $C\alpha$ et 11,10Å sur les $C\epsilon$ (voir figure 6.13A).

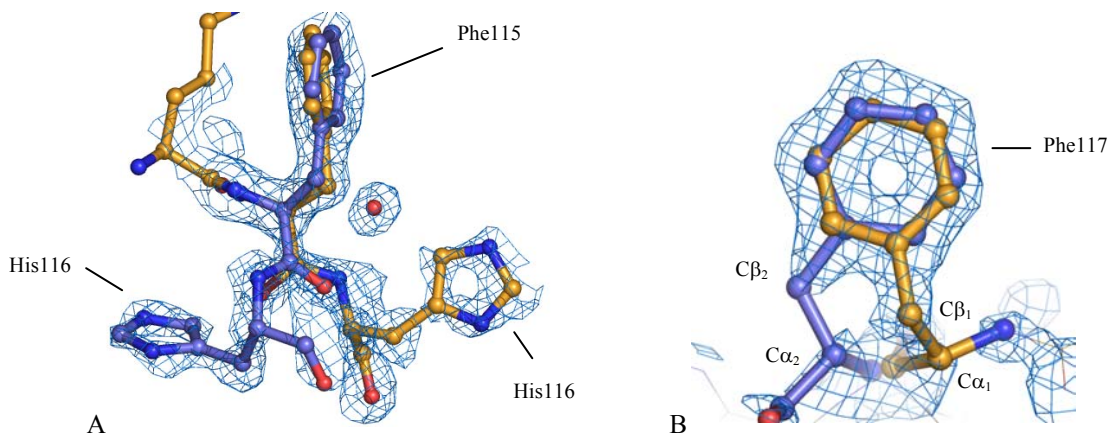


Figure 6.13 : (A) Représentation de la double conformation des résidus Phe115 et His116 entourés par la densité électronique $2|F_o|-|F_c|$ à $0,8\sigma$. La première configuration de cette chaîne a les atomes de carbone représentés en jaune et, dans la seconde conformation, ces derniers sont en bleu. (B) Représentation du résidu Phe117 avec sa carte de densité électronique $2|F_o|-|F_c|$ à 1σ . Les C α et C β ne sont pas à la même position mais le cycle aromatique est à la même place.

L'histidine 116 dans la première conformation fait essentiellement des interactions de type intramoléculaire avec le résidu glutamine 100 et quelques molécules d'eau. Alors que la seconde conformation interagit en grande partie avec des résidus d'un symétrique cristallin tel que l'histidine 121. Ceci entraîne un décalage de la chaîne principale au niveau des acides aminés suivants. La phénylalanine 117 a les deux cycles aromatiques bien positionnés dans l'espace malgré une divergence au niveau de la numérotation expliquée par le positionnement différent des carbones α et β , dissemblables de 2,12 Å et 2,72 Å respectivement (voir figure 6.13B). Mais le décalage de la chaîne latérale est rattrapé sur la double conformation des résidus Ser118 et Asp119 qui vont permettre d'avoir une Phe120 positionnée de la même façon dans les deux conformères (voir figure 6.14).

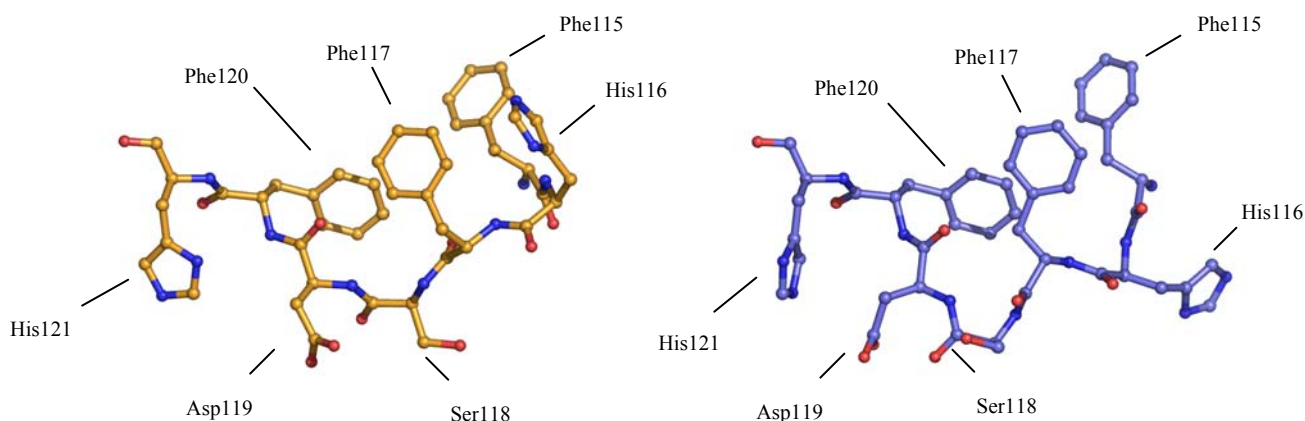


Figure 6.14 : Représentation des deux conformations de la boucle entre les brins *c* et *d* du motif 3. Les résidus Phe115 ont presque la même position mais par la suite les His116 sont très distants ce qui va entraîner un décalage au niveau de la chaîne principale jusqu'au résidu Phe120.

Dans la seconde conformation, le résidu His121 va être sous une conformation différente du fait que le résidu aspartate 119 est plus proche de lui. Le plan du cycle aromatique de l'histidine 121 ne sera plus alors perpendiculaire au plan de la chaîne latérale de l'aspartate 119 ; ils seront "parallèles" pour limiter l'encombrement stérique généré par cette nouvelle conformation. Grâce ou à cause de cette seconde conformation, un grand nombre de liaisons hydrogène vont être établies entre cette histidine 121 et l'histidine 117 d'un autre symétrique protéique et l'empilement cristallin va être favorisé.

Il a été observé la formation d'une double conformation de la phénylalanine 117 suite à une mutation au niveau de l'arginine 58 de la γ D humaine, à l'origine de cataractes aculéiformes (Basak et al. 2003). Il est intéressant de constater que, malgré une structure hautement conservée, les γ cristallines ne possèdent pas, toutes, une flexibilité au niveau de cette boucle. Cependant, cette boucle a été caractérisée comme l'une des principales différences structurales entre les γ cristallines de basse et de haute température de transition (Norledge et al. 1997). Dans la protéine γ F de bœuf, White et al. (1989) ont voulu montrer l'importance des résidus 117, 120 et 124 formant un cœur hydrophobe, dans la conformation de la boucle entre les brins *c* et *d* du motif 3 et donc dans leur participation à la transition de phase. Ces trois résidus semblent d'encombrement plus important chez les protéines de haute température de transition que chez les cristallines de basse température (voir tableau 6.5).

Groupe	Nom et Organisme	Séquence													
Haute Température	γ E de bœuf	115	-	-	F	-	-	I	-	-	-	F	-	-	126
	γ F de bœuf	115	-	-	F	-	-	I	-	-	-	F	-	-	126
	γ E de rat	115	-	-	F	-	-	F	-	-	-	F	-	-	126
	γ F de rat	115	-	-	F	-	-	F	-	-	-	F	-	-	126
Basse Température	γ B de bœuf	115	-	-	L	-	-	V	-	-	-	L	-	-	126
	γ D de bœuf	115	-	-	F	-	-	I	-	-	-	L	-	-	126
	γ C d'homme	115	-	-	L	-	-	I	-	-	-	L	-	-	126
	γ A de rat	115	-	-	L	-	-	I	-	-	-	M	-	-	126
	γ B de rat	115	-	-	L	-	-	I	-	-	-	L	-	-	126
	γ C de rat	115	-	-	L	-	-	V	-	-	-	L	-	-	126

Tableau 6.5 : Alignement des acides aminés 117, 120 et 124 des γ cristallines caractérisées pour les hautes et les basses températures de transition. Les protéines à haute température de transition possèdent des acides aminés de fort encombrement, alors que ce n'est pas le cas pour les protéines dont la température de transition de phase est faible.

Cet encombrement différent permettrait d'avoir différentes conformations de la boucle suivant les protéines. Norledge et al. (1997) et White et al. (1989) ont conclu que la boucle *cd* pouvait avoir différentes conformations entre les cryoprotéines (haute température de transition) et les cristallines de basse température de transition.

Dans la structure résolue de la γ E de rat, les résidus Phe117, Phe120 et Phe124 ont gardé leur emplacement et donc forment toujours ce cœur hydrophobe, malgré cette double conformation de la chaîne principale. L'acide aminé Phe115 qui fait partie aussi de cette zone d'hydrophobicité et qui est au départ de la double conformation, a un positionnement de son cycle aromatique bien conservé. Ainsi, la présence de cette double conformation ne modifie pas la présence et la taille de la zone hydrophobe située entre les motifs en clé grecque 3 et 4. La γ E cristalline de rat a été caractérisée expérimentalement comme une cristalline du groupe II, groupe ayant une température de transition élevée (Siezen et al. 1988). Quand les boucles *cd* des protéines de faible et de haute température de transition, sont comparées, les chaînes principales des γ F et γ B cristallines de bœuf ressemblent grandement aux deux conformations de la γ E cristalline (voir figure 6.15).

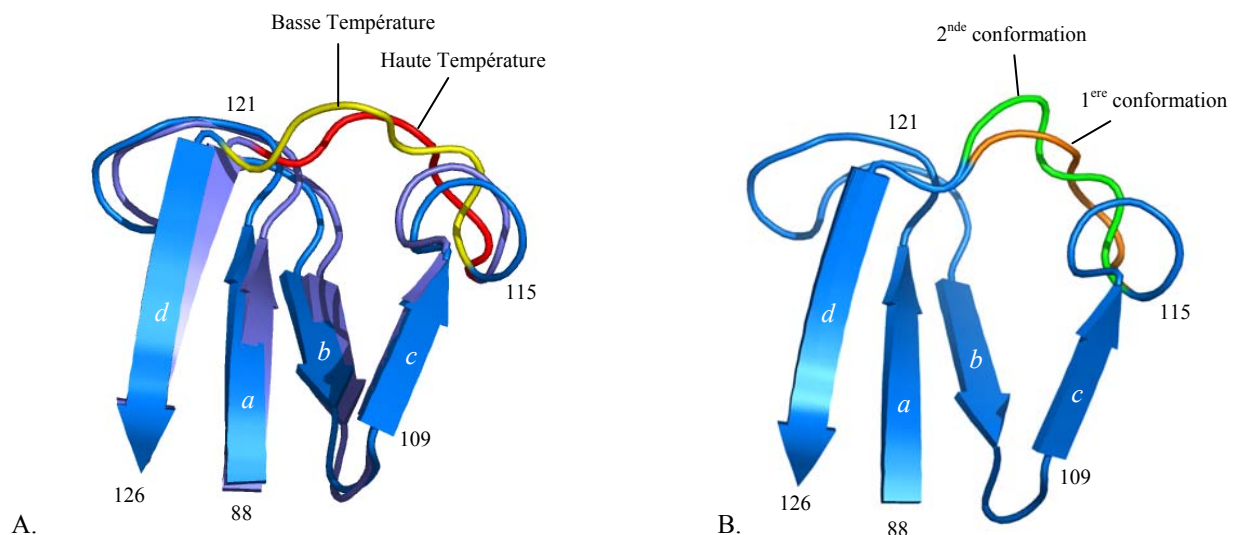


Figure 6.15 : Représentation des boucles *cd* du motif 3 des γ cristallines. (A) Les chaînes principales des γ IVa (ou γ F violet, rouge et haute température de transition) et γ II (ou γ B, bleu, jaune et basse température de transition) cristallines de bœuf sont alignées et les boucles de couleur montrent la différence majeure qu'il existe entre les protéines de basse et de haute température de transition. (B) La double conformation de la chaîne principale (vert et orange) de la γ E cristalline de rat semble être relativement similaire aux deux autres conformations.

La présence de cette double conformation peut amener à se poser les questions suivantes : cette double conformation n'est-elle pas la « réplique » des conformations exprimant les propriétés de transition des deux groupes de protéines ? Si c'est le cas, la conformation de cette boucle dans les γ cristallines est-elle ou n'est-elle pas alors la raison principale qui conduit aux caractéristiques thermodynamiques distinctes des deux groupes de γ cristallines comme l'avaient pensé Norledge et White ? N'est-ce pas plutôt la taille et l'encombrement de la zone hydrophobe qui est en cause ?

Pour tenter de savoir si la double conformation correspond aux conformations des cristallines de basse et de haute température de transition, une comparaison de la déviation des deux chaînes principales de la cristalline γ E avec les chaînes principales de la γ B de bœuf (basse température, PDB id. 1AMM) et de la γ F de bœuf (haute température, PDB id. 1A45) a été effectuée, grâce au programme *Lsqkab*, en figure 6.16 :

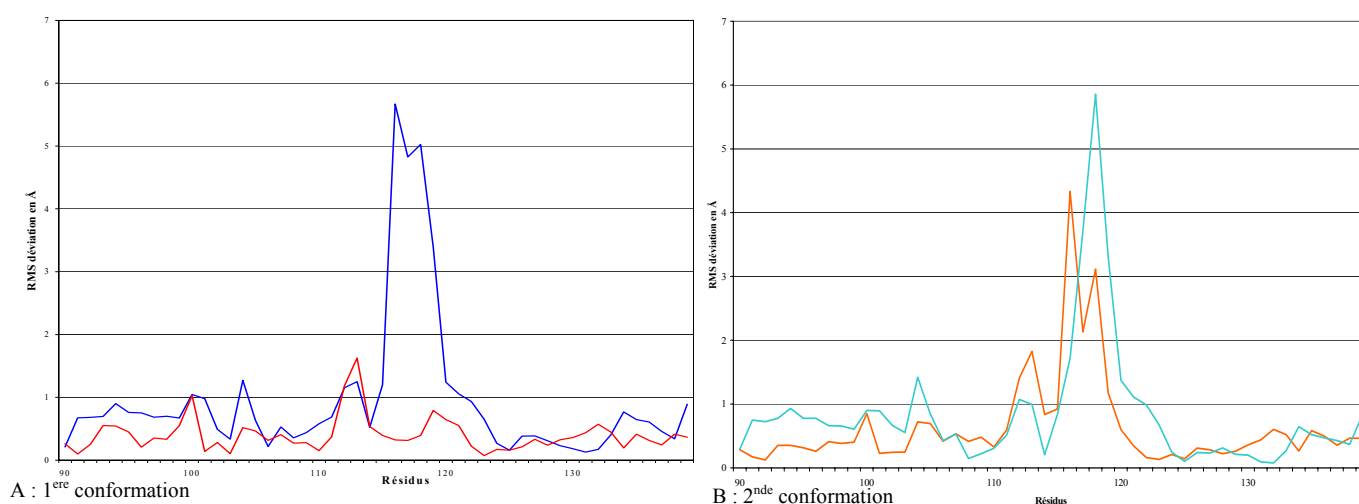


Figure 6.16 : Déviation des C α des résidus 90 à 140. (A) Comparaison de la première conformation de la γ E cristalline avec la γ B de bœuf (Bleu) et de la γ F de bœuf (Rouge). Cette conformation s'apparente plus aux cristallines de haute température de transition. (B) Comparaison de la seconde conformation de la γ E cristalline avec la γ B de bœuf (Bleu) et de la γ F de bœuf (Orange). La seconde conformation ne s'apparente pas vraiment avec un des deux groupes des cristallines.

Dans la comparaison de la première conformation de la γ E cristalline avec des protéines des deux groupes, il apparaît que cette forme de la boucle est apparentée au groupe des cristallines à haute température de transition, puisque la déviation des C α est faible avec ce groupe de cristallines (- de 1Å). Dans la comparaison de la seconde conformation de la boucle, l'assimilation d'un groupe de cristallines est moins évident même si, encore une fois, cette forme semble être apparentée au même groupe de cristallines. En effet, la déviation maximum avec la γ F excède, à peine les 4,3Å, alors qu'avec la déviation maximum atteint, presque les 6Å.

Il est à retenir que cette analyse a été effectuée sur deux protéines de chaque groupe dont les structures cristallographiques sont connues : la γ B et la γ D (Chirgadze et al. 1996) de bœuf pour les protéines de faible température de transition ; et la γ F et γ E de bœuf (Mayer et al. 2003). Les résultats sont étroitement similaires, et permettent d'émettre les hypothèses que cette boucle *cd* est différente suivant les deux groupes de protéines, mais que celle ci devrait être relativement conservée à travers ces deux familles aux propriétés

différentes. D'après cette analyse, les deux conformations de la γ E cristalline ne semblent pas être identifiables aux deux familles de cristallines, seule la première conformation est très proche du groupe des protéines à haute température de transition, ce qui est en accord avec les sous groupes décrits par Siezen et *al.* (1988). La seconde conformation ne paraît pas être assimilée à la conformation du groupe des protéines à basse transition de phase. Donc il est possible d'avancer que les deux conformations de la cristalline γ E de rat ne représentent pas les conformations de la boucle *cd* de chacun des groupes dont la température de transition est différente. Dans ce cas, la conformation de cette boucle reste donc la principale différence structurale existante entre ces deux groupes, malgré la double conformation dans la γ E cristalline. Si ce n'est pas la conformation de la boucle qui influence le changement de température, il serait intéressant de vérifier l'importance de certains acides aminés sur la température de transition. Par exemple, suite à une mutagenèse dirigée, l'influence de résidus tels que les phénylalanines 117,120 et 124 peut être évaluée par diffusion lumière comme l'ont fait Broide et al (1991), Thomson et al. (1987) et Siezen et al (1991) sur les protéines natives. Si des différences de température de transition, dues à des mutations, sont observées, cela montrerait l'importance de ces résidus dans ce comportement de transition de phase. Il serait alors possible d'émettre l'hypothèse que des acides aminés hydrophobes encombrants, tels que les phénylalanines, favorisent la stabilisation de cette zone, menant à une température de transition plus haute. Cette étude n'a pas été faite au cours de cette thèse par manque de temps et car ce n'était pas l'objet final de cette étude.

6.3.5 Le tripeptide RGD et la partie C terminale

Au niveau de la position 59, se trouvent trois acides aminés, arginine glycine aspartate (RGD), qui sont souvent retrouvés lors d'interactions moléculaires telles que l'adhésion cellulaire, la phagocytose des protéines, l'agrégation de plaquettes sanguines ou la liaison avec le complément. Beaucoup de protéines adhésives présentes dans la matrice extracellulaire et dans le sang contiennent ce tripeptide (Ruoslahti et Pierschbacher 1986). De nombreux membres de la famille des intégrines (Takagi 2004), reconnaissent le motif RGD comme leur ligands. L'intégrine α V β 3 dont la structure a été résolue avec un peptide mimant le RGD, montre clairement l'interaction faite par ce motif (Xiong et al. 2002). Or cette séquence est présente dans un très grand nombre de γ cristallines à l'inverse des β cristallines (Najmudin et al. 1993). Pierschbacher et Ruoslahti (1987) ont voulu démontrer, en comparant le RGD des structures de la γ B cristalline, de l' α protéase lytique et de la thermolysine, que

les séquences adjacentes ont une importance toute relative et que la conformation de ce tripeptide varie suivant ces dernières. Ces trois protéines ne présentent pas la même conformation au niveau de ces trois peptides mais cette dernière n'en diffère pas énormément. Ce qui est relativement important comme l'affirme, Reed et al. (1988), c'est que la présence de chaînes latérales chargées n'est pas suffisante. Il faut aussi que ces résidus puissent initier des brins β nichés et stabilisés par des liaisons hydrogène inter résidus. Ceci est le cas pour les γ cristallines, où le RGD est en début du brin β *d* du motif 1. Les résidus environnants favorisent, aussi, un type de conformation du RGD et d'interaction avec le ligand (Hynes 1987; Takagi 2004). Par exemple, dans les interactions physiologiques intégrines α V β 3 ligand, beaucoup de résidus des sous unités α et β sont impliqués dans la reconnaissance du ligand et ces acides aminés, en dehors du motif RGD, ont une contribution essentielle dans cette association (Takagi 2004). En revanche, il n'a jamais été démontré un environnement spécifique favorable aux interactions moléculaires.

Les acides aminés proches du motif RGD dans le modèle discuté pendant cette thèse sont en grande majorité des résidus de surface possédant deux ou plusieurs conformations avec des chaînes latérales de taille importante. Ces résidus ont apparemment la possibilité de faire de nombreuses interactions électrostatiques puisqu'ils regroupent une arginine Arg36, un aspartate Asp38, une sérine Ser39, une phénylalanine Phe172 et une tyrosine Tyr173 (voir figure 6.17).

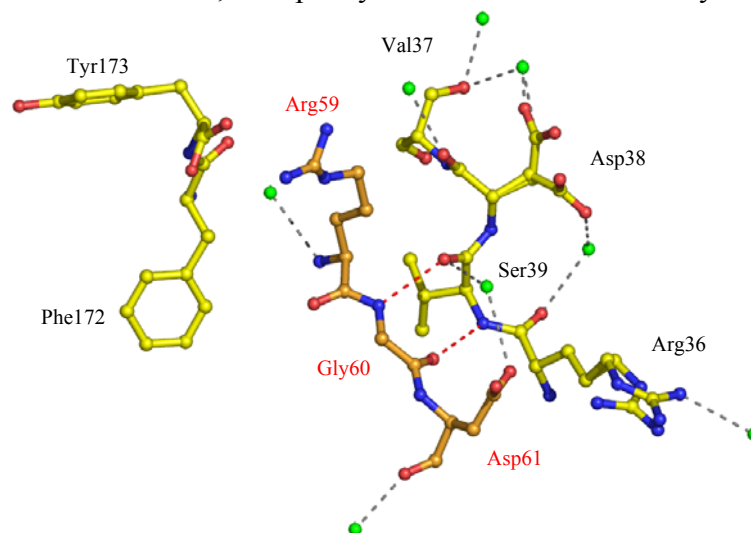


Figure 6.17 : Représentation du tripeptide RGD et des acides aminés proches dans la structure de la γ E cristalline deutériée en solvant hydrogéné. Contrairement au RGD, les proches résidus sont soit une double conformation (Arg36, Asp38) soit font partie d'une zone flexible (Phe172, Tyr173). En vert, sont représentées les molécules d'eau voisines et en pointillé les liaisons hydrogène dont celles, en rouge, entre le RGD et la sérine 39.

Ce motif RGD se situe au niveau du brin β *c* du motif 2 en clé grecque de la γ cristalline. La structure est principalement stabilisée par des liaisons hydrogène (en pointillé

rouge sur la figure) avec le brin βd du motif 1, spécialement entre les résidus valine 37 et glycine 60. D'autres interactions impliquant des molécules d'eau stabilisent le système. Ce proche environnement pourrait permettre de nombreuses interactions spécifiques avec d'autres protéines. Cependant, aucune interaction de ce type avec une γ cristalline a été caractérisée dans cette zone, hormis les contacts cristallins de la γ B cristalline de bœuf (Najmudin et al. 1993) et la γ D cristalline humaine (Basak et al. 2003).

Dans le premier cas, l'arginine 59, dite « critique », a des facteurs de températures B hauts et interagit avec le carbonyle de l'histidine 117 d'une molécule symétrique. Najmudin et al. (1993) ont comparé la structure de la γ B avec le RGD de la protéine VP1 du virus de la fièvre aphteuse (Acharya et al. 1989). Ils ont conclu que la conformation identique des tripeptides de ces deux protéines et la flexibilité de l'arginine 59, pouvaient expliquer la capacité de liaison de la γ B cristalline.

Dans le cas de la γ D humaine, Basak et al. (2003) ont mis en avant la mutation R58H. Comme il a été précisé auparavant, cette mutation chez la γ D cristalline humaine, qui remplace l'arginine 58 en une histidine, est à l'origine des cataractes aculéiformes. Dans le cristal de la forme native, l'arginine 58 interagit avec l'aspartate 156 d'un symétrique moléculaire. La mutation R58H, a un impact sur la surface moléculaire générale puisque le contact cristallin est essentiellement remplacé par des interactions avec des molécules d'eau causées par une réorganisation locale de certaines chaînes latérales.

Ainsi, malgré le manque de preuve d'interaction *in vivo* des γ cristallines avec d'autres protéines *via* ce RGD, il est clair que la spécificité de liaison doit résider dans la conformation du tripeptide RGD et le rôle des séquences adjacentes est déterminant pour adopter cette conformation appropriée. En conclusion, la chaîne principale et les chaînes latérales de la séquence RGD sont bien définies par des interactions tertiaires excepté pour la fin de la chaîne latérale de l'arginine 59 qui est donc libre d'interagir avec un récepteur.

Comme précisé juste avant, la partie C terminale est proche du tripeptide RGD. Dans toutes les structures étudiées pendant la thèse, les résidus C terminaux de 171 à 173 montrent une densité des chaînes principales et latérales quasi inexistante ou trop complexe à déterminer jusqu'à la fin de l'affinement. Ce désordre met en évidence la large flexibilité de la

queue C terminale observée dans pratiquement toutes les structures de γ cristallines, mise à part pour la γ B de bœuf où les résidus phénylalanine 172 et tyrosine 173 ont été observés sur des cartes de densité électronique à $0,6\sigma$ (Najmudin et al. 1993). Dans notre étude, seule, la chaîne latérale du dernier acide aminé, la tyrosine 173, est positionnable du fait qu'elle permet un contact cristallin entre deux monomères protéiques (avec les résidus 66 et 70 d'un symétrique cristallin). Tous ces résidus C terminaux ont été placés avec une occupation nulle pour éviter de fournir toute information fautive. Le fait que cette partie C terminale, extrêmement flexible, se retrouve à proximité du tripeptide RGD peut éventuellement « cacher » ce dernier à des récepteurs. Ainsi, toute interaction du RGD avec un récepteur en serait défavorisée.

6.3.6 Autres détails sur les structures

Comme il a été précisé dans le chapitre 2, les γ cristallines n'ont pas de fonction enzymatique connue à ce jour. Seules quelques suppositions à propos de réponses à des stress oxydatifs, des rôles protecteurs vis à vis de radicaux, des capacités réductrices, ont été évoquées.

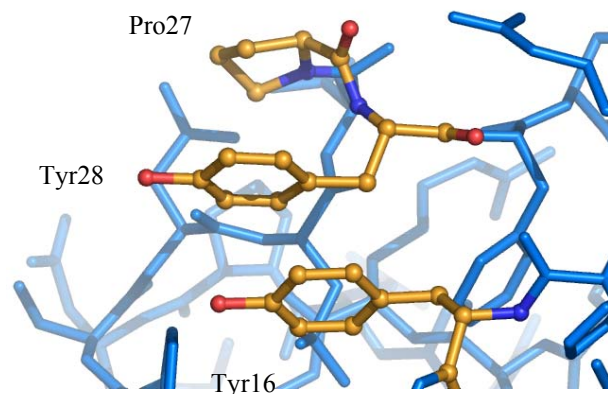
6.3.6.1 Analyse des résidus cystéines

Comme les cristallines ont été considérées comme très proches des oxido réductases (Wistow et Piatigorsky 1988; Rao et Zigler 1991), et que le nombre de cystéines est relativement important dans toutes les γ cristallines, il est remarquable d'observer que les structures, hydrogénées ou deutériées, en H_2O ou en D_2O , contiennent des cystéines uniquement sous forme réduite. Sur les six cystéines présentes dans la γ E de rat, quatre n'ont pas la possibilité d'effectuer un pont disulfure du fait qu'elles se trouvent trop éloignées les unes des autres. En revanche, le couple Cys18 et Cys78, dont les atomes de soufre sont orientés favorablement, n'ont pas la possibilité de faire une liaison intramoléculaire du même genre que celle observée dans la γ B cristalline de bœuf entre les résidus 18 et 22 (McDermott et al. 1988) puisque les deux atomes de soufre sont distants de $3,87 \text{ \AA}$. Selon le programme *Prochec*, la distance idéale de séparation est de 2 \AA . De plus, les facteurs thermiques et la densité électronique démontrent clairement que les acides aminés 18 et 78 ne font pas de la liaison covalente. Comme l'ont suggéré Najmudin et al. (1993) et bien d'autres, il est fort possible que les γ cristallines, de part leur grand nombre de cystéines réduites, peuvent jouer un rôle réducteur à l'intérieur même du cristallin.

6.3.6.2 Empilement des résidus aromatiques

En surface, au niveau du motif 1 du domaine N terminal, les résidus tyrosines 16 et 28 ont un empilement au niveau de leur cycle aromatique et avec le cycle de la proline 27 (voir figure 6.18). Un tel empilement peut contribuer (Vollhardt et Schore 1995) à la stabilité du motif et donc de la structure générale de la protéine en délocalisant les électrons des orbitales p , le long des liaisons π des différents cycles aromatiques. Les « courants de cycle » générés par les électrons en circulation de chacun des cycles, engendrent deux champs magnétiques locaux qui vont se superposer (Vollhardt et Schore 1995).

Figure 6.18 : Empilement des tyrosines 16 et 28 et de la proline 27. Le « courant de cycle » dû aux électrons en circulation des cycles aromatiques vont créer des champs électriques locaux qui vont s'empiler. Il est impossible dans cette conformation protéique de former des carbones C_e tétraédriques par modification radicalaire.



Balasubramanian et al. (1990) ont indiqué que les tyrosines présentes dans les cristallines, préférentiellement les γ cristallines et surtout les γ A et γ E, sont sujet à une forte oxydation. Ce genre d'oxydation, produite par une des espèces réactives de l'oxygène, telles que le superoxide (O_2^\bullet), l'eau oxygénée (H_2O_2), le radical hydroxyle ($^\bullet OH$), et le singulet d'oxygène (1O_2), peut engendrer la formation de dihydroxyphénylalanine ou de dityrosine, liaison entre les carbones C_e des tyrosines voisines. Cette modification oxydative fait partie des nombreux marqueurs du vieillissement entraînant la formation d'agrégat de haut poids moléculaire des protéines du cristallin, retrouvés dans les cataractes (Guptasarma et Balasubramanian 1992).

Guptasarma et Balasubramanian (1992) ont publié que la réaction avec les différentes espèces de l'oxygène citées précédemment, n'engendre pas la formation de dityrosine inter ou intramoléculaire mais qu'une photo réaction des états excités des chromophores tryptophane et tyrosine peut conduire à la formation de dityrosine. La formation de dityrosine peut aussi, impliquer un transfert d'électron par réaction radicalaire, mais cette réticulation covalente ne s'effectue qu'en présence d'un agent transfert d'électron

(le Fer par exemple) comme pour la myoglobine de cachalot, où la formation de dityrosine a été démontrée (Tew et Ortiz de Montellano 1988).

En examinant la structure de la γ E cristalline résolue pendant cette thèse, il s'avère que seuls les résidus 16 et 38, et éventuellement les résidus 133 et 138 (mais les chaînes latérales ne sont pas orientées dans la même direction), ont des cycles aromatiques proches, bien que la distance entre deux C_α équivalents soit supérieure à 4 Å. La densité électronique et les facteurs thermiques B montrent bien l'absence de liaison de ce type. Une photolyse directe ou un traitement à l'eau oxygénée en présence de Fer pourrait éventuellement créer une réticulation covalente entre ces résidus.

6.4 Analyse de la structure du solvant

La haute résolution a permis de localiser un nombre substantiel de molécules de solvant grâce au programme *ARP-wARP*. L'occupation a été déterminée suivant les résidus proches capables de faire des liaisons hydrogène avec les molécules d'eau par le programme *ShelX*.

Une molécule d'eau est considérée comme présente dans le modèle quand elle valide certains critères : la molécule d'eau doit faire au minimum une liaison hydrogène de 2,4 à 3,5 Å avec un atome d'oxygène ou d'azote de la protéine ou avec une autre molécule de solvant (H_2O , D_2O ou acétate) ; la densité électronique doit être bien définie au dessus de 1σ dans la carte $2|F_o|-|F_c|$ et ne présente pas de désaccord à 3σ dans la carte $|F_o|-|F_c|$. De plus, toutes les molécules d'eau dont le facteur B dépassait 60 \AA^2 et pour qui la densité électronique n'était pas bien définie, ont été supprimées au cours des différents affinements.

Plus d'une centaine de sites pour l'eau légère (ou l'eau lourde) ont été identifiés dans chacune des structures. Le nombre de molécules d'eau est à peu près équivalent à celui trouvé pour la structure de la γ E cristalline résolue par Norledge et al. (1997), comme le montre le tableau 6.6. Certaines molécules d'eau sont impliquées dans des occupations alternées de chaînes latérales comme au niveau de la boucle *cd* du motif 3. Ces molécules ont la même occupation que l'atome de la chaîne latérale auquel elles sont rattachées par une liaison hydrogène. Mais la majorité du solvant est bien ordonnée avec des facteurs B représentés dans le tableau 6.6. En plus des molécules d'eau, des molécules d'acétate ont été

positionnées dans toutes les structures. Aucune molécule de magnésium n'a été trouvée dans les différentes structures malgré sa présence dans le solvant de cristallisation.

<i>Protéine</i>	<i>HγEh2o</i>	<i>HγEd2o</i>	<i>DγEh2o</i>	<i>DγEd2o</i>
<i>Résolution</i>	1,45	1,44	1,36	1,72
<i>Nombre de molécules d'H₂O (D₂O) par protéine</i>	121	161	145	120
<i>Facteur B moyen en Å² des molécules d'eau</i>	29,7	28,2	26,3	30,7
<i>Nombre de ligands par protéine</i>	2 Acétates	2 Acétates	3 Acétates	2 Acétates
<i>Facteur B moyen en Å² des acétates</i>	43,0	55,4	31,3	39,3

Tableau 6.6 : Nombre de molécules de solvant identifiées pour la γ E cristalline de rat hydrogénée ou deutériée en solvant hydrogéné ou deutérié et pour le modèle 1A5D. Les facteurs B sont calculés après affinement par *ShelX*.

Il existe certaines différences quantitatives entre les structures résolues qui sont la conséquence de nombreux paramètres. Premièrement les cristaux ne sont pas identiques. Il est tout à fait possible de ne pas observer des atomes (molécules d'eau et double conformation de chaîne latérale) sur des cristaux provenant de la même goutte de cristallisation. Ensuite, les données cristallographiques ne sont pas les mêmes suivant les structures. Ces structures n'ont pas été déterminées à la même résolution et le nombre de données par rapport aux paramètres ne sont pas identiques. Par exemple, dans la structure de la protéine deutériée en solvant hydrogéné, la présence de trois acétates n'est pas à rediscuter. En revanche, dans les autres structures, à l'emplacement de l'acétate manquant, deux atomes d'oxygène, représentant des molécules d'eau, sont positionnés au niveau des deux atomes d'oxygène de l'acétate. Au niveau du groupement méthyle, l'atome de carbone est en revanche très mal définie. Donc deux molécules d'eau vont remplacer cet acétate dans les deux structures de la protéine hydrogénée et celle de la protéine deutériée en milieu deutérié.

6.4.1 Comparaison du solvant entre les différentes structures

Une étude comparative des facteurs thermiques *B* sur le solvant entourant la γ E cristalline a été entreprise sur les 98 molécules d'eau identiques et présentes dans toutes les structures (voir figure 6.19A). Les molécules n'étant pas présentes dans toutes les structures,

ne sont pas prises en compte. Une comparaison topologique a été effectuée sur ces mêmes molécules (voir figure 6.19B).

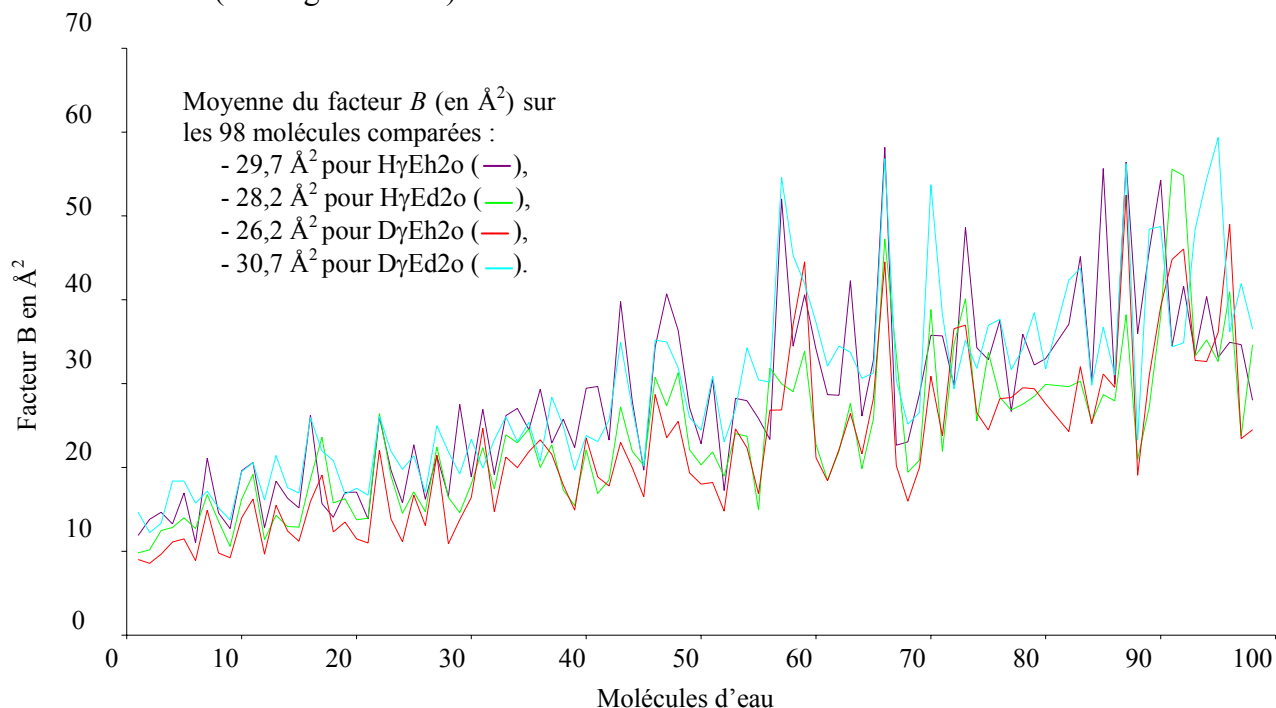


Figure 6.19a : Comparaison des facteurs B des 98 molécules d'eau après affinement par *ShelX*.

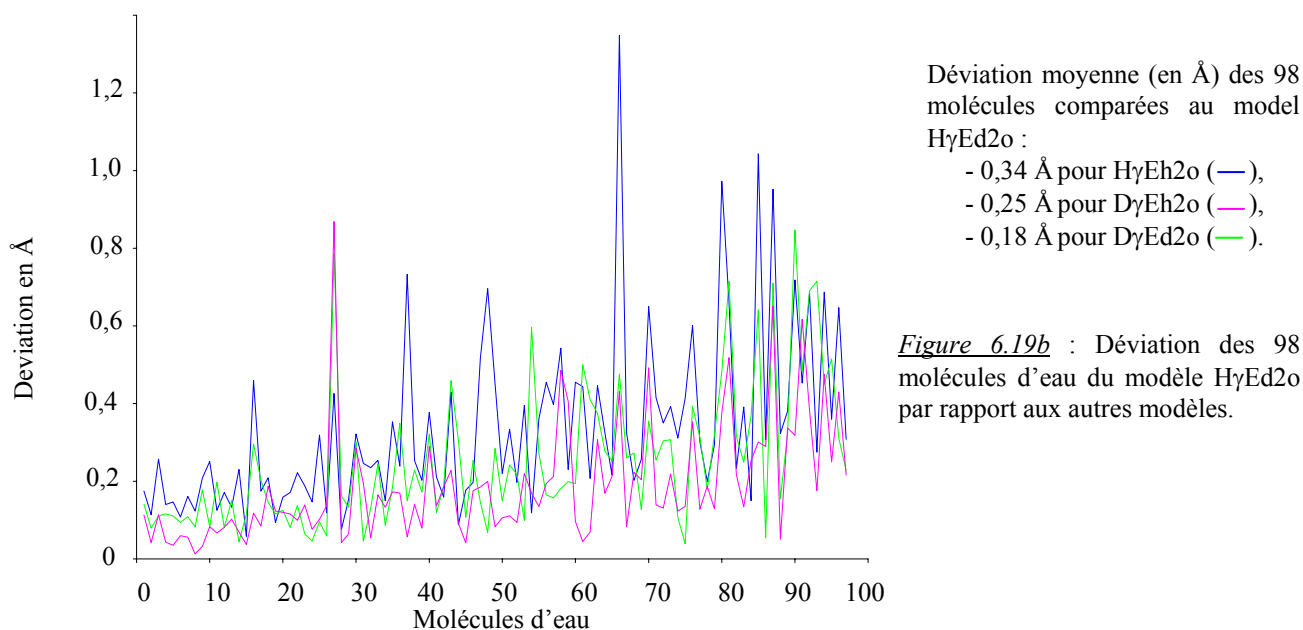


Figure 6.19b : Déviation des 98 molécules d'eau du modèle H γ Ed2o par rapport aux autres modèles.

Les 98 molécules d'eau présentent la même caractéristique que les atomes de la chaîne protéique en ce qui concerne le facteur B . Les facteurs B des molécules d'eau des modèles hydrogéné en H_2O et deutérié en D_2O sont plus élevés que ceux des modèles hétérogènes et les explications doivent être sensiblement les mêmes que celles données dans le chapitre 6.2.3, et ne seront pas rediscutées pour les molécules d'eau.

Les variations topologiques entre molécules d'eau identiques, sont extrêmement faibles entre tous les modèles. En revanche, il existe de nombreuses molécules d'eau qui ne sont pas présentes ou qui ont été enlevées soit à cause d'une densité trop faible après affinement, soit à cause d'un facteur B trop haut ($>60 \text{ \AA}^2$) ou tout simplement à cause de leur absence.

6.4.2 Les liaisons hydrogène avec la protéine et la solvation

L'étude des différentes interactions du solvant avec la protéine peut être décrite pour un seul modèle, celui de la protéine deutériée en eau légère. Ce modèle est choisi à la vue de ses meilleures statistiques et parce que les observations des autres modèles sont sensiblement identiques. Les liaisons hydrogène ont été analysées grâce au programme *Contact (CCP4)*. Les probables liaisons hydrogène ou molécules d'eau à proximité de l'extrémité C terminale (de l'aspartate 171 à la tyrosine 173), ont été supprimées du fait de la grande flexibilité de cette zone.

Sur ce modèle, un total de 145 molécules d'eau a été positionné. Elles ont été numérotées de la molécule 3001 à la molécule 3145. Certaines molécules possèdent des occupations partielles du fait de la proximité d'un acide aminé en double conformation ou parce que la molécule a tout simplement deux positions très proches l'une de l'autre. C'est le cas pour la molécule d'eau 3094.

Les molécules ont la possibilité de faire des liaisons hydrogène avec des atomes accepteur ou donneur de proton. Une molécule d'eau peut, idéalement, établir jusqu'à 4 liaisons hydrogène, mais dans les structures, ces molécules d'eau avec une coordination possible de 4 sont très rares. Le tableau 6.7 représente les molécules d'eau avec les coordinations qu'elles établissent.

	<i>Coordination des molécules d'eau</i>				<i>Total</i>
	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	
<i>Nombre de molécules d'eau</i>	23 (15,9 %)	36 (24,8 %)	39 (26,9 %)	35 (24,1 %)	145 (100 %)
<i>Facteur thermique B (en \AA^2)</i>	31,7	28,3	26,3	18,5	26,3

Tableau 6.7 : Distribution des molécules d'eau en fonction de leur nombre de coordination.

Quand le nombre de liaisons hydrogène augmente, le facteur thermique B diminue. En fait, la molécule d'eau est de plus en plus stabilisée et donc son mouvement thermique sera moindre, ce qui sera visible au niveau de la diminution du facteur B . Il est possible de se rendre compte du même phénomène si l'analyse est faite sur le nombre de liaison à la protéine. En effet, lorsque le nombre de liaisons à la protéine augmente, le facteur B diminue (voir tableau 6.8). La protéine joue un rôle stabilisant vis à vis des molécules d'eau reliées, ce qui diminue ainsi le mouvement thermique.

	Nombre de liaisons à la protéine					Total
	0	1	2	3	4	
Nombre de molécules d'eau	12	53	55	21	4	145
Facteur thermique B (en Å^2)	34,6	29,5	24,0	20,3	18,3	26,3

Tableau 6.8 : Distribution des molécules d'eau en fonction du nombre de liaisons à la protéine.

Parmi toutes les molécules d'eau, 133 font, au minimum, une liaison avec les atomes de la protéine, soit 91,7 %. Cette eau est considérée comme la première couche d'hydratation. Seules, 12 molécules d'eau, soit 8,3 %, font uniquement des liaisons avec d'autres molécules d'eau. Elles représentent les couches supérieures de l'hydratation. Toutes les molécules sont listées en annexe E.2 avec les différents contacts qu'elles établissent.

Toutes les molécules d'eau qui sont comprises dans la première couche d'hydratation font des liaisons hydrogène différentes avec la protéine. Certaines interagissent avec les atomes d'azote ou le carbonyle de la chaîne principale, d'autres avec les atomes N et O des chaînes latérales. Dans le tableau 6.9, toutes les interactions avec des accepteurs ou des donneurs de protons de ce type sont listées, ainsi que celles avec les molécules d'eau et d'acétate. En moyenne, le facteur B est moins important quand la molécule d'eau considérée est liée à la chaîne principale, que quand elle est rattachée à une chaîne latérale. La liaison hydrogène est aussi moins grande (CP : 2,85 Å, CL : 2,82 Å, H₂O : 2,84 Å) quand les molécules d'eau sont reliées à un atome d'oxygène de la protéine ou d'une autre molécule d'eau. A l'inverse, une liaison avec un atome d'azote (CP : 2,93 Å et CL : 2,99 Å) est plus longue.

<i>Atomes</i>	<i>Nombres</i>	<i>Distance moyenne en Å</i>	<i>Facteur B moyen de l'eau en Å²</i>
<i>C=O de la CP</i>	78	2,85	23,47
<i>N de la CP</i>	33	2,93	22,62
<i>O des CL</i>	77	2,82	25,50
<i>N des CL</i>	50	2,99	24,87
<i>H-O-H</i>	134	2,84	25,02
<i>O des acétates</i>	5	3,01	24,91

Tableau 6.9 : Répartition des liaisons hydrogène en fonction de leur interaction avec les atomes de la protéine et du solvant. La distance qui sépare l'oxygène de l'eau à l'accepteur ou le donneur est donnée en Å, et le facteur *B* en Å². CP : Chaîne Principale, CL : Chaîne Latérale

Quand ce sont les atomes d'oxygène et d'azote de la protéine qui sont considérés, les résultats sont similaires à celui de Najmudin et al. (1993). 147 oxygènes de la fonction carbonyle de la chaîne principale sont associés par des liaisons hydrogène et 22 ne sont pas impliqués dans des liaisons. Par ces 147 interactions, 44,6% sont faites avec des atomes d'azote de la chaîne principale, 18,8% avec des atomes donneur de protons des chaînes latérales, et seulement 36,6% des interactions se font avec une molécule d'eau. Pour les 163 atomes d'azote de la chaîne principale qui sont considérés (les 3 résidus C terminaux sont enlevés ainsi que les N des 7 prolines), ceux ci sont répartis en, 29 n'ayant pas d'interaction et, 133 faisant au moins une liaison hydrogène. Ces 133 atomes interagissent pour 61,0% avec le C=O de la chaîne principale, 17,5% avec des atomes des chaînes latérales, et seulement 21,4% font une liaison avec une molécule d'eau. Pour synthétiser, les atomes d'azote et d'oxygène de la chaîne principale sont plus impliqués dans le maintien de la structure secondaire que dans les interactions avec les molécules d'eau environnantes. Ceci est peut être une des explications de la faible teneur en eau du nucléus du cristallin, là où les β et γ cristallines sont les protéines les plus représentées.

6.4.3 L'eau autour des résidus chargés

Barlow et Thornton (1983) avaient indiqué que les γ cristallines possédaient, sur leur surface, un grand nombre de charges, à comparer d'autres protéines. Cette densité de charge au niveau de la surface est très importante mais la charge totale est faible. Les charges positives et négatives sont bien équilibrées avec les arginines et les aspartates formant des

interactions intramoléculaires et intermoléculaires. Il a été suggéré que ces associations de charges dans les γ cristallines avait pour but de minimiser les interactions avec l'eau, ce qui devrait contribuer à un meilleur empilement dans le cristallin (Blundell et al. 1981; Wistow et al. 1983). Ces associations de charges sont aussi observables dans la γ E cristalline. Mais les groupes chargés situés à la surface, ont aussi la possibilité d'interagir avec des molécules d'eau (voir tableau 6.10), comme l'ont pensé Kumaraswamy et ses collaborateurs (1996) dans une étude de la γ B cristalline de bœuf.

Résidus	Nombre d'acides aminés dans la γ E cristalline	Nombre d'atomes chargés ou polaires dans la γ E cristalline	Liaisons hydrogène avec les molécules d'eau
Arginine	21	63	25
Asparagine	5	10	10
Aspartate	13	26	19
Glutamate	9	18	9
Glutamine	9	18	16
Histidine	9	18	6
Lysine	1	1	0
Serine	14	14	10
Thréonine	4	4	8
Tyrosine	16	16	18
Total	101	188	121

Tableau 6.10 : Contribution des différents résidus polaires ou chargés dans l'hydratation de la γ E cristalline d'après une étude similaire sur la γ B cristalline (Kumaraswamy 1996).

Comme dans la γ B, l'hydratation est dominée par des interactions avec les chaînes latérales des arginines, puis par celles des aspartates. Les 21 arginines de la γ E cristalline ont la possibilité de faire des liaisons hydrogène avec des molécules d'eau ou un autre résidu polaire comme dans le cas de l'arginine 79 (voir figure 6.20A). Mais elles ont aussi la possibilité de faire des ponts salins soit intramoléculaires, c'est le cas du couple Arg114-Asp106, soit intermoléculaires (voir tableau 6.4 et figure 6.11). Les interactions intermoléculaires sont réalisées par NH1 et NH2 terminaux, alors que les interactions intramoléculaires sont établies par l'azote N ϵ et aussi par les azotes terminaux.

Ce nombre considérable de résidus polaires ou chargés a des effets énormes sur l'organisation du solvant autour des γ cristallines. Il va avoir formation d'un réseau de molécules d'eau très structuré au voisinage proche de la protéine grâce aux groupes exposés.

Kumaraswamy et al. (1996) suggèrent que le grand nombre d'associations ioniques, plus spécialement de type intramoléculaire, va tendre à réduire les nombres de sites d'interactions avec les molécules d'eau. Ces regroupements de fonction acide et basique vont diminuer la charge totale de la surface de la protéine. Dans la γ B cristalline, il est décrit qu'à travers ces appariements, il y a la formation d'un réseau où les charges sont associées soit directement soit *via* les molécules d'eau. Ce type de réseau est présent dans la γ E cristalline (voir figure 6.20B). Les résidus chargés ne sont séparés que par une ou deux molécules d'eau dans ce type de réseau.

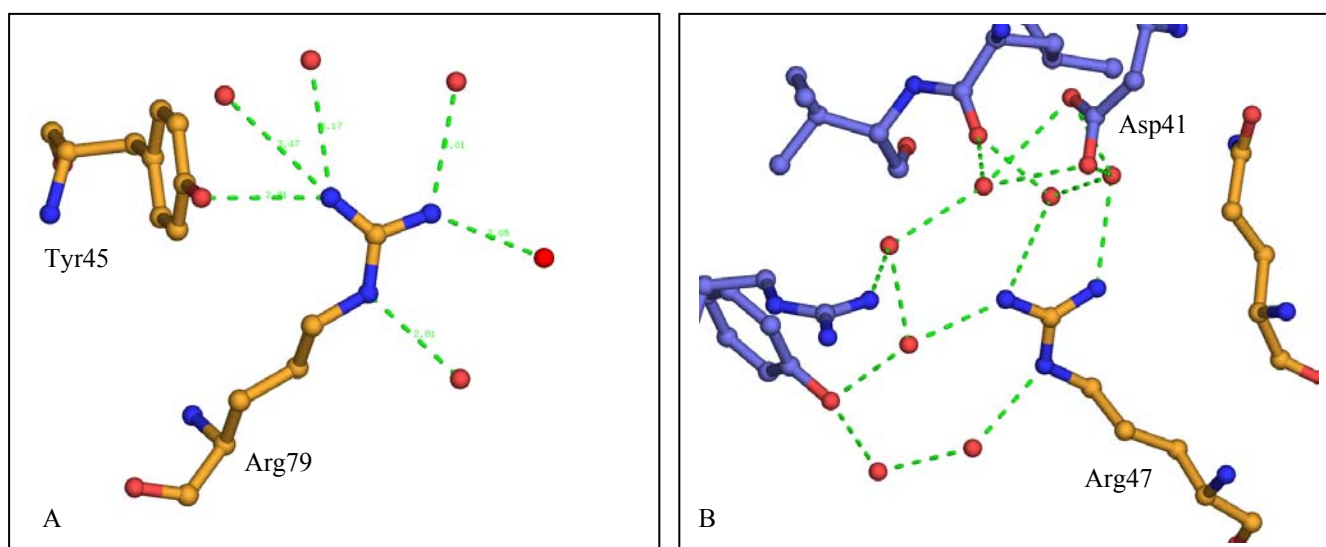


Figure 6.20 : Organisation des molécules d'eau autour des résidus chargés. A) Représentation de l'arginine 79 faisant des liaisons hydrogène avec des molécules d'eau et avec la tyrosine 45. B) Représentation du réseau de molécules d'eau qui connecte l'arginine 47 à l'aspartate 41 d'un équivalent cristallin.

6.4.4 Distribution des molécules d'eau

Les 145 molécules d'eau sont distribuées essentiellement sur la surface de la protéine (voir figure 6.21). Elles interagissent presque équitablement avec la chaîne principale (29,4%), la chaîne latérale (33,7 %) et les autres molécules d'eau (35,5%). Le reste des interactions est réalisé avec les acétates (1,5%).

Najmundin et al. (1993) ont décrit les contacts des 230 molécules d'eau identifiées autour de la γ B cristalline de bœuf résolue à 1,47 Å. Ils ont trouvé seulement 5 molécules d'eau positionnées faisant deux liaisons entre les protéines et ils ont situé 13 molécules d'eau faisant un lien entre les domaines N et C terminaux.

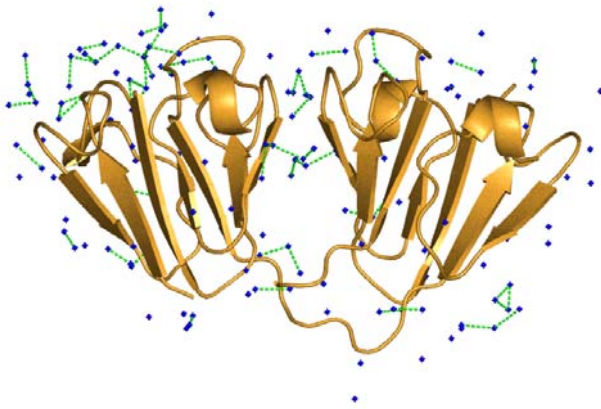


Figure 6.21 : Représentation des molécules d'eau (en bleu) dispersées tout autour de la γE cristalline. Les pointillés verts correspondent aux différents réseaux formés par l'eau.

Dans la protéine γE cristalline, 13 molécules d'eau (3011, 3027, 3071, 3079, 3081, 3082, 3085, 3104, 3117, 3118, 3119, 3133 et 3135) ont été positionnées de façon intermoléculaire. En revanche il y a 14 molécules d'eau (3001, 3003, 3004, 3005, 3020, 3035, 3039, 3041, 3043, 3061, 3068, 3084, 3099 et 3169) qui interagissent avec les deux domaines de la protéine. Ces molécules forment deux réseaux très bien définis.

Toutes les molécules d'eau ne sont pas équitablement réparties entre les deux domaines. Il y a plus de molécules d'eau du côté N terminal (89) que du côté C terminal (66) de la protéine, et certaines sont en contact avec les deux domaines (12).

Certaines des molécules d'eau interagissant avec les feuillettes β , se joignent au motif en clé grecque par ses bords (voir figure 6.22). Dans ce cas, elles étendent la structure en feuillet comme l'avaient remarqué Thanki et al (1991) sur une analyse de 24 structures à haute résolution.

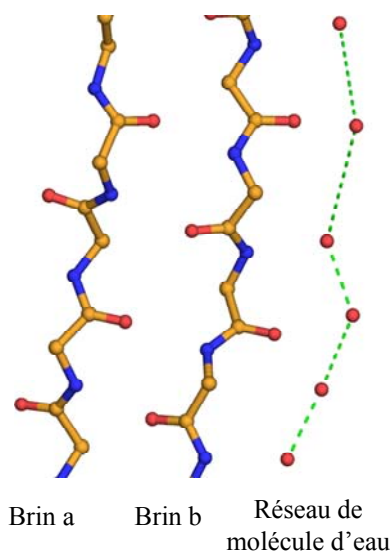
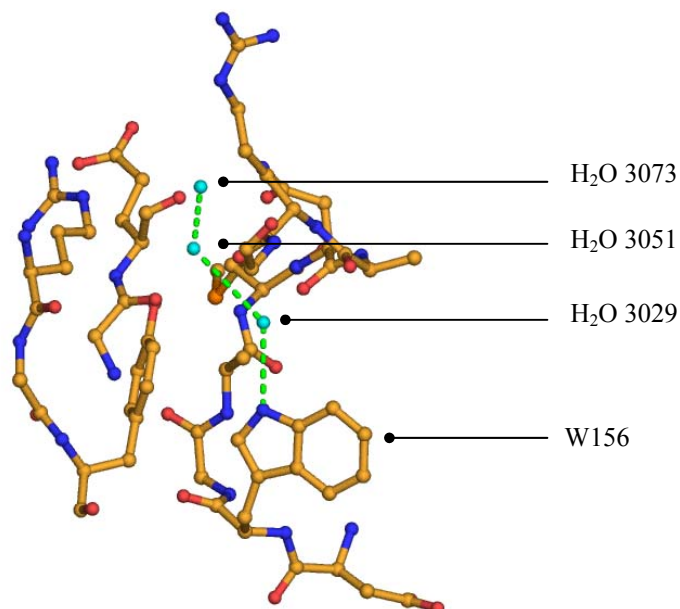


Figure 6.22 : Représentation de l'extension du motif 1 en clé grecque au niveau du brin b par les molécules d'eau (sphère bleu).

Il existe dans le modèle, deux molécules d'eau fortement enfouies dans la protéine. La 3002 pour le domaine N terminal et la 3029 pour le domaine C terminal, font une liaison hydrogène avec le noyau indole de deux tryptophanes symétriquement équivalents dans les domaines N et C terminaux (W68 et W156 respectivement). Celles-ci sont connectées aux molécules d'eau de surface par une sorte de canal pour le solvant où se trouvent les molécules 3051 et 3073 (ou la 3009 pour l'autre domaine) (voir figure 6.23).

Figure 6.23 : Représentation d'une des deux molécules d'eau les plus enfouies dans la protéine qui fait une liaison hydrogène avec le tryptophane 156. Cette molécule, la 3029, est reliée à la surface par d'autres molécules d'eau, 3051 et 3073.



Ces molécules d'eau enfouies sont observables dans toutes les structures des γ cristallines déjà résolues étant donné que le tryptophane est un résidu hautement conservé dans les γ cristallines. De ce fait, elles semblent avoir un rôle essentiel dans la stabilisation des γ cristallines.

6.4.5 Réseaux de molécules d'eau

Dans la γ E cristalline, il y a plusieurs réseaux de molécules d'eau. Le réseau le plus étendu représente une association de 13 molécules d'eau qui forment des liens entre les molécules de la première et de la seconde couche d'hydratation et les molécules d'eau d'un équivalent cristallin. Ce genre de réseau contribue grandement à l'empilement des molécules dans le cristal. La figure 6.24 représente les plus grands réseaux trouvés autour de la γ E cristalline.

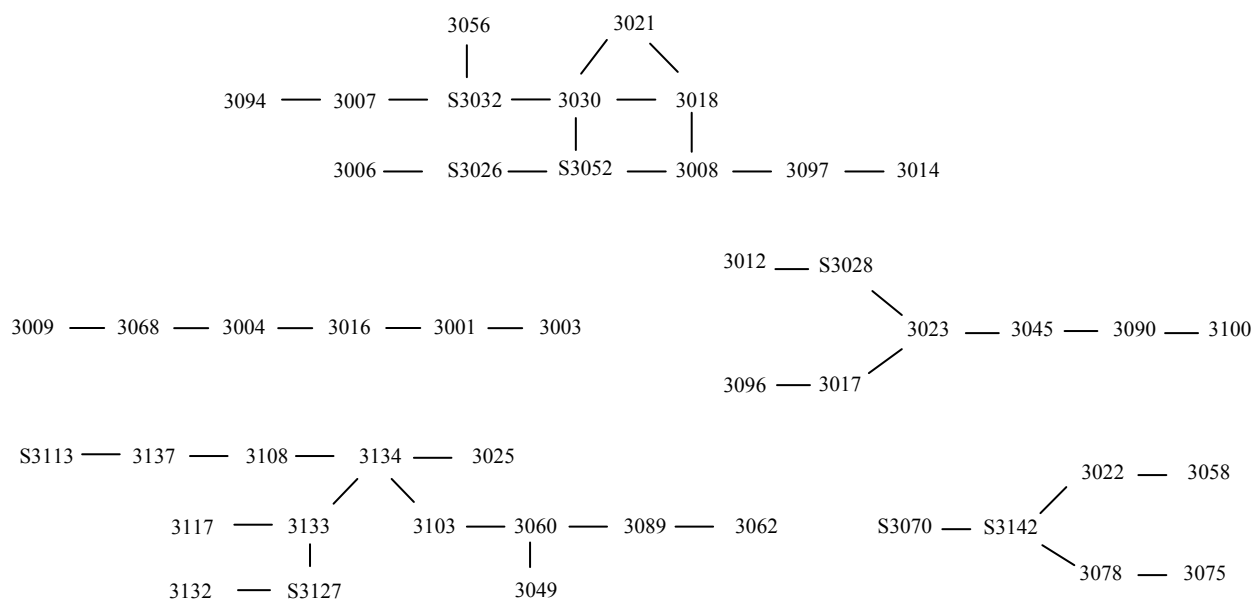


Figure 6.24 : Schéma des 5 plus grands réseaux de molécules d'eau de la γE cristalline deutériée en H_2O . S signifie que cette molécule d'eau appartient à un équivalent cristallin.

6.5 Bilans

Les structures cristallographiques de la γE cristalline de rat ont été étudiées sous différentes formes isotopiques. L'état cristallin de la protéine est très proche de l'état physiologique au regard de la concentration. Les cristaux de γE cristalline ont une concentration de 850 mg.ml^{-1} (voir annexe F), ce qui peut être considéré comme équivalent à la région centrale du cristallin de rat dont la concentration protéique est de 890 mg.ml^{-1} (Philipson 1969).

La protéine γE cristalline, qu'elle soit sous forme hydrogénée ou deutériée, en solvant H_2O ou D_2O , a une structure tridimensionnelle identique. La déviation de la chaîne principale n'excède pas $0,156 \text{ \AA}$, et celle des chaînes latérales ne dépasse pas $0,158 \text{ \AA}$. En revanche, certaines molécules d'eau ne sont pas présentes dans tous les modèles. Mais ceci est certainement la cause des résolutions et des statistiques après affinement, différentes. Près d'une centaine de molécules d'eau sont quand même présentes dans toutes les structures.

Il est très intéressant de noter la différence du facteur thermique B en fonction du marquage isotopique. Lorsque les deutériums remplacent les hydrogènes, que ce soit sur des positions échangeables ou non, il y a toujours une diminution du facteur d'agitation thermique, mis à part pour la protéine deutériée en D_2O . Ce dernier cas ne peut faire partie du

comparatif à cause d'une moins bonne résolution, qui va certainement augmenter les valeurs du facteur B de quelques Å^2 . A de nombreuses reprises dans la littérature, l'influence du deutérium sur la rigidité des liaisons covalentes a été discutée (Cioni et Strambini 2002). Sa présence a contribué à la diminution de l'agitation thermique et, peut être, au positionnement de certains atomes lors de l'affinement. Cet avantage pourrait être développé par les cristallographes afin de localiser certains atomes ou groupes d'atomes. Mais tout ceci reste hypothétique car il faudrait mener une étude comparative entre deux jeux de données collectés sur le même cristal après trempage en solution hydrogénée et en solution deutériée. Or les dommages radiatifs permettent rarement ce genre d'étude hormis sur des cristaux de grande taille.

L'obtention des jeux de données, traités jusqu'à 1,36 Å de résolution, a permis d'observer des détails qui n'avaient pas été perçus avec la structure de la protéine résolue à 2,3 Å (Norledge et al. 1997). Les contacts cristallins sont certainement les mêmes qu'il existe *in vivo* car les concentrations protéiques sont identiques dans le cristal et dans la cellule. La haute flexibilité du tripeptide RGD et de la partie C terminale relance l'hypothèse d'un site d'interaction possible avec d'autres protéines. L'absence de ponts disulfure et l'empilement de résidus aromatiques confirment les observations de la structure résolues à 2,3 Å .

Mais l'élément de la protéine, le plus captivant, est la présence de la double conformation de la boucle *cd* du motif 3. Cette boucle qui est certainement responsable de propriétés différentes vis-à-vis du changement de phase, a une double conformation qui n'avait pas été observée dans les autres γ cristallines. Cette double conformation permet pourtant de garder un cœur hydrophobe composé par d'imposants résidus tels que les phénylalanines. Ceci discrédite l'hypothèse d'avoir dans la γE cristalline, les deux formes de la boucle, correspondant aux formes de la boucle des deux groupes de cristallines de haute et basse température de transition.

Une étude de la solvation de la γE cristalline, a été menée pour servir de base pour la cristallographie aux neutrons. Les 145 molécules d'eau et les 3 acétates sont distribués tout autour de la protéine. Seulement cinq réseaux de molécules d'eau ont été identifiés. Ils interagissent par de liaisons hydrogène avec la protéine puisque seulement 8,3 % des molécules d'eau se trouvent sur les couches d'hydratation supérieures. En effet, les interactions faites pour la γE cristalline sont essentiellement réalisées avec la protéine elle-même (intra ou intermoléculaires). L'étude neutronique devrait alors préciser la nature exacte

de ces interactions, tout comme l'état de charge de nombreux résidus de surface. Ce nombre de 145 molécules d'eau est faible en comparaison des 394 molécules d'eau référencées pour la γ B cristalline de bœuf (Kumaraswamy 1996), mais il reste en accord avec celui de la γ E résolues par Norledge et al. (1997) et avec le faible contenu de la maille cristalline qui est de 37%. De plus, ceci correspond à faible contenu en eau des cellules présentes au centre du cristallin (voir chapitre 2).

Ayant une bonne base avec la diffraction aux rayons X, la suite logique est d'obtenir des cristaux de volume supérieur à 1 mm^3 , afin d'envisager la diffraction neutronique. Le chapitre suivant, va décrire les différentes techniques testées et les différents problèmes rencontrés pour finir sur les conclusions et les perspectives de cette thèse.

Bibliographie :

- Acharya, R., Fry, E., Stuart, D., Fox, G., Rowlands, D., and Brown, F. 1989. The three-dimensional structure of foot-and-mouth disease virus at 2.9 Å resolution. *Nature* **337**: 709-716.
- Balasubramanian, D., Du, X., and Zigler, J.S., Jr. 1990. The reaction of singlet oxygen with proteins, with special reference to crystallins. *Photochem Photobiol* **52**: 761-768.
- Barlow, D.J., and Thornton, J.M. 1983. Ion-pairs in proteins. *J Mol Biol* **168**: 867-885.
- Basak, A., Bateman, O., Slingsby, C., Pande, A., Asherie, N., Ogun, O., Benedek, G.B., and Pande, J. 2003. High-resolution X-ray crystal structures of human gammaD crystallin (1.25 Å) and the R58H mutant (1.15 Å) associated with aculeiform cataract. *J Mol Biol* **328**: 1137-1147.
- Benjamin, L., and Benson, G. 1962. A deuterium isotope effect on excess enthalpy of methanol-water solutions. *J. Phys. Chem.:* 858-861.
- Betzl, C., Gourinath, S., Kumar, P., Kaur, P., Perbandt, M., Eschenburg, S., and Singh, T.P. 2001. Structure of a serine protease proteinase K from *Tritirachium album limber* at 0.98 Å resolution. *Biochemistry* **40**: 3080-3088.
- Blundell, T., Lindley, P., Miller, L., Moss, D., Slingsby, C., Tickle, I., Turnell, B., and Wistow, G. 1981. The molecular structure and stability of the eye lens: x-ray analysis of gamma-crystallin II. *Nature* **289**: 771-777.
- Broide, M.L., Berland, C.R., Pande, J., Ogun, O.O., and Benedek, G.B. 1991. Binary-liquid phase separation of lens protein solutions. *Proc Natl Acad Sci U S A* **88**: 5660-5664.
- Chirgadze, Y., Driessen, H., Wright, G., Slingsby, C., Hay, R.E., and Lindley, P.F. 1996. Structure of the bovine Eye lens gammaD crystallin at 1.95Å. *Acta Crysta* **D52**: 712-721.
- Cioni, P., and Strambini, G.B. 2002. Effect of heavy water on protein flexibility. *Biophys J* **82**: 3246-3253.
- Garland, D., Russell, P., and Zigler, J.S., Jr. 1988. The oxidative modification of lens proteins. *Basic Life Sci* **49**: 347-352.
- Guptasarma, P., and Balasubramanian, D. 1992. Dityrosine formation in the proteins of the eye lens. *Curr Eye Res* **11**: 1121-1125.
- Guptasarma, P., Balasubramanian, D., Matsugo, S., and Saito, I. 1992. Hydroxyl radical mediated damage to proteins, with special reference to the crystallins. *Biochemistry* **31**: 4296-4303.
- Hynes, R.O. 1987. Integrins: a family of cell surface receptors. *Cell* **48**: 549-554.
- Kumaraswamy, V.S., Lindley, P. F., Slingsby, C., Glover, I. D. 1996. An eye lens protein-water structure: 1.2 angstrom resolution structure of gamma B-crystallin at 150K. *Acta Crystallogr D Biol Crystallogr* **52**: 611.
- Lee, B., and Richards, F.M. 1971. The interpretation of static protein structures: estimation of static accessibility. *J Mol Biol* **55**: 379-400.
- Matthews, B. 1975. *The Proteins*, pp. 403-590.
- McDermott, M.J., Gawinowicz-Kolks, M.A., Chiesa, R., and Spector, A. 1988. The disulfide content of calf gamma-crystallin. *Arch Biochem Biophys* **262**: 609-619.
- Najmudin, S., Nalini, V., Driessen, H., Slingsby, C., Blundell, T., S., M.D., and Lindley, P.F. 1993. Structure of the bovine eye lens protein gammaB(gammaII)-crystallin at 1.47 Å. *Acta Crystallogr D Biol Crystallogr* **49**: 223-233.
- Norledge, B.V., Hay, R.E., Bateman, O.A., Slingsby, C., and Driessen, H.P. 1997. Towards a molecular understanding of phase separation in the lens: a comparison of the X-ray structures of two high Tc gamma-crystallins, gammaE and gammaF, with two low Tc gamma-crystallins, gammaB and gammaD. *Exp Eye Res* **65**: 609-630.
- Philipson, B. 1969. Distribution of protein within the normal rat lens. *Invest Ophthalmol* **8**: 258-270.
- Pierschbacher, M.D., and Ruoslahti, E. 1987. Influence of stereochemistry of the sequence Arg-Gly-Asp-Xaa on binding specificity in cell adhesion. *J Biol Chem* **262**: 17294-17298.
- Rao, P.V., and Zigler, J.S., Jr. 1991. Zeta-crystallin from guinea pig lens is capable of functioning catalytically as an oxidoreductase. *Arch Biochem Biophys* **284**: 181-185.
- Reed, J., Hull, W.E., von der Lieth, C.W., Kubler, D., Suhai, S., and Kinzel, V. 1988. Secondary structure of the Arg-Gly-Asp recognition site in proteins involved in cell-surface adhesion. Evidence for the occurrence of nested beta-bends in the model hexapeptide GRGDSP. *Eur J Biochem* **178**: 141-154.

- Ruoslahti, E., and Pierschbacher, M.D. 1986. Arg-Gly-Asp: a versatile cell recognition signal. *Cell* **44**: 517-518.
- Ruoslahti, E., and Pierschbacher, M.D. 1987. New perspectives in cell adhesion: RGD and integrins. *Science* **238**: 491-497.
- Siezen, R.J., Wu, E., Kaplan, E.D., Thomson, J.A., and Benedek, G.B. 1988. Rat lens gamma-crystallins. Characterization of the six gene products and their spatial and temporal distribution resulting from differential synthesis. *J Mol Biol* **199**: 475-490.
- Slingsby, C., and Clout, N.J. 1999. Structure of the crystallins. *Eye* **13 (Pt 3b)**: 395-402.
- Takagi, J. 2004. Structural basis for ligand recognition by RGD (Arg-Gly-Asp)-dependent integrins. *Biochem Soc Trans* **32**: 403-406.
- Tew, D., and Ortiz de Montellano, P.R. 1988. The myoglobin protein radical. Coupling of Tyr-103 to Tyr-151 in the H₂O₂-mediated cross-linking of sperm whale myoglobin. *J Biol Chem* **263**: 17880-17886.
- Thanki, N., Umrana, Y., Thornton, J.M., and Goodfellow, J.M. 1991. Analysis of protein main-chain solvation as a function of secondary structure. *J Mol Biol* **221**: 669-691.
- Thomson, J.A., Schurtenberger, P., Thurston, G.M., and Benedek, G.B. 1987. Binary liquid phase separation and critical phenomena in a protein/water solution. *Proc Natl Acad Sci USA* **84**: 7079-7083.
- Tuominen, V.U., Myles, D.A., Dauvergne, M.T., Lahti, R., Heikinheimo, P., and Goldman, A. 2004. Production and preliminary analysis of perdeuterated yeast inorganic pyrophosphatase crystals suitable for neutron diffraction. *Acta Crystallogr D Biol Crystallogr* **60**: 606-609.
- Vollhardt, and Schore. 1995. La stabilité particulière du sextet d'électrons dans une structure cyclique. In *Traité de chimie organique*. (ed. D. Université), pp. 549-593.
- White, H.E., Driessen, H.P., Slingsby, C., Moss, D.S., and Lindley, P.F. 1989. Packing interactions in the eye-lens. Structural analysis, internal symmetry and lattice interactions of bovine gamma IVA-crystallin. *J Mol Biol* **207**: 217-235.
- Wistow, G., Turnell, B., Summers, L., Slingsby, C., Moss, D., Miller, L., Lindley, P., and Blundell, T. 1983. X-ray analysis of the eye lens protein gamma-II crystallin at 1.9 Å resolution. *J Mol Biol* **170**: 175-202.
- Wistow, G.J., and Piatigorsky, J. 1988. Lens crystallins: the evolution and expression of proteins for a highly specialized tissue. *Annu Rev Biochem* **57**: 479-504.
- Wlodawer, A., Borkakoti, N., Moss, D., and Howlin, B. 1986. Comparison of two independently refined models of ribonuclease-A. *Acta Cryst.* **B42**: 379-387.
- Wlodawer, A., Savage, H., and Dodson, G. 1989. Structure of insulin: results of joint neutron and X-ray refinement. *Acta Crystallogr B* **45 (Pt 1)**: 99-107.
- Xiong, J.P., Stehle, T., Zhang, R., Joachimiak, A., Frech, M., Goodman, S.L., and Arnaout, M.A. 2002. Crystal structure of the extracellular segment of integrin alpha Vbeta3 in complex with an Arg-Gly-Asp ligand. *Science* **296**: 151-155.

Chapitre 7

Cristallisation de la γ E cristalline, conclusions et perspectives

Abstract

Obtaining crystals is currently the bottleneck in protein structure determination by X-ray crystallography. However, as it was mentioned in chapter 2, in view of the low neutron fluxes, it is necessary to use large crystals ($> 1 \text{ mm}^3$) for the collecting diffraction data. The intent of this chapter is to develop some useful techniques for crystallising γ E crystalline protein. The most frequently used crystallisation methods were the vapour diffusion and the dialyse techniques to determine crystallisation conditions. To obtain single crystals, micro-, normal and macro-seeding techniques, and the technique of crystallisation in capillaries were used with success, but these crystals were like plates with a very fine depth (third dimension). The growth of this dimension was never successful but with a thermal technique, we hope to obtain huge crystals. To finish this chapter, conclusions and perspectives are developed.

7.1 Des cristaux de grande taille pour les neutrons

L'obtention de cristaux est une des étapes déterminantes et nécessairement obligatoires pour acquérir la structure d'une protéine par cristallographie. Ces dernières années, plusieurs groupes de scientifiques cherchent à développer avec succès, des techniques de cristallisation à haut débit afin de minimiser les quantités de protéines (Santarsiero et al. 2002). En revanche, comme il a déjà été précisé dans le chapitre 2, pour la cristallographie aux neutrons, le problème ne réside pas seulement dans le fait d'obtenir des cristaux mais il est essentiel d'avoir des cristaux de taille importante. En effet, sachant que le flux de neutrons qui arrive sur l'échantillon est faible et pour éviter des collectes de données trop longues en temps, des cristaux dont le volume dépasse 1 mm^3 sont indispensables. Tout au long de ma thèse, différentes techniques ont été utilisées pour obtenir des cristaux uniques, avec une troisième dimension et de grande taille. Ce chapitre va énoncer les différentes méthodes utilisées pour enfin conclure cette thèse et définir les perspectives éventuelles à mes travaux.

7.2 Cristallisation de la γ E cristalline

Le principe fondamental de cristallisation doit se faire suivant la courbe de solubilité de la protéine. Pour qu'un cristal se développe, des précipitants sont ajoutés de manière à croiser la courbe de solubilité et d'obtenir une solution sursaturée (voir figure 7.1A). La détermination des conditions est souvent un long travail basé sur des essais, des erreurs et un criblage poussé. Du comportement physico-chimique de la protéine à l'influence des sels, du pH ou de la température, de nombreux paramètres ont un impact sur la croissance cristalline. Et il existe de nombreuses techniques pour cristalliser les macromolécules biologiques (Ducruix et Giegé 1992; Bergfors 1999). Pour la γ E cristalline, la diffusion de vapeur et la dialyse, sont les deux techniques qui ont permis d'obtenir les conditions de cristallisation de la γ E cristalline. Pour ensuite, tenter d'obtenir des cristaux uniques tridimensionnels de taille plus importante, des techniques comme la cristallisation en capillaire, la croissance après ensemencement (micro-, normal et macro-) et la croissance contrôlée par effet de température, ont été testées.

7.2.1 Méthodes de diffusion de vapeur et de dialyse

Ces techniques sont généralement les plus populaires car elles sont faciles à réaliser, elles requièrent peu d'échantillon et permettent des criblages (surtout la diffusion de vapeur). La cristallisation par dialyse est une variante de la technique de diffusion de vapeur. Dans la

micro-dialyse, l'échantillon protéique est séparé de l'agent précipitant par une membrane semi-perméable qui permet aux petites molécules tels que les ions, les additifs, les tampons et les sels de diffuser mais pas la protéine. Des boutons de dialyse (Hampton Research) comme en figure 7.1A, ont été utilisés. La migration de l'état de la solution sur le diagramme de solubilité, où la concentration protéique est constante et la concentration en agent précipitant augmente, est représentée dans la figure 7.2.

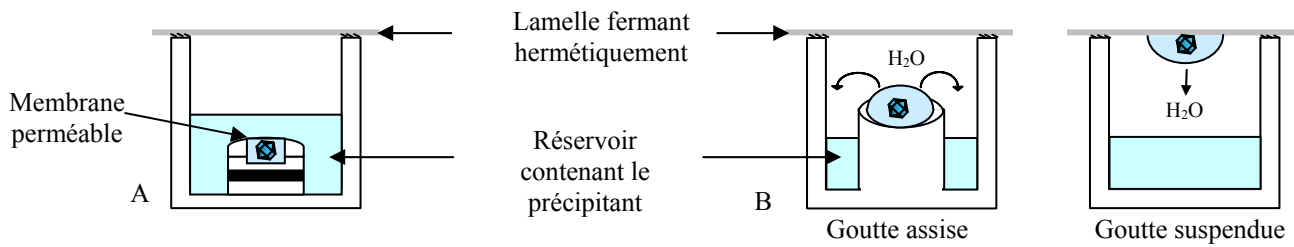


Figure 7.1 : Cristallisation par diffusion de précipitant ou de vapeur. A) Schéma d'un bouton de dialyse recouvert d'une membrane semi perméable qui va laisser diffuser l'agent précipitant dans le réservoir contenant la protéine. B) Schéma d'une goutte assise (a) et d'une goutte suspendue. L'évaporation d'eau va conduire à un équilibre entre le réservoir et la goutte.

La diffusion de vapeur a été réalisée en goutte assise et en goutte suspendue (voir figure 7.1B). Dans les deux cas, la protéine et l'agent cristallisant vont se concentrer, suite à l'évaporation d'eau de la goutte. Dans certains cas, la solution atteint la zone de nucléation, puis les concentrations de la protéine et de l'agent précipitant vont diminuer pour arriver dans une zone métastable où le cristal va croître (voir figure 7.2).

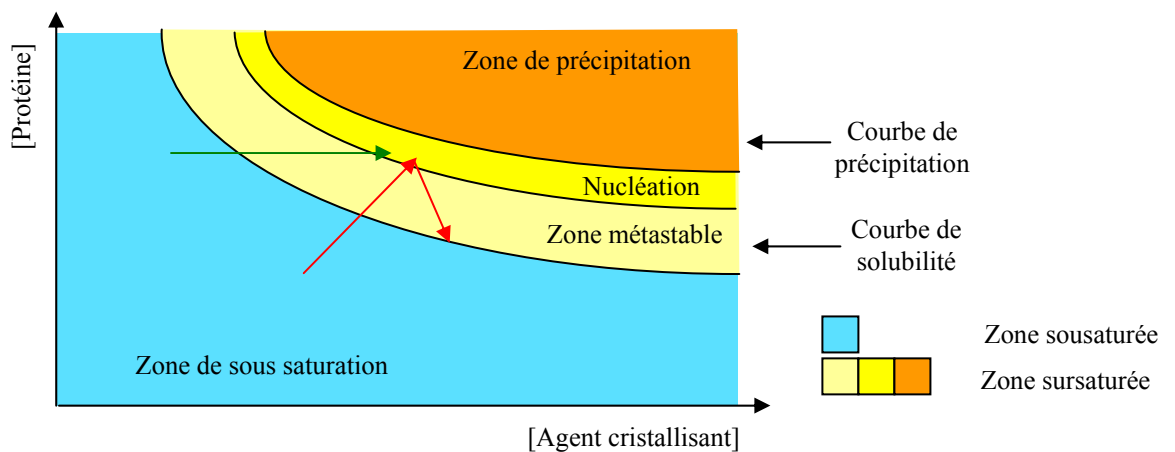


Figure 7.2 : Diagramme de solubilité et corrélation entre les concentrations en protéine et en agent cristallisant dans des expériences de cristallisation par dialyse (en vert), et par diffusion de vapeur (en rouge).

Comme il a été mentionné dans le chapitre 5, de nombreuses conditions de cristallisation ont été trouvées (voir tableau 5.2) par criblage avec les kits de Hampton Research et Molecular Dimension Limited. Mais la plupart du temps, de gros amalgames

d'aiguilles étaient obtenus (voir figure 7.3A). Après optimisation du pH, de la concentration en protéine et de la température, ces « oursins » ont donné des assemblages de plaques ; la seconde dimension était importante, mais rarement la troisième (voir figure 7.3B).

7.2.2 Obtention de cristaux unique

Afin de limiter la combinaison de plusieurs plaques qui est essentiellement due à une très forte nucléation, différentes techniques ont été testées. Parmi ces dernières, les techniques d'ensemencement et de cristallisation en capillaire ont été les plus prometteuses.

7.2.2.1 Ensemencement

La cristallisation à partir d'ensemencement permet d'éviter certains problèmes inhérents à la formation de cristaux formés spontanément car la nucléation spontanée requière un plus haut degré de supersaturation proche de la précipitation (voir figure 7.2). Cette zone de nucléation peut être très petite suivant les conditions. Quand la sursaturation est réduite, la chance de former un nucléus stable est réduite (Ducruix et Giegé 1992). En revanche, l'apport de nucléus au niveau de la zone métastable peut favoriser et stimuler la croissance cristalline à partir de ces nucléus. Par des dilutions de ces nucléus, la nucléation, et donc la formation, d'assemblage sont diminuées. Parmi les différentes possibilités d'ensemencement (micro-, macro- et ensemencement en « ligne »), la méthode du micro-ensemencement a fourni à de nombreuses reprises, des plaques individuelles (voir figure 7.3C).

7.2.2.2 Cristallisation en capillaire par la technique d'acuponcture

La cristallisation en capillaire par la technique acuponcture permet d'établir un gradient en protéine et en précipitant dans le capillaire grâce à une diffusion lente du précipitant à travers un gel d'agarose (Garcia-Ruiz 1994; Garcia-Ruiz et Moreno 1997; Garcia-Ruiz et al. 1998). Le long du capillaire, différentes zones reproduisant le diagramme de solubilité vont être créés (voir figure 7.4). A certains endroits, la nucléation est faible et des cristaux uniques vont pouvoir grandir (voir figure 7.3D), mais ces cristaux se présentent toujours sous forme de plaques et la taille est généralement limitée par le diamètre interne du capillaire.

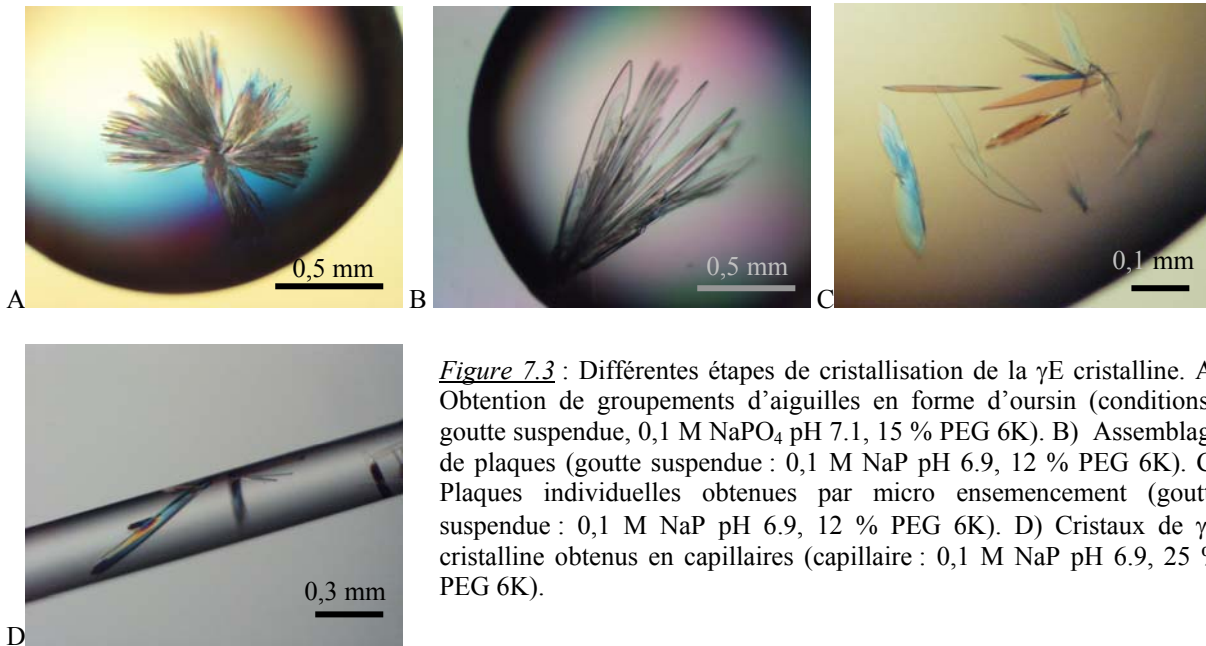
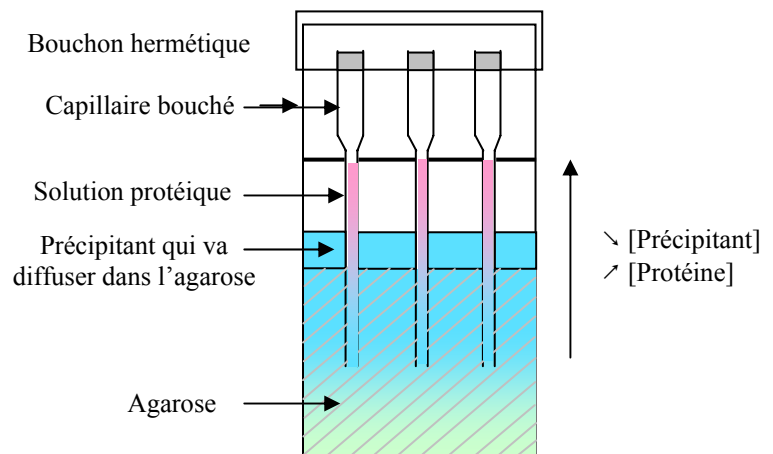


Figure 7.3 : Différentes étapes de cristallisation de la γ E cristalline. A) Obtention de groupements d'aiguilles en forme d'oursin (conditions : goutte suspendue, 0,1 M NaPO₄ pH 7.1, 15 % PEG 6K). B) Assemblage de plaques (goutte suspendue : 0,1 M NaP pH 6.9, 12 % PEG 6K). C) Plaques individuelles obtenues par micro ensemencement (goutte suspendue : 0,1 M NaP pH 6.9, 12 % PEG 6K). D) Cristaux de γ E cristalline obtenus en capillaires (capillaire : 0,1 M NaP pH 6.9, 25 % PEG 6K).

Figure 7.4 : Système de cristallisation en capillaire. La solution protéique est enfermée dans un capillaire fermé. Ce capillaire est introduit dans une couche d'agarose de faible pourcentage, et le précipitant est déposé en surface de l'agarose. Le précipitant et la protéine vont en diffusant à travers le gel, former doucement deux gradients inverses. Ceci permet d'explorer des nombreuses zones du diagramme de solubilité.



7.2.3 Des cristaux en trois dimensions

L'obtention d'une troisième dimension de taille raisonnable reste l'étape la plus importante et la plus limitante pour envisager la cristallographie aux neutrons avec la protéine γ E cristalline. Dans de trop nombreuses tentatives, des cristaux de grandes tailles ont été obtenus (voir chapitre 5.4.1). Ces derniers présentaient la particularité de ne pas être en fait un seul cristal, mais un empilement de plaques cristallines. Cette superposition de réseaux cristallins a mené à des dédoublements voir triplements des tâches de diffraction en cristallographie aux rayons X.

Après tous les essais effectués, il s'est avéré que je n'ai jamais réussi à obtenir une dernière dimension utilisable sur un cristal unique par les précédentes techniques citées. Par ces techniques, il est très difficile de contrôler la croissance cristalline. Une collaboration avec Monika Budayova-Spano (EMBL) a été entreprise afin de contrôler cette croissance par effet de température. Ce paramètre physique est très important pour la croissance cristalline

(Ducruix et Giegé 1992; Bergfors 1999). Après ensemencement de la solution de cristallisation (protéine, tampon et agent précipitant) avec des germes cristallins, à une température élevée, la croissance du cristal ensemencé est observée. Les essais de cristallogénèse ont démontré que la solubilité de notre système protéique augmente avec la température (solubilité directe). Ainsi en abaissant la température, nous favorisons la croissance cristalline. A l'inverse, à l'apparition d'imperfections, la température est augmentée avec précision pour dissoudre les bords imparfaits du cristal (Budayova-Spano 2005). Des cristaux uniques de γ E cristalline avec des dimensions sont aux environs de 1mm x 0,4mm x 0,05 mm lors de premiers essais (voir figure 7.5), mais la troisième dimension semble faire toujours défaut pour envisager la cristallographie aux neutrons.

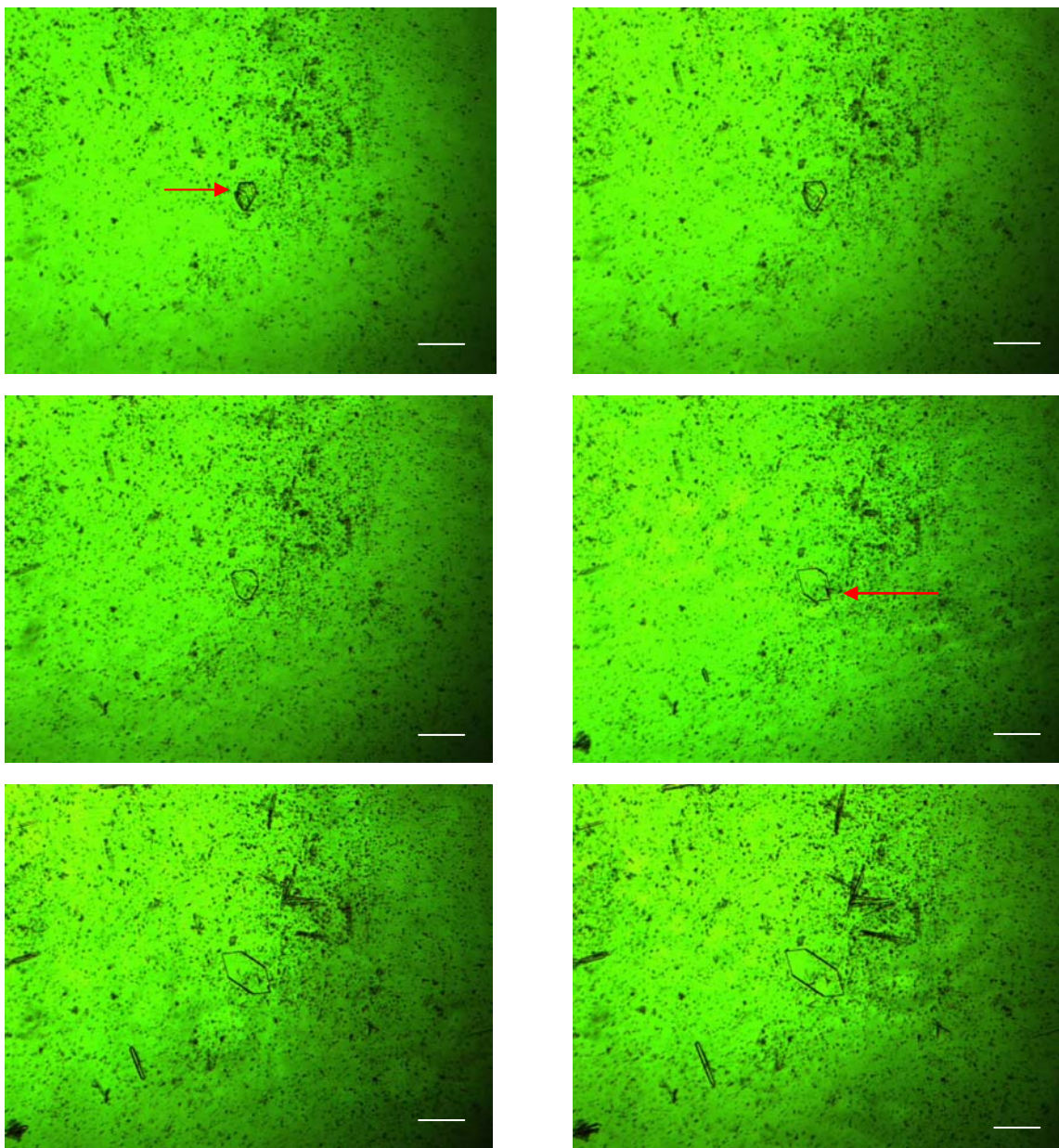


Figure 7.5 : Induction de la croissance cristalline de la γ E cristalline par variation de température. Les défauts (flèches rouges par exemple) du cristal sont corrigés en faisant varier avec précision la température. Le temps total de cette séquence est d'environ 10 jours et la barre d'échelle représente 0,5 mm.

7.3 Conclusions et perspectives

Localiser la position d'atomes d'hydrogène ou de solvant dans les protéines est essentiel pour répondre aux questions concernant les mécanismes enzymatiques, les interactions avec des ligands, les effets de solvant et dynamiques structurales. La cristallographie aux neutrons permet la localisation de ces atomes d'hydrogène importants dans des structures biologiques sous forme cristalline, qui ne sont pas visibles avec une analyse aux rayons X uniquement.

Afin de comprendre les différents types d'interactions dans les γ cristallines, protéines clés impliquées dans l'opacification du cristallin de certaines formes de cataractes de mammifères, les études de cette thèse ont été menées pour envisager une étude en cristallographie des neutrons. Cette étude devrait être grandement facilitée par l'utilisation des protéines deutériées. Le fait de substituer les atomes d'hydrogène par des atomes de deutérium va réduire le bruit de fond causé par la diffusion incohérente des hydrogènes, renforçant les données à haute résolution et fournissant un gain dans le rapport signal sur bruit. Cette substitution a la possibilité d'être faite sur les atomes d'hydrogène échangeables de la protéine par trempage dans des solutions deutériées, ce qui ne représente que 25% des hydrogènes d'une protéine en général. Pour effectuer un remplacement total, il faut que ce remplacement intervienne au niveau même de la synthèse protéique, c'est à dire au niveau de l'expression *in vivo*.

Les premiers objectifs de ma thèse étaient donc de développer des systèmes fiables d'expression pour une γ cristalline marquée avec le deutérium : un système bactérien (*E. coli*) et un système lévurien (*P. pastoris*). L'expression de cette protéine en milieu minimum hydrogéné n'a pas été identique suivant les deux systèmes et l'utilisation du système de levure s'est retrouvée limitante à cause du faible taux d'expression. En revanche, la protéine a été surexprimée en grande quantité, dans le système bactérien en milieu hydrogéné. *Escherichia coli* est un bon candidat à la production de protéine recombinante, en revanche, cette bactérie a une croissance cellulaire fortement ralentie en milieu minimum deutérié. Pendant cette thèse, nous avons développé un système d'adaptation au deutérium, sur milieu minimum solide puis en milieu minimum liquide. Cette adaptation a permis de réduire les temps de culture qui peuvent rester très longs en comparaison d'une culture hydrogénée. Pour palier aux coûts onéreux des produits deutériés, nous avons développé, au sein du laboratoire de deutériation de l'ILL, l'utilisation de bio-réacteur pour la production de biomasse bactérienne et donc de

protéines deutériées. La production de la γ E cristalline a été effectuée en milieu hydrogéné et en milieu deutérié avec des rendements légèrement inférieurs.

Avant d'envisager une étude sur une protéine marquée au deutérium, il semblait important de vérifier plusieurs points : le pourcentage de substitution, ainsi qu'aucune modification physico-chimique ou structurale de la protéine, afin que l'étude de diffraction des neutrons soit identique à une étude de diffraction aux rayons X.

Après leur purification, les cristallines hydrogénées et deutériées ont été caractérisées par spectrométrie de masse. Cette étude a révélé une différence dans le taux de modification post traductionnelle qu'est la coupure de la méthionine N-terminale par la MAP. Afin de travailler avec des solutions protéiques homogènes, la protéine MAP a été coexprimée avec la γ cristalline sous forme hydrogénée. Les spectres de masse des protéines hydrogénées et deutériées montraient que les protéines étaient pures et que l'incorporation isotopiques étaient de 98%. N'ayant pas d'activité biochimique connue, la caractérisation biochimique de ces deux protéines s'est limitée à des comportements migratoires sur gel et à des protéolyses limitées. Les protéines migrent de façon identique en condition native et possèdent le même point isoélectrique. La très haute stabilité des γ cristallines, vis à vis des protéases est conservée après remplacement des hydrogènes. Les protéines ont été aussi caractérisées d'un point de vue spectral, en infrarouge et en dichroïsme circulaire, et ne présentent pas de grande différence entre elles, même si la protéine deutériée semble avoir gagné une très légère stabilité. Les protéines hydrogénées et deutériées sont donc identiques d'un point de vue biochimique et spectral.

Par la suite, les deux protéines ont été cristallisées dans des conditions similaires et des collectes de diffraction aux rayons X ont été effectuées sur les protéines hydrogénées et deutériées cristallisées en tampon hydrogéné et en tampon deutérié. La structure de ces protéines a été résolue entre 1,36 Å et 1,72 Å. Les différents modèles obtenus ont été comparés et ne présentaient pas de différence majeure. La différence la plus remarquable a été relevée au niveau du facteur thermique B qui subit l'influence stabilisante du deutérium.

Avec une meilleure résolution que celle de la protéine déjà publiée à 2,3 Å, les structures résolues ont révélé les grandes caractéristiques des γ cristallines : les quatre motifs en clé grecque, la stabilisation interdomaine, les contacts cristallins, le tripeptide RGD et la partie C-terminale flexible. Les éléments les plus remarquables sont certainement l'existence de la double conformation de la boucle entre les motifs c et d du motif 3, qui maintient la

présence d'un cœur hydrophobe fortement encombré, caractéristique des γ cristallines dont la température de transition est élevée.

Un autre élément important est la solvation de la γ E cristalline qui est relativement faible en comparaison de la γ B cristalline de bœuf résolue par Kumaraswamy et al. (1996). Mais ceci reste en accord avec les données de la structure à 2,3 Å, et avec le faible contenu en eau des cellules situées au centre du cristallin. De ce fait, seulement 145 molécules d'eau ont été localisées à 1,36 Å et interagissant principalement avec la protéine.

Cette étude cristallographique aux rayons X est une bonne base pour la cristallographie aux neutrons et pourra certainement être bénéfique pour la prochaine analyse par diffraction des neutrons, du solvant et des différentes interactions que fait la protéine. Malheureusement, aucun essai en cristallographie des neutrons n'a pu être réalisé par un manque de cristaux de volume important. Des cristaux dont la plus grande dimension dépassait le millimètre ont été couramment obtenus, mais l'obtention de la troisième dimension d'une grandeur utilisable est restée le problème majeur de cette thèse.

Hormis la cristallogénèse contrôlée par variation précise de température, d'autres possibilités d'évolution de ce sujet sont possibles. Etant donné le grand nombre de γ cristallines existantes, il est possible de changer légèrement le système protéique. Ceci a été entrepris au cours de ma thèse, avec la protéine γ C cristalline de *Rattus norvegicus*, dont la structure n'est pas encore connue, avec la protéine γ D cristalline humaine, résolue à 1,25 Å (Basak et al. 2003), et avec la γ B cristalline de bœuf, qui a fourni dans le passé des cristaux tridimensionnels (0,5 x 0,5 x 1 mm) (Najmudin et al. 1993; Kumaraswamy 1996). En ce qui concerne la γ C cristalline, la protéine a été produite sous forme hydrogénée et deutériée et des formes cristallines hydrogénées ont été obtenues, mais je n'ai pas réussi à améliorer les conditions. Pour le projet concernant γ D cristalline humaine, j'ai produit et purifié les protéines hydrogénées et deutériées. Les études préliminaires de cristallisation n'ont abouti qu'à des cristaux en forme de plaques. Enfin, je n'ai réussi à obtenir le clone codant pour la protéine γ B cristalline de bœuf qu'au mois de septembre 2004. La protéine a été produite sous forme hydrogénée et reste la meilleure voie exploitable pour une étude de diffraction de neutrons étant les dimensions des cristaux qu'il possible d'obtenir.

Les techniques de substitution isotopique développées au cours de cette thèse, peuvent aussi bien s'appliquer à d'autres domaines comme la diffusion des neutrons aux

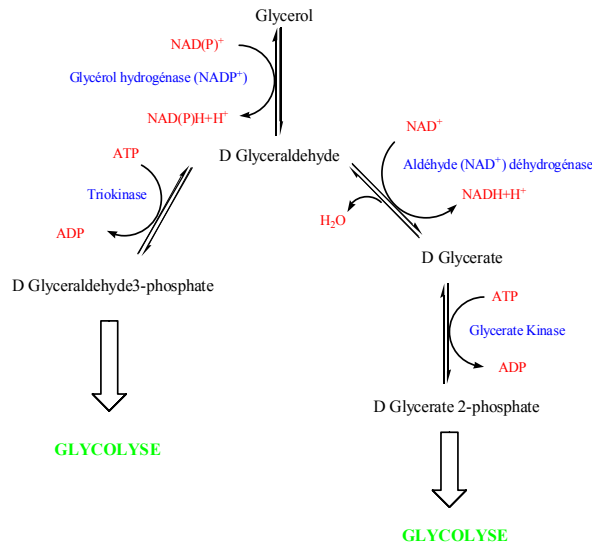
petits angles mais les travaux décrits précédemment restent une approche essentielle pour faciliter et développer la diffraction neutronique qui demeure une technique idéale pour toute étude sur les interactions protéine-protéine, protéine-solvant et solvant-solvant.

Bibliographie :

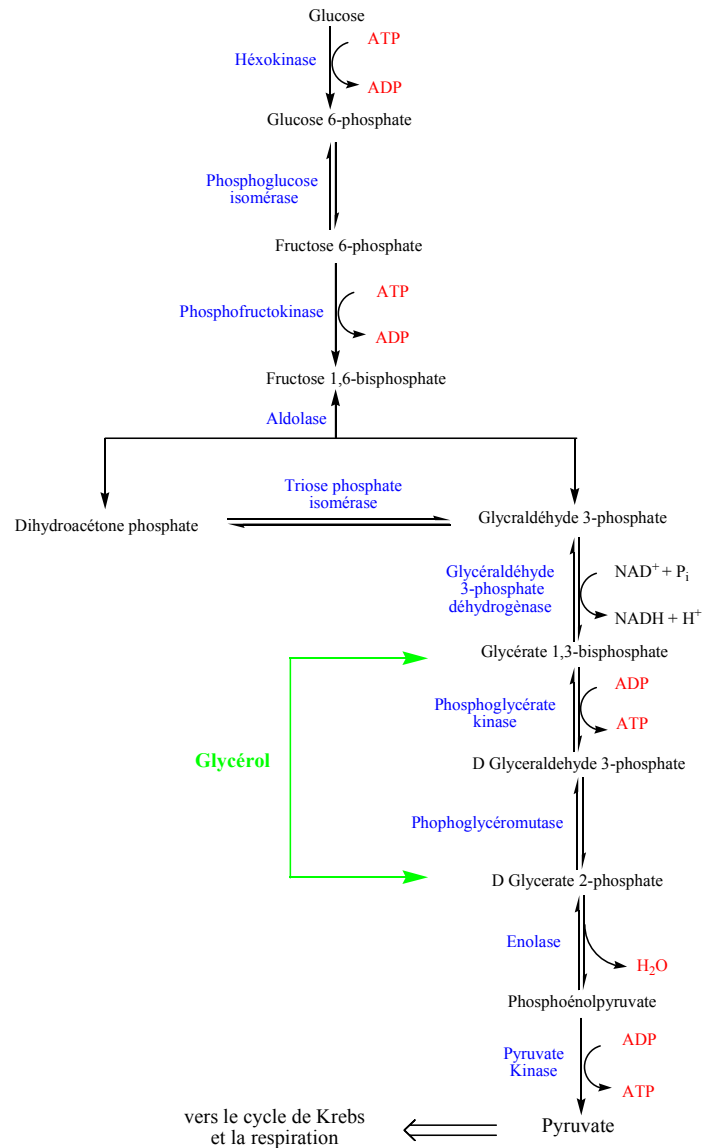
- Basak, A., Bateman, O., Slingsby, C., Pande, A., Asherie, N., Ogun, O., Benedek, G.B., and Pande, J. 2003. High-resolution X-ray crystal structures of human gammaD crystallin (1.25 Å) and the R58H mutant (1.15 Å) associated with aculeiform cataract. *J Mol Biol* **328**: 1137-1147.
- Bergfors, T. 1999. *Protein crystallization : techniques, strategies and tips A laboratory Manual*. International University Line, La Jolla.
- Budayova-Spano, M. 2005. En préparation.
- Ducruix, A., and Giegé, R. 1992. *Crystallization of nucleic acids and proteins : A practical approach*. Oxford university Press, New York.
- Garcia-Ruiz, J., and Moreno, A. 1997. Growth kinetics of protein crystals in the gel acupuncture technique. *Journal of Crystal Growth* **178**: 393-401.
- Garcia-Ruiz, J., Moreno, A., Otalora, F., and Rondon, D. 1998. Teaching Protein Crystallization by the Gel Acupuncture method. *Journal of Chemical Education* **75**: 442-446.
- Garcia-Ruiz, J.M. 1994. Investigations on protein crystal growth by the gel acupuncture method. *Acta Crystallogr D Biol Crystallogr* **50**: 484-490.
- Kumaraswamy, V.S., Lindley, P. F., Slingsby, C., Glover, I. D. 1996. An eye lens protein-water structure: 1.2 angstrom resolution structure of gamma B-crystallin at 150K. *Acta Crystallogr D Biol Crystallogr* **52**: 611.
- Najmudin, S., Nalini, V., Driessen, H., Slingsby, C., Blundell, T., S., M.D., and Lindley, P.F. 1993. Structure of the bovine eye lens protein gammaB(gammaII)-crystallin at 1.47 Å. *Acta Crystallogr D Biol Crystallogr* **49**: 223-233.
- Santarsiero, B., Yegian, D., Lee, C., Spraggon, G., Gu, J., Scheibe, D., Uber, D., Cornell, E., Nordmeyer, R., Kolbe, W., et al. 2002. An approach to rapid protein crystallization using nanodroplets. *Journal of Applied Crystallography*: 278-281.

Annexes A : Intégration du glycérol dans la glycolyse

A.1 Modifications du glycérol



A.2 Intégration des dérivés du glycérol dans la glycolyse

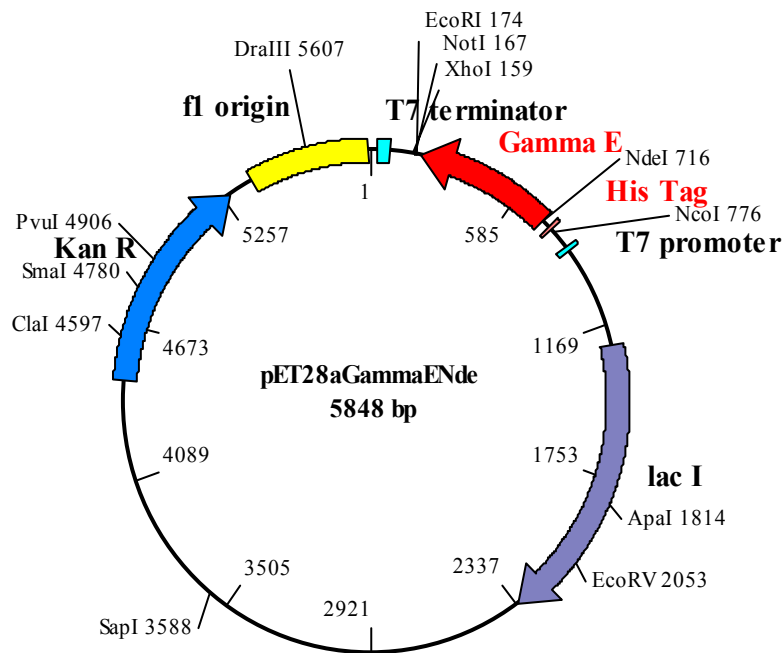


Annexe B : Constructions plasmidiques et analyse de séquence

B.1 Plasmides contenant la γ E cristalline :

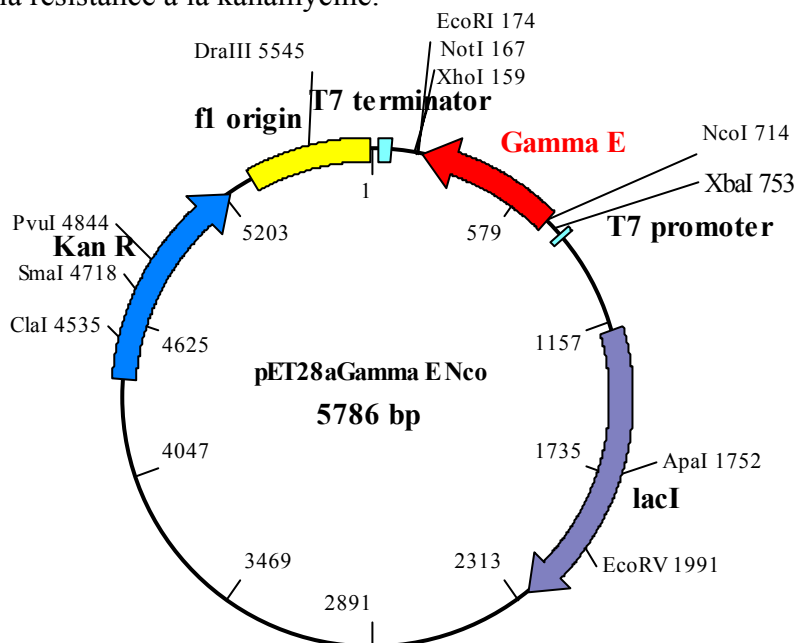
B.1.1 Le pEt28 γ ENde

Ce plasmide sert à exprimer la γ E cristalline avec une queue polyhistidines et il possède le gène codant pour la résistance à la kanamycine.



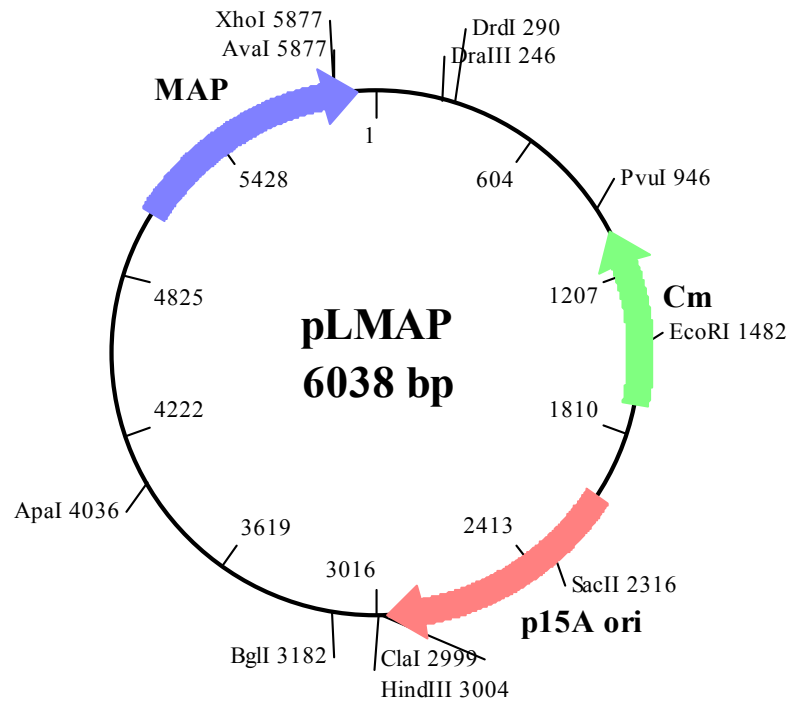
B.1.2 Le pEt28 γ ENco

Ce plasmide sert à exprimer la γ E cristalline sans queue polyhistidines et il possède le gène codant pour la résistance à la kanamycine.



B.2 Plasmide contenant la MAP (Méthionine Amino Peptidase)

Ce plasmide permet de surexprimer la MAP et il possède le gène codant pour la résistance au chloramphénicol (Cm).



B.3 γ E cristalline de rat et codons rares

B.3.1 Séquence de l'ADN complémentaire de la γ E cristalline de rat

```

1- ATG GGG AAG ATC ACC TTC TAT GAG GAC CGC GGC TTC CAG GGC CGC CAC TAT GAG
TGC AGC ACA GAC CAC TCC AAC CTG CAG CCC TAC TTC AGC CGC TGC AAC TCT GTG CGC
GTG GAC AGT GGC TGC TGG ATG CTC TAT GAG CAG CCC AAC TTC ACA GGC TGC CAG TAT
TTC CTT CGT CGC GGG GAC TAC CCT GAC TAC CAG CAG TGG ATG GGT TTC AGC GAC TCT
GTC CGC TCC TGC CGC CTC ATC CCC CAC TCC AGC TCT CAC AGA ATC AGG ATC TAC GAG
CGA GAG GAC TAC AGA GGC CAG ATG GTG GAG ATC ACA GAC GAC TGC CCC CAC CTG CAG
GAC CGC TTC CAC TTC AGT GAC TTC CAC TCT TTC CAC GTG ATG GAG GGC TAC TGG GTC
CTC TAT GAG ATG CCC AAC TAC CCG GGG CCG CAG TAC CTG CTG AGG CCT GGA GAA TAC
AGG CGC TAC CAC GAC TGG GGC GCC ATG AAT GCC AGG GTA GGC TCT CTG AGG AGA ATC
ATG GAT TTC TAT -523

```


B.3.2 Tableau des codons rares présents dans la γ E cristalline de rat

<i>Acide Aminé</i>	Codon rare	Fréquence dans l'ADNc de la γ E (en %)	Taille des fragments éventuellement générés
Arginine	CGA	0,5 %	8,5 KDa
	CGG	1 %	17,0 et 17,1 KDa
	AGG	2,5 %	11,0 et 18,0 ; 18,2 ; 20,0 et 20,5 KDa
	AGA	1,5%	10,0 et 12,0 KDa
Glycine	GGA	0,5 %	18,1 KDa
	GGG	1,5 %	- ; 7,2 et 17,0 KDa
Proline	CCC	2,5 %	3,2 ; 5,7 ; 9,8 ; 13 ; 17 KDa

Annexes C: Milieux de cultures

C.1. Milieux riches

Luria Bertani (LB)

Composition du milieu

<i>Composants</i>	<i>Quantités par litre</i>
Bacto-Tryptone	10 g
Extrait de bactéries et de levures	5 g
NaCl	10 g
NaOH	Jusqu'à pH = 7,0
+/- Agar	15 g
+/- Antibiotique(s)	

La solution est stérilisée par autoclave. L'agar fondu est éventuellement remis en suspension par agitation douce. Si nécessaire, un ou plusieurs antibiotiques sont ajoutés quand la température du milieu est aux environs de 50°C. Approximativement 25 ml de cette solution, sont coulés dans une boîte de Pétri stérile.

Milieu SOC

<i>Composants</i>	<i>Quantités pour 1l</i>
Bacto-Tryptone	20 g
Extrait de bactéries et de levures	5 g
NaCl	0,5 g
KCl 1M	2,5 ml
NaOH	Jusqu'à pH = 7,0

La solution est stérilisée par autoclave et 10 ml de MgCl₂ et 20 ml de Glucose 1M sont ajoutés stérilement.

C.2 Milieu Enfors

Solution minérale

Composants	Quantités par litre
CaCl ₂ .2H ₂ O	0,5 g
FeCl ₃ .6H ₂ O	16,7 g
ZnSO ₄ .7H ₂ O	0,18 g
CuSO ₄ .5H ₂ O	0,16 g
MnSO ₄ .4H ₂ O	0,15 g
CoCl ₂ .6H ₂ O	0,18 g
EDTA	20,1 g
H ₂ O	qsp 1l

Pour éviter toute formation de complexe ou de précipité, les différents composants sont ajoutés successivement et dans cet ordre précis, en attendant la dissolution complète de chaque élément. La solution est stérilisée par filtration à travers une membrane 0,2 µm (Millipore).

Pour préparer une solution minérale deutériée, les composants sont dissouts dans une petite quantité de D₂O déionisée. Cette solution est ensuite évaporée par évaporation « flash » au rotovapor. La pâte obtenue est diluée dans une faible quantité de D₂O. Cette opération est répétée quatre fois pour remplacer tous les hydrogènes en deutériums. La solution est ensuite stérilisée par filtration.

Milieu Enfors

<i>Composants</i>	<i>Quantités par litre</i>
(NH ₄) ₂ SO ₄	6,86 g
KH ₂ PO ₄	1,56 g
Na ₂ HPO ₄ .2H ₂ O	6,48 g
(NH ₄) ₂ -H-Citrate	0,49 g
Solution MgSO ₄ 1M	1 ml
Sol. Minérale	1 ml
Glycérol 100%	5 g
+/- Antibiotique(s)	
H ₂ O	qsp 1l

Comme pour la solution minérale, les éléments sont ajoutés dans l'ordre de la liste pour éviter toute précipitation. La solution est stérilisée par filtration à travers une membrane 0,2 µm (Millipore). Le pH est vérifié mais l'ajout des différents composants devrait établir un pH aux alentours de 6,8-7.

Pour préparer du milieu deutérié, les éléments solides ((NH₄)₂SO₄, KH₂PO₄, Na₂HPO₄.2H₂O et (NH₄)₂-H-Citrate) subissent quatre évaporations « flash » et quatre dilutions dans du D₂O. 1 ml de MgSO₄ dissout dans de l'eau lourde est ajouté au milieu qui a été complété par ~ 900 ml de D₂O, ainsi que 1 ml de la solution minérale deutériée. Les antibiotiques nécessaires et 6,45 ml de glycérol deutérié sont ajoutés avant de filtrer sur une membrane 0,2 µm.

Pour toute conservation de solution deutériée, les bouteilles sont fermées hermétiquement.

Solution de « Feeding »

<i>Composants</i>	<i>Quantités pour 250 ml</i>
Glycérol (hydrogéné ou deutérié)	10 g
MgSO ₄ (hydrogéné ou deutérié)	0,25 ml
Sol. Minérale (hydrogénée ou deutériée)	ou 0,25
+ Antibiotique(s)	
H ₂ O (D ₂ O)	qsp 250 ml

C.3 Antibiotiques et IPTG

<i>Dénomination</i>	<i>Solution stock</i>	<i>Solution de travail</i>
Ampicilline	100 mg/ml dans H ₂ O (D ₂ O) conservée à -20°C	50 µg/ml
Kanamycine	30 mg/ml dans H ₂ O (D ₂ O) conservée à -20°C	30 µg/ml
Chloramphénicol	34 mg/ml dans le l'éthanol conservée à -20°C	50 µg/ml
isopropyl-β-D-thiogalactopyranoside, IPTG	2,38 g dans 10 ml H ₂ O (D ₂ O) conservée à -20°C	0,02-0,24 mg/ml

Les composés sont dissouts dans de l'eau légère ou lourde, ou de l'éthanol, puis la solution est filtrée sur une membrane 0,2 µm, puis gardée à -20°C.

Annexe D : Tampons de purification

D. 1 Purification de la protéine avec une queue poly histidines

Tampon de lyse

<i>Composants</i>	<i>Quantité pour 500 ml</i>	<i>Concentration Finale</i>
Tris/ HCl pH = 7,5	3,02 g	50 mM
Imidazole	10 ml, 1M stock	10 mM
Glycérol	7,8 ml	2 %

Juste avant l'utilisation, 1 ml du stock de 1,4 dithio-D,L-threitol 1M (DTT) est additionné aux 500ml. Une pastille de cocktail d'antiprotéases « complete EDTA free » est ajoutée à 50 ml de tampon de lyse qui sont ensuite agités jusqu'à dissolution complète.

1^{er} Tampon de lavage de la colonne de Nickel

Ce premier tampon est identique au tampon de lyse.

2nd Tampon de lavage de la colonne de Nickel

<i>Composants</i>	<i>Quantité pour 500 ml</i>	<i>Concentration Finale</i>
Tris/ HCl pH = 7,5	3,02 g	50 mM
Imidazole	20 ml, 1M stock	20 mM
Glycérol	7,8 ml	2 %

Tampon d'élution de la colonne

<i>Composants</i>	<i>Quantité pour 500 ml</i>	<i>Concentration Finale</i>
Tris/ HCl pH = 7,5	3,02 g	50 mM
Imidazole	8,51 g	250 mM
Glycérol	7,8 ml	2 %

Avant chaque utilisation, tous les tampons sont complétés par 2mM DTT final.

D.2 Purification de la protéine sans une queue poly histidines

Tampon de lyse

<i>Composants</i>	<i>Quantité pour 2l</i>	<i>Concentration Finale</i>
Hépes pH = 6.4	47,16 g	50 mM
NaCl	0,58 g	5 mM
DTT	1,04 g	5 mM
+ 1/2 pastille « complete EDTA free »		

1^{er} Tampon de lavage de la colonne d'affinité SP

Ce premier tampon est identique au tampon de lyse.

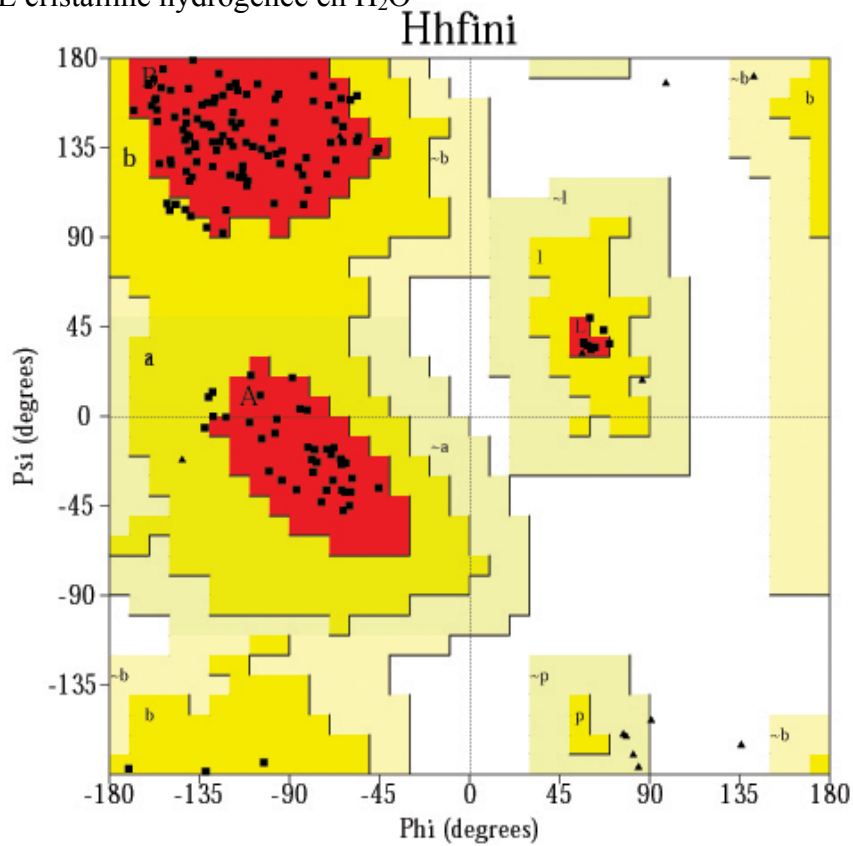
Tampon d'élution de la colonne d'affinité SP

<i>Composants</i>	<i>Quantité pour 0,5 ml</i>	<i>Concentration Finale</i>
Hépes pH = 6.4	11,79 g	50 mM
NaCl	5,26 g	300 mM
DTT	0,52 g	10 mM

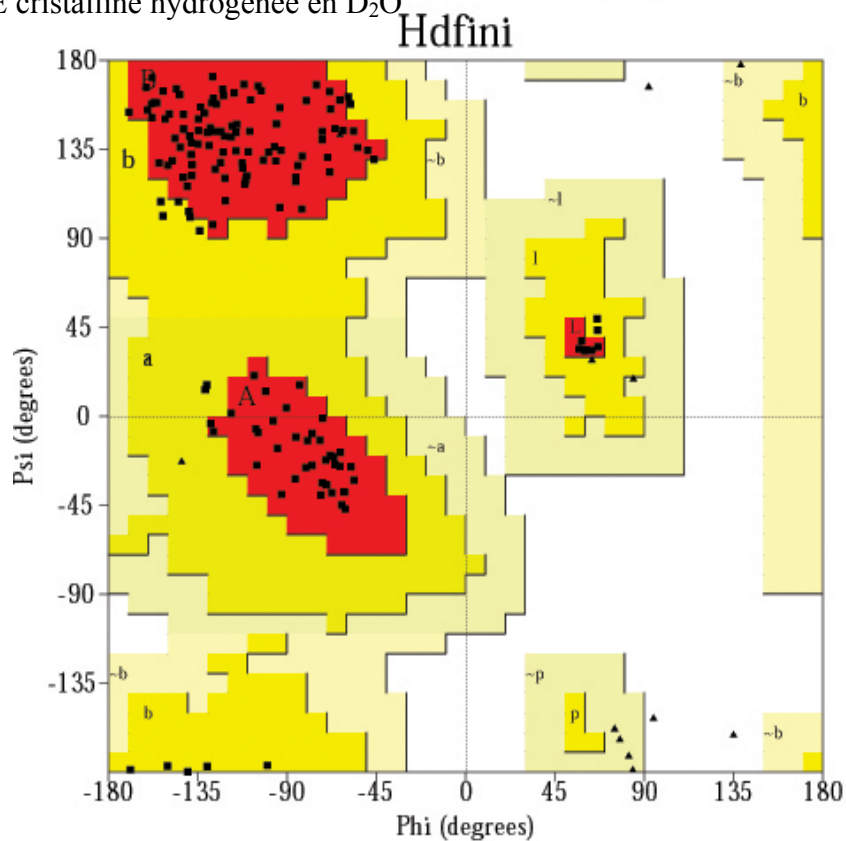
Annexe E : Structure des γE cristallines

E.1 Diagrammes de Ramachandran

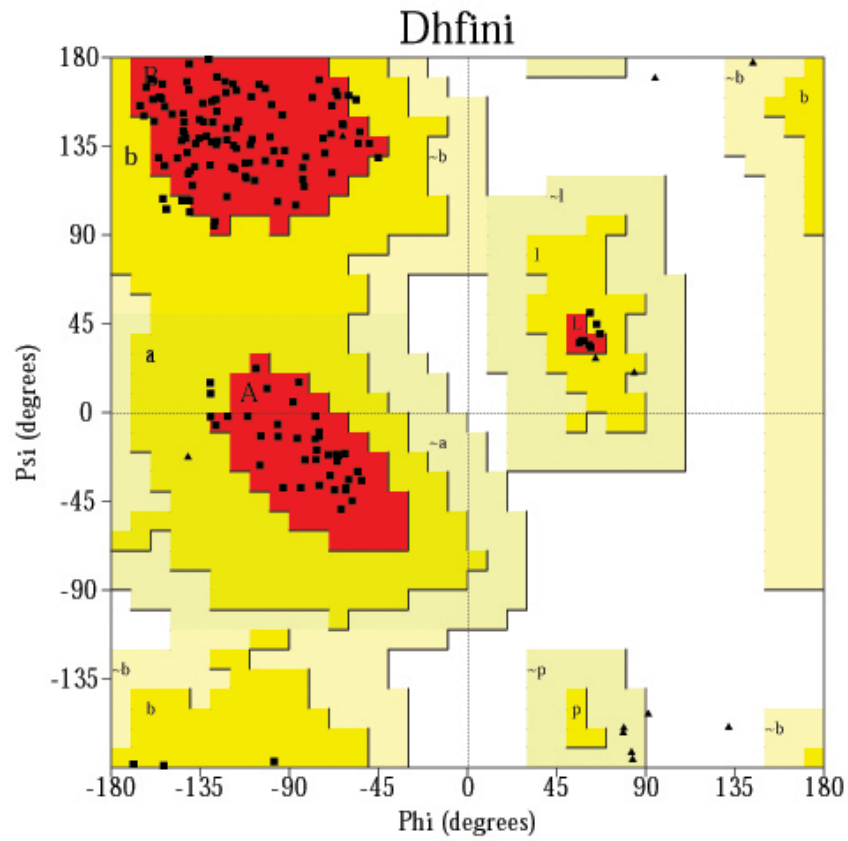
E.1.1 γE cristalline hydrogénée en H₂O



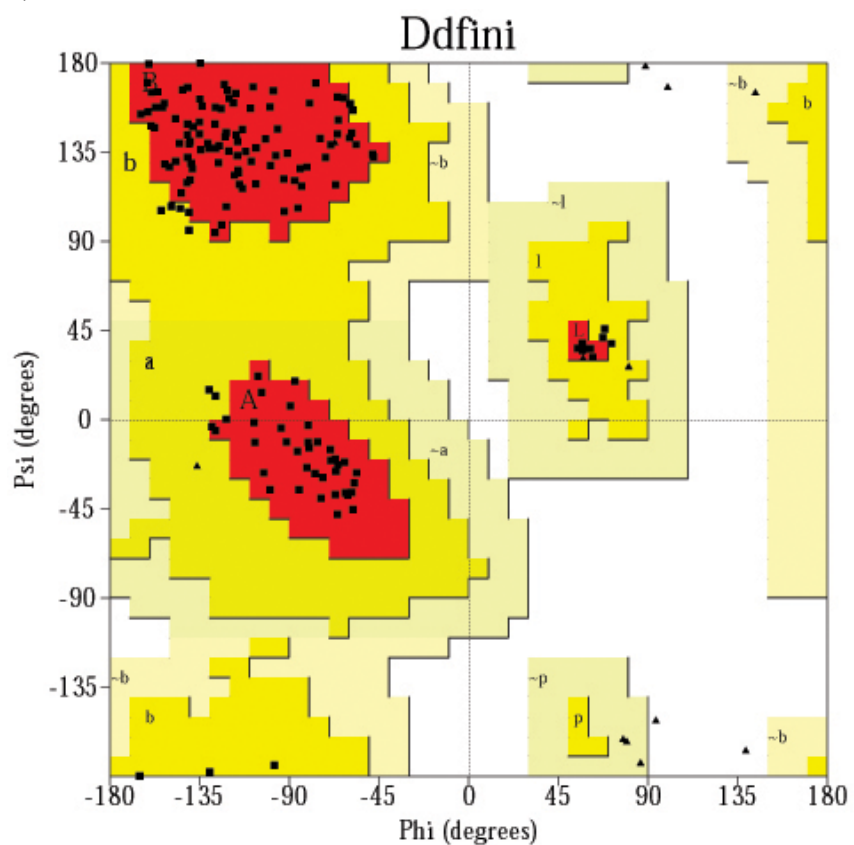
E.1.2 γE cristalline hydrogénée en D₂O



E.1.3 γE cristalline deutériée en H₂O



E.1.4 γE cristalline deutériée en D₂O



E.2 Molécules d'eau de la γE cristalline deutériée en H₂O

Pour définir cette liste de molécules d'eau entourant la γE cristalline deutériée en H₂O, plusieurs critères ont été considérés :

- des critères géométriques comme la distance d'interaction avec la protéine ou le solvant ou l'orientation des liaisons, comme définis dans le chapitre 6. Ces molécules d'eau forment des liaisons hydrogène qui sont principalement dues à une interaction électrostatique, qui est une interaction attractive à longue portée inversement proportionnelle à la distance interatomique. Les distances des liaisons hydrogène sont comprises entre 2,4 et 3,5 Å. Une molécule d'eau complète, c'est à dire avec le maximum d'interactions possibles, doit former un tétraèdre.
- des critères visuels. Toute molécule d'eau qui ne présentait pas de densité électronique sphérique est soit partagée en deux molécules d'eau, soit non considérée.
- des critères d'agitation thermiques. Toute molécule d'eau qui présente une agitation thermique très élevée ($B > 60\text{Å}^2$), après affinement par *ShelX*, n'était pas considérée, tout comme celles qui se trouvaient à proximité de la partie C terminale, flexible, de la protéine.

Dans le tableau suivant, sont représentés les atomes d'oxygène des molécules d'eau avec leur facteur thermique *B* et les liaisons établies avec la protéine et le solvant.

<i>Molécules d'eau</i>	<i>Facteur B en Å²</i>	<i>Contacts</i>			<i>Distance en Å</i>	<i>Remarques</i>
		<i>Résidu</i>	<i>Numéro</i>	<i>Atome</i>		
W3001	9,02	Gln	1142	N	3,01	Réseau B
		Gln	1054	O	2,70	
		Wat	3003	O	2,81	
		Wat	3016	O	2,76	
W3002	8,56	Trp	1068	NE1	2,89	Molécule la plus enfouie
		Ser	1072	O	2,82	
		Ser	1074	O	3,03	
W3003	9,66	Gln	1142	NE2	2,94	Réseau B
	9,66	Gly	1140	O	2,81	
		Met	1069	O	2,75	
		Wat	3001	O	2,81	
W3004	11,11	Gln	1054	N	2,99	Réseau B
		Gln	1142	O	2,7	
		Wat	3016	O	2,69	
		Wat	3068	O	2,91	

<i>Molécules d'eau</i>	<i>Facteur B en Å²</i>	<i>Contacts</i>			<i>Distance en Å</i>	<i>Remarques</i>
		<i>Résidu</i>	<i>Numéro</i>	<i>Atome</i>		
W3005	11,48	Gln	1054	NE2	3,00	
		Leu	1144	O	2,73	
		Trp	1156	O	2,77	
		Wat	3139	O	2,94	
W3006	8,89	Leu	1080	O	2,8	Réseau A
		Asp	1021	OD1	3,02	
		Asp	1021	OD2	3,07	
		Wat	3026	O	2,85	
W3007	14,92	Gln	1026	O	2,8	Réseau A
		Sym Wat	3032	O	2,91	
		Wat	3094	O	2,79	
W3008	9,78	Asp	1073	OD1	2,97	Réseau A
		Ser	1074	OG	2,92	
		Wat	3018	O	3,09	
		Wat	3097	O	2,73	
		Sym Wat	3052	O	2,91	
W3009	9,24	Ser	1074	O	2,83	Réseau B
		Ser	1072	OG	2,88	
		Tyr	1055	OH	2,67	
		Wat	3068	O	2,78	
W3010	13,98	Gly	1001	N	2,79	
		Gly	1001	O	3,11	
		Ser	1019	O	2,79	
		Wat	3116	O	2,88	
W3011	16,25	Sym Glu	1095	N	2,87	
		Sym His	1121	ND1	3,02	
		Asp	1113	OD1	2,73	
		Sym Wat	4174	O	2,83	
W3012	9,65	Arg	1076	NH1	2,8	Réseau C
		Sym Wat	3028	O	2,89	
		Glu	1046	OE2	2,68	
W3013	15,49	Act	2378	OT	2,78	
		Act	2378	O	3,23	
		Cys	1041	N	2,82	
		Ile	1081	O	2,61	
W3014	12,42	Arg	1031	NH1	2,85	Réseau A
		Ser	1030	OG	2,74	
		Wat	3097	O	2,6	
W3015	11,21	Arg	1079	NE	2,8	
		Asp	1021	OD2	3,27	
W3016	15,93	Arg	1141	NH1	2,95	Réseau B
		Wat	3001	O	2,76	
		Wat	3004	O	2,69	
W3017	19,11	Gln	1026	OE1	2,86	Réseau C
		Wat	3023	O	2,78	
		Wat	3096	O	2,71	
W3018	12,31	Val	1075	N	2,96	Réseau A
		Wat	3021	O	2,96	
		Wat	3008	O	3,09	
		Wat	3030	O	3,12	

<i>Molécules d'eau</i>	<i>Facteur B en Å²</i>	<i>Contacts</i>			<i>Distance en Å</i>	<i>Remarques</i>
		<i>Résidu</i>	<i>Numéro</i>	<i>Atome</i>		
W3019	13,5	Arg	1014	N	2,86	
		Glu	1007	OE1	2,71	
		Wat	3122	O	2,82	
		Wat	3088	O	2,69	
W3020	11,49	Gln	1054	OE1	3,01	
		Tyr	1045	OH	2,66	
		Wat	3084	O	2,88	
W3021	10,99	Phe	1029	O	2,88	Réseau A
		Val	1075	O	2,84	
		Wat	3018	O	2,96	
		Wat	3030	O	2,9	
W3022	22,08	Ser	1019	N	2,8	Réseau E
		Thr	1020	OG1	3,22	
		Wat	3142	O	2,75	
		Wat	3058	O	2,62	
W3023	13,86	Wat	3017	O	2,78	Réseau C
		Sym Wat	3028	O	2,91	
		Wat	3045	O	2,84	
W3024	11,14	Asp	1073	N	2,91	
		Wat	3098	O	2,96	
W3025	16,71	Tyr	1065	N	2,98	
		Asn	1033	O	2,78	
		Asn	1033	OD1	2,73	
		Wat	3134	O	2,72	
W3026	13,07	Arg	1146	NH1	2,91	Réseau A
		Wat	3052	O	2,77	
		Wat	3006	O	2,85	
W3027	21,47	Glu	1017	N	2,94	
		Sym Gln	1100	NE2	3,29	
		Sym Gln	1100	OE1	2,83	
		Wat	3082	O	2,73	
W3028	10,89	Leu	1080	O	2,8	Réseau C
		Sym Wat	3023	O	2,91	
		Sym Wat	3012	O	2,89	
W3029	13,78	Trp	1156	NE1	2,95	Molécule la plus enfouie
		Arg	1162	O	2,89	
		Wat	3051	O	2,78	
		Asn	1160	O	2,82	
W3030	16,41	Arg	1076	NE	3,45	Réseau A Molécule centrale du réseau
		Ser	1030	O	3,33	
		Wat	3021	O	2,9	
		Wat	3018	O	3,12	
		Sym Wat	3052	O	2,86	
		Sym Wat	3032	O	2,94	
W3031	24,68	His	1124	ND1	2,79	
		Arg	1090	NH1	2,96	
		Tyr	1092	OH	2,74	
W3032	14,69	Tyr	1129	OH	2,72	Réseau A
		Sym Wat	3007	O	2,91	
		Sym Wat	3030	O	2,94	
		Sym War	3056	O	2,86	

<i>Molécules d'eau</i>	<i>Facteur B en Å²</i>	<i>Contacts</i>			<i>Distance en Å</i>	<i>Remarques</i>
		<i>Résidu</i>	<i>Numéro</i>	<i>Atome</i>		
W3033	21,23	Met	1101	N	3,08	
		Gln	1100	OE1	2,68	
W3034	19,98	Gly	1128	N	2,9	
		Wat	3140	O	3,03	
W3035	21,89	Phe	1056	O	2,79	
		Tyr	1062	OH	2,74	
		Wat	3110	O	3,01	
W3036	23,29	Thr	1105	N	2,91	
		Asp	1106	OD1	2,95	
W3037	21,56	Gln	1112	N	2,87	
		Gly	1164	O	3,24	
W3038	17,9	His	1087	ND1	2,75	
		Arg	1168	NH1	2,97	
		Wat	3055	O	2,73	
W3039	14,9	Pro	1147	O	2,76	
		Tyr	1150	OH	2,89	
		Wat	3064	O	2,72	
		Wat	3072	O	2,64	
W3040	23,53	Asn	1024	N	2,9	
		Ser	1023	N	2,97	
		Ser	1023	OG	3,24	
		Wat	3063	O	2,74	
W3041	18,88	Gln	1142	OE1	3,02	
		Tyr	1133	OH	2,63	
		Wat	3043	O	2,96	
		Wat	3061	O	3,03	
W3042	17,77	Thr	1004	N	3,22	
		Arg	1036	O	2,78	
		Asp	1038	OD2	2,55	
		Thr	1004	OG1	2,77	
W3043	23	Arg	1167	NH2	2,97	
		Tyr	1138	O	2,83	
		Wat	3041	O	2,96	
W3044	19,96	Glu	1017	OE1	2,6	
		Thr	1004	OG1	2,71	
		Sym Wat	3047	O	2,59	
W3045	16,49	Gln	1047	NE2	2,75	
		Wat	3023	O	2,84	
		Wat	3090	O	3,07	
W3046	28,7	Glu	1127	OE2	2,65	
		Wat	3126	O	2,95	
W3047	23,54	Gly	1099	N	2,93	
		Arg	1098	NE	2,57	
		Sym Wat	3044	O	2,59	
W3048	25,5	Glu	1134	O	2,75	Réseau D
W3049	19,36	Asp	1008	OD1	2,76	
		Wat	3060	O	2,67	
W3050	18,01	Tyr	1153	N	2,88	
		His	1121	O	2,82	

<i>Molécules d'eau</i>	<i>Facteur B en Å²</i>	<i>Contacts</i>			<i>Distance en Å</i>	<i>Remarques</i>
		<i>Résidu</i>	<i>Numéro</i>	<i>Atome</i>		
W3051	18,2	Arg	1162	O	2,79	
		Tyr	1143	OH	2,62	
		Wat	3073	O	2,57	
W3052	14,79	Tyr	1129	OH	2,75	Réseau A
		Wat	3026	O	2,77	
		Sym Wat	3008	O	2,91	
		Sym Wat	3030	O	2,86	
W3053	24,59	Glu	1103	N	2,99	
W3054	22,35	Tyr	1065	OH	2,71	
W3055	16,84	His	1087	O	3,11	
		Thr	1105	O	2,96	
		Wat	3038	O	2,73	
		Sym Wat	4171	O	2,76	
W3056	26,83	Gln	1026	OE1	2,73	Réseau A
		Sym Wat	3032	O	2,86	
W3057	21,76	Thr	1105	OG1	3,09	
		Sym Wat	4171	O	2,69	
W3058	30,31	Glu	1017	O	2,74	Réseau E
		Tyr	1016	OH	2,74	
		Wat	3022	O	2,62	
W3059	44,53	Asn	1049	ND2	2,78	
W3060	21,17	Arg	1009	N	2,94	Réseau D
		Wat	3089	O	2,95	
		Wat	3049	O	2,67	
		Wat	3103	O	2,83	
W3061	18,42	Gln	1142	NE2	3,05	
		Phe	1056	O	2,67	
		Wat	3110	O	3,12	
		Wat	3041	O	3,03	
W3062	22,14	Arg	1009	O	2,72	Réseau D
		Wat	3089	O	2,73	
W3063	26,45	Asn	1024	O	2,69	
		Wat	3040	O	2,74	
W3064	21,59	Act	2380	OT	2,54	
		Arg	1146	NE	2,97	
		Arg	1146	NH2	3,02	
		Wat	3039	O	2,72	
W3065	28,25	His	1121	ND1	3,21	
		Glu	1093	O	2,81	
		Wat	4174	O	2,59	
W3066	44,53	Asn	1024	ND2	3,04	
W3067	20,21	Gly	1099	O	2,86	
		Tyr	1097	O	2,93	
		Sym Wat	3143	O	2,55	
W3068	15,97	Gln	1054	NE2	3,08	Réseau B
		Gly	1052	O	3	
		Gly	1157	O	2,81	
		Wat	3009	O	2,78	
		Wat	3004	O	2,91	
W3069	20,01	His	1015	O	2,89	
		Tyr	1028	O	2,73	

<i>Molécules d'eau</i>	<i>Facteur B en Å²</i>	<i>Contacts</i>			<i>Distance en Å</i>	<i>Remarques</i>
		<i>Résidu</i>	<i>Numéro</i>	<i>Atome</i>		
W3070	18,71	His	1022	ND1	3,23	Réseau E
		Asp	1021	O	2,76	
		Thr	1020	OG1	2,49	
		Wat	3142	O	2,85	
W3071	20,49	Arg	1059	NH1	2,97	
		Sym Pro	1136	O	2,78	
		Wat	4175	O	2,7	
W3072	31,72	Glu	1149	O	2,83	
		Wat	3039	O	2,64	
W3073	36,97	Asn	1160	ND2	3,21	
		Asn	1160	OD1	2,56	
		Wat	3051	O	2,57	
W3074	26,5	Asn	1024	ND2	2,9	
W3075	24,46	Met	1159	N	2,87	Réseau E
		Tyr	1143	OH	2,86	
		Wat	3078	O	3,08	
W3076	28,2	Val	1163	N	2,98	
		Val	1163	O	3,17	
		Wat	3083	O	2,97	
W3077	31,32	Pro	1082	O	3,21	
W3078	29,46	Wat	3075	O	3,08	Réseau E
		Sym Wat	3078	O	2,89	
W3079	29,39	Asp	1061	N	2,87	
		Sym Ser	1085	OG	3,19	
W3080	27,6	Arg	1031	NH2	3,11	
W3081	33,01	Sym His	1154	NE2	3,19	
		Glu	1093	OE2	2,58	
		Sym Tyr	1153	OH	2,88	
W3082	24,27	Sym Gln	1100	NE2	3,07	
		Glu	1017	O	2,71	
		Wat	3027	O	2,73	
W3083	25,26	Phe	1120	N	2,91	
		Asp	1119	OD1	2,93	
		Wat	3076	O	2,97	
W3084	25,21	Arg	1079	NH1	3,11	
		Phe	1050	O	2,82	
		Wat	3020	O	2,88	
		Wat	3099	O	3,14	
W3085	31,15	Sym Asp	1061	O	3,29	
		Ser	1085	OG	3,27	
		Sym Wat	3087	O	3,01	
		Sym Wat	3104	O	3,03	
W3086	29,57	Arg	1088	NH2	2,92	
		Glu	1127	OE1	2,66	
		Sym Wat	4171	O	2,82	
W3087	44,67	Arg	1058	O	3,12	
		Arg	1059	O	2,9	
		Tyr	1062	OH	2,77	
		Sym Wat	3085	O	3,01	

<i>Molécules d'eau</i>	<i>Facteur B en Å²</i>	<i>Contacts</i>			<i>Distance en Å</i>	<i>Remarques</i>
		<i>Résidu</i>	<i>Numéro</i>	<i>Atome</i>		
W3088	19,06	Arg	1031	NH2	3,23	
		Asp	1008	OD2	2,51	
		Wat	3019	O	2,69	
W3089	30,9	Wat	3062	O	2,73	Réseau D
		Wat	3060	O	2,95	
W3090	39,17	Thr	1051	OG1	2,69	Réseau C
		Wat	3045	O	3,07	
		Wat	3100	O	2,71	
W3091	44,81	His	1015	N	3,05	
		His	1015	ND1	3,15	
W3092	46,06	Pro	1027	O	2,97	
W3093	32,77	Arg	1168	O	2,87	
		Asp	1107	OD1	2,91	
		Asp	1107	OD2	2,96	
W3094	20,03	Arg	1014	NH1	2,98	Réseau A
		Arg	1014	NH1	2,97	
		Gln	1026	O	3	
		Pro	1027	O	3,23	
		Wat	3007	O	2,79	
W3095	36,09	Tyr	1092	OH	2,61	
		Sym Wat	3123	O	3,25	
W3096	49,01	Wat	3017	O	2,71	Réseau C
W3097	23,43	Ser	1030	O	2,97	Réseau A
		Wat	3014	O	2,6	
		Wat	3008	O	2,73	
W3098	24,52	Asn	1033	OD1	2,97	
		Asp	1073	OD2	3,02	
		Wat	3024	O	2,96	
W3099	32,64	Arg	1079	NH2	2,83	
		Wat	3084	O	3,14	
W3100	52,5	Asn	1049	OD1	2,95	Réseau C
		Wat	3090	O	2,71	
W3101	32,03	Tyr	1138	OH	2,68	
W3102	28,34	Ser	1023	OG	2,75	
W3103	29,51	Wat	3134	O	2,77	Réseau D
		Wat	3060	O	2,83	
W3104	36,56	Asp	1061	O	2,39	
		Sym Ser	1085	O	2,66	
		Gln	1067	OE1	3,22	
		Sym Wat	3085	O	3,03	
W3105	23,75	Ala	1161	O	2,81	
		Phe	1120	O	2,6	
W3106	30,88	Gly	1013	N	3,04	
		Gln	1012	OE1	3,17	
W3107	37,05	Asp	1008	OD2	2,77	
W3108	32,88	Gln	1066	NE2	3,1	Réseau D
		Wat	3134	O	2,86	
		Wat	3137	O	3,29	
W3109	39,07	Gln	1054	NE2	3,25	
		Gly	1052	O	2,78	

<i>Molécules d'eau</i>	<i>Facteur B en Å²</i>	<i>Contacts</i>			<i>Distance en Å</i>	<i>Remarques</i>
		<i>Résidu</i>	<i>Numéro</i>	<i>Atome</i>		
W3110	33,49	Gln	1067	O	2,86	
		Wat	3035	O	3,01	
		Wat	3061	O	3,12	
W3111	20,03	Pro	1082	O	2,82	
		Ser	1084	OG	2,43	
W3112	43,03	Glu	1103	OE1	3,28	
		Glu	1103	OE2	2,67	
		Wat	3126	O	2,77	
W3113	45,57	Asn	1049	O	2,69	Réseau D
		Sym Wat	3137	O	2,89	
W3114	23,06	Sym Wat	4171	O	2,86	
W3115	30,66	Arg	1059	N	2,91	
		Gly	1040	O	3,24	
W3116	58,26	Sym Wat	1010	O	2,88	
W3117	30,95	Sym Arg	1168	NH2	2,86	Réseau D
		Gln	1066	OE1	2,88	
		Wat	3133	O	3,18	
W3118	27	Sym Asn	1137	ND2	3,02	
		Asp	1038	O	2,73	
		Wat	3071	O	2,7	
W3119	25,33	Arg	1094	NH1	2,94	
		Sym Asp	1113	O	2,95	
		Wat	3065	O	2,59	
		Sym Wat	3011	O	2,83	
W3120	34,26	Gly	1001	O	3,04	
W3121	26,31	Glu	1134	OE1	2,65	
W3122	26,86	Wat	3019	O	2,82	
W3123	30,29	Gly	1013	O	2,93	
		Sym Wat	3095	O	3,25	
W3124	52,81	His	1124	NE2	2,78	
		Val	1125	O	3,22	
W3125	39,14	Gly	1001	N	2,93	
		Ser	1039	OG	2,76	
W3126	49,22	Wat	3112	O	2,77	
		Wat	3046	O	2,95	
W3127	30,39	Asp	1107	N	2,93	Réseau D
		Wat	3132	O	3,2	
		Sym Wat	3133	O	2,68	
W3128	17,15	Phe	1115	N	2,65	
		Gln	1112	O	2,38	
		Leu	1111	O	2,81	
		Phe	1117	O	2,53	
W3129	27,05	Thr	1105	OG1	2,69	
		Wat	3138	O	2,87	
		Wat	4148	O	2,71	
W3130	21,93	Sym Wat	3086	O	2,82	
		Sym Wat	4177	O	2,86	
		Sym Wat	3057	O	2,69	
		Sym Wat	3055	O	2,76	

<i>Molécules d'eau</i>	<i>Facteur B en Å²</i>	<i>Contacts</i>			<i>Distance en Å</i>	<i>Remarques</i>
		<i>Résidu</i>	<i>Numéro</i>	<i>Atome</i>		
W3131	43,79	Asn	1137	OD1	2,84	
W3132	48,83	Asp	1107	O	2,89	Réseau D
		Wat	3127	O	3,2	
W3133	29,17	Sym Arg	1168	NH2	3,18	Réseau D
		Asp	1064	OD2	2,9	
		Wat	3117	O	3,18	
		Wat	3134	O	3,12	
		Sym Wat	3127	O	2,68	
W3134	21,03	Asp	1064	OD2	2,68	Réseau D Molécule centrale du réseau
		Wat	3108	O	2,86	
		Wat	3025	O	2,72	
		Wat	3133	O	3,12	
		Wat	3103	O	2,77	
W3135	43,43	Gln	1112	NE2	3,27	
		Sym Arg	1098	NH1	3,09	
		Sym Asp	1096	OD2	2,77	
		Gln	1112	OE1	2,96	
W3136	55,94	Wat	3129	O	2,71	
		Wat	3138	O	2,58	
		Glu	1103	O	2,96	
W3137	47,04	Tyr	1065	OH	2,78	Réseau D
		Sym Wat	3113	O	2,89	
		Wat	3108	O	3,29	
W3138	27,92	Arg	1088	NH1	2,97	
		Wat	3129	O	2,87	
		Wat	4148	O	2,58	
W3139	36	Act	2380	O	3,29	
		Act	2380	OT	3,19	
		Wat	3005	O	2,94	
W3140	44,42	Met	1126	O	2,77	
		Wat	3034	O	3,03	
W3141	39,38	Asp	1155	O	2,68	
W3142	54,62	Thr	1020	OG1	2,62	
		Wat	3022	O	2,75	
		Sym Wat	3078	O	2,89	
		Wat	3070	O	2,85	
W3143	31,73	Tyr	1006	OH	2,62	
		Sym Wat	3067	O	2,55	
W3144	44,7	Ser	1085	N	3,12	
		Ser	1085	OG	2,6	

Annexe E : concentration des γ E cristallines dans le cristal

- Nombre x de maille dans un cristal :

$$x = \frac{Volume_{cristal}}{Volume_{maille}}$$

- Nombre de protéine dans une maille en $P2_1$: 4
- Nombre y de protéine dans un cristal avec une maille en $P2_1$:

$$y = x \times 4$$

- Masse d'une molécule de γ E cristalline hydrogénée de 21132 Da :

$$m = \frac{21132}{N}$$

où N est le nombre d'Avogadro, $N = 6,022 \cdot 10^{23} \text{ mole}^{-1}$

- Concentration C en protéine dans le cristal :

$$C = \frac{y \times m}{Volume_{cristal}} = \frac{x \times 4 \times 21132}{Volume_{cristal} \times N} = \frac{Volume_{cristal} \times 4 \times 21132 \times Volume_{maille}}{Volume_{cristal} \times N} = \frac{84528 \times Volume_{maille}}{N}$$

Applications numériques :

Sachant que le volume de la maille de $H\gamma E h_2O$ est de $1,65 \cdot 10^{-19} \text{ cm}^3$, la concentration C sera alors de :

$$C = 0,85 \text{ g/cm}^3 = 850 \text{ mg/ml}$$