



Université de Franche-Comté

École Doctorale Louis Pasteur

Thèse de Doctorat

Spécialité Chimie

présentée par

Catalina David

Mécanismes de réduction des ponts S-S dans les
protéines étudiés par spectroscopie Raman et
modélisation moléculaire

Thèse dirigée par Mironel Enescu

Laboratoire de Chimie Physique et Rayonnements - Alain Chambaudet - UMR CEA E4

Soutenue le 18 novembre 2009 devant la Commission d'Examen

Rapporteurs :	M. Dauchez	Professeur, Université de Reims Champagne-Ardenne (Reims)
	S. Antonczak	Professeur, Université de Nice Sophia-Antipolis (Nice)
Examineurs :	S. Pin	Chargé de recherche CNRS (CEA-Saclay)
	C. Ramseyer	Professeur, Université de Franche-Comté (Besançon)
	M. Enescu	Professeur, Université de Franche-Comté (Besançon)

Remerciements

Toute mon estime et mes premiers remerciements vont à mon directeur de thèse, Pr. Mironel Enescu. Je ne saurais lui exprimer toute ma reconnaissance pour sa rigueur scientifique et ses encouragements qui, durant toutes ces années de travail, m'ont permis de trouver les ressources nécessaires pour mener à terme ce travail. Je lui suis également reconnaissante pour sa disponibilité et ses qualités pédagogiques. J'ai beaucoup appris à ses côtés et je lui adresse toute ma gratitude pour m'avoir fait part de son savoir en chimie physique, chimie théorique et spectroscopie Raman.

Ce travail est un travail d'équipe et je tiens à remercier surtout Sarah Foley qui m'a beaucoup aidée pendant les expériences de chimie physique.

Je remercie Pr. Michel Fromm directeur du LCPR-AC, qui m'a accueilli dans son laboratoire ainsi que tous les permanents du laboratoire : Manuel Grivet, Christelle Brun-Bergeon, Christophe Mavon, Franck Berger, Jean-Baptiste Sanchez, Jean-Emmanuel Groetz.

Je tiens à remercier tout particulièrement Jean-Emmanuel Groetz pour tous ses nombreux conseils en ce qui concerne l'utilisation de LINUX et surtout de m'avoir initié à l'utilisation de L^AT_EX. Je remercie aussi mes autres collègues Nicolas, François, Omar et Zayd. Je remercie beaucoup Bruno pour sa disponibilité et sa patience surtout pour la lecture et les corrections de ce mémoire.

Je remercie tous les membres du jury pour avoir accepté de lire ce travail et de participer à la soutenance de cette thèse.

Merci beaucoup à Béatrice et Boutheïna pour leur bonne humeur et je n'oublierai pas non plus Elsa!

Je voudrais aussi remercier Sylvie Petot-Visini pour son aide et son joli sourire de chaque matin.

Enfin, je voudrais remercier tout particulièrement ma famille, ma mère, ma grand-mère et mon frère qui m'ont toujours soutenus et mon cher mari qui a eu toujours confiance en moi.

Résumé

Les ponts disulfure ont un rôle essentiel dans la stabilité des structures natives des protéines. L'étude du mécanisme de leur réduction en milieu biologique est très important pour la compréhension de leurs fonctions. Le but de notre travail a été d'analyser le rôle que la structure protéique joue dans la réactivité des ponts disulfure. Nous avons étudié le mécanisme de réduction des ponts disulfure dans les protéines par différents agents réducteurs en utilisant deux approches originales : la spectroscopie Raman et le calcul moléculaire. La spectroscopie Raman nous a permis de détecter directement l'avancement de la réaction de réduction des ponts. Les changements conformationnels qui accompagnent la coupure des ponts disulfure des protéines ont été analysés en détectant la bande Raman amide I. Les cinétiques de la réaction de réduction des ponts disulfure par le DTT et la TCEP ont été mesurées pour les albumines sériques et la lysozyme. Les hétérogénéités de ces cinétiques ont été analysées afin de mettre en évidence des différences de réactivité entre les ponts disulfure participants à la réaction. Ces résultats expérimentaux ont été affinés en utilisant le calcul moléculaire. Une étude détaillée a été effectuée dans le cas des quatre ponts disulfure de la lysozyme. Les enthalpies et les trajectoires de réaction ont été déterminées par la méthode QM-FE. Les résultats ainsi obtenus ont été discutés en relation avec les données expérimentales et les différents modèles de réaction.

Abstract

The disulfide bridges have a crucial role in the stability of native protein structures. The study of the mechanism of their reduction in biological environment is very important in order to understand their functions. The aim of our study was to analyze the role that protein structure plays in the reactivity of the disulfide bridges. We studied the mechanism of the disulfide bridges reduction in proteins by various reducing agents using two novel approaches : Raman spectroscopy and molecular modelling. Raman spectroscopy allowed us to directly detect the progress of the reduction reaction. The conformational changes following the cleavage of the protein disulfide bridges were analyzed by detecting the Raman band amide I. The kinetics of the reduction of the disulfide bridges by DTT and TCEP were measured for the serum albumins and for the lysozyme. The heterogeneity of these kinetics were analyzed to highlight the different reactivities of the disulfide bridges participating in the reaction. These experimental results were been refined using the molecular modelling. A detailed study has been done in the case of the four disulfide bridges of lysozyme. The reaction enthalpies and reaction paths were determined by using the QM-FE method. The results were discussed in relation with experimental data and with different reaction models.

Table des matières

Remerciements	i
Résumé	iii
Abstract	vii
Table des matières	ix
Liste des figures	xi
Liste des tableaux	xvii
Liste des abréviations	xvii
Introduction	1
1 Méthode expérimentale : la spectroscopie Raman	13
1.1 Principes	14
1.2 Méthode expérimentale	27
2 Méthodes de modélisation moléculaire	31
2.1 Calculs <i>ab initio</i>	32
2.2 Calculs MM	40
2.3 Calculs QM/MM	47
3 Attribution des bandes S-S dans les spectres Raman des protéines	51
3.1 Analyse des bandes Raman S-S de systèmes moléculaires simples	54
3.2 Analyse des bandes Raman S-S des albumines sériques	59
3.3 Analyse des bandes Raman S-S de la lysozyme	63
3.4 Les bandes S-S dans les spectres Raman d'autres protéines	66
3.5 Simulation des bandes Raman des quatre ponts S-S de la lysozyme	69

4 Réduction des ponts S-S dans les protéines : application de la spectroscopie Raman	79
4.1 Introduction	80
4.2 Réduction des ponts S-S dans les albumines sériques	82
4.3 Réduction des ponts S-S dans la lysozyme	105
5 Modélisation moléculaire de la réaction de réduction des ponts S-S dans la lysozyme par la TCEP	117
5.1 Introduction	118
5.2 Principe de la méthode	120
5.3 Méthodes de calcul	122
5.4 Résultats	128
5.5 Discussion	137
Conclusions	141
Annexe	143
Bibliographie	147

Liste des figures

1	Le dépliement de la structure protéique induit par le clivage des ponts S-S.	2
2	Pont S-S modèle. La conformation du pont est caractérisée par les cinq angles dièdres : $\chi_1, \chi_2, \chi_3, \chi'_1, \chi'_2$.	5
3	Structure secondaire des protéines et couplage des vibrations CO.	6
1.1	Représentation schématique de conditions expérimentales pour définir les directions de polarisation de la diffusion Raman.	26
1.2	Schéma du spectromètre Raman.	29
2.1	Exemple de traitement QM/MM sur le système lysozyme+TCEP : le système modèle est représenté par des billes et le reste du système par des traits.	49
3.1	Spectre Raman de la lysozyme en poudre mesuré en polarisation parallèle (rouge) et perpendiculaire (noir).	53
3.2	La L-cystine.	54
3.3	Bande Raman S-S de la cystine mesurée en solution aqueuse à une concentration de 16 mM et modélisée par deux fonctions de Lorentz.	55
3.4	Bande Raman S-S de la cystine mesurée en poudre et modélisée par une seule fonction de Lorentz (courbe rouge).	56
3.5	DTT oxydé.	56
3.6	Bande Raman S-S du DTT oxydé mesurée en poudre et modélisée par une seule fonction de Lorentz (courbe rouge).	57
3.7	Bande Raman S-S du DTT oxydé mesurée en solution aqueuse à une concentration de 16 mM et modélisée par une somme de trois fonctions de Lorentz.	58
3.8	Le glutathion oxydé.	58
3.9	Bande Raman S-S du GSSG mesurée en solution aqueuse à une concentration de 16 mM et modélisée par une somme de deux fonctions de Lorentz.	59

3.10	Bande Raman S-S du GSSG mesurée en poudre et modélisée par une somme de deux fonctions de Lorentz.	60
3.11	Spectre Raman d'une solution aqueuse de HSA native à 1 mM et pH=7,0.	61
3.12	Spectre Raman d'une solution aqueuse de HSA native à 0,5 mM (rouge) et en présence de TCEP à 20 mM (noir); les mesures ont été effectuées à pH=2,0.	62
3.13	Spectre Raman d'une solution aqueuse de HSA native à 1 mM et pH=7,0 mesurée en polarisation parallèle (rouge) et en polarisation perpendiculaire (noir).	63
3.14	Spectre Raman d'une solution aqueuse de BSA native à 1 mM et pH=7,0.	64
3.15	Spectre Raman d'une solution aqueuse de BSA native à 1 mM et pH=7,0 mesurée en polarisation parallèle (rouge) et en polarisation perpendiculaire (noir).	65
3.16	Spectres Raman d'une solution aqueuse de BSA native à 1 mM et pH=7,0 (rouge), pH=3,6 (bleu), pH=2,6 (noir).	66
3.17	Spectre Raman d'une solution aqueuse de lysozyme native à 2 mM et pH=3,2.	67
3.18	(a) Spectre Raman S-S d'une solution aqueuse de lysozyme native à 2 mM et (b) en présence de TCEP à 60 mM, après 8 heures de réaction; les mesures ont été effectuées à pH=3,2. La concentration de la protéine dans les deux cas est la même.	68
3.19	Spectre Raman d'une solution aqueuse de lysozyme native à 2 mM mesuré en polarisation parallèle (rouge) et perpendiculaire (noir).	69
3.20	Spectre Raman d'une solution aqueuse de BLG native à 2 mM et pH=7,0.	70
3.21	Spectre Raman d'une solution aqueuse de RNase A native à 2 mM(rouge) et en présence de TCEP à 10 mM (noir), à pH=8,13. La concentration de la protéine dans les deux cas est la même.	71
3.22	Spectre Raman d'une solution aqueuse de TRYPS native à 2 mM (rouge) et en présence de TCEP à 16 mM (noir), à pH=3,0. La concentration de la protéine dans les deux cas est la même.	72
3.23	Spectres Raman simulés des quatre ponts S-S de la lysozyme : cys6-cys127 (rouge), cys30-cys115 (bleu), cys64-cys80 (vert) et cys76-cys94 (noir). . . .	74
3.24	Spectres Raman simulés de la lysozyme : spectre Raman total de la lysozyme native (noir), de la lysozyme totalement réduite (bleu) et de la lysozyme en polarisation perpendiculaire (rouge).	75
3.25	Système modèle du pont cys30-cys115 utilisé dans les calculs QM/MM. La partie du système traitée en QM est représentée par des billes.	76

3.26	Spectre Raman de la lysozyme simulé par des calculs QM/MM.	76
4.1	Réaction d'échange des ponts disulfure entre la protéine et le DTT.	81
4.2	Réduction des ponts disulfure par la TCEP.	81
4.3	Spectre Raman d'une solution aqueuse de HSA à 1 mM et à pH 7,8 : (a) spectre de la protéine native (b) spectre de la protéine (même concentration) en présence du DTT à 30 mM.	83
4.4	Spectre Raman d'une solution aqueuse de BSA à 1 mM et à pH 7,8 : (a) spectre de la protéine native (b) spectre de la protéine (même concentration) en présence du DTT à 30 mM.	84
4.5	Les concentrations de ponts S-S de la HSA et du DTT oxydé sont représentées en fonction de la concentration de DTT ajouté. Les triangles représentent la contribution à la bande Raman S-S de la protéine, centrée sur 507 cm^{-1} , les carrés la contribution à la bande centrée sur 524 cm^{-1} et les cercles la contribution à la bande Raman S-S du DTT oxydé.	85
4.6	Les concentrations de ponts S-S de la BSA et du DTT oxydé sont représentées en fonction de la concentration de DTT ajouté. Les triangles représentent la contribution à la bande Raman S-S de la protéine, centrée sur 507 cm^{-1} , les carrés la contribution à la bande centrée sur 525 cm^{-1} et les cercles la contribution à la bande Raman S-S du DTT oxydé.	86
4.7	Courbe de titrage de l'acidité des groupes thiol dans la BSA obtenue par spectroscopie Raman.	87
4.8	Analyse de la bande Raman amide I d'une solution aqueuse de BSA à 1 mM (a) et d'une solution mixte : BSA à 1 mM et DTT à 30 mM (b). Le pH des solutions est de 7,8.	89
4.9	Analyse de la bande Raman amide I d'une solution aqueuse de HSA à 1 mM (a) et d'une solution mixte : HSA à 1 mM et DTT à 30 mM (b). Le pH des solutions est de 7,8.	90
4.10	Dépendance du rapport des surfaces des bandes 6 et 7 en fonction de la concentration de DTT ajouté (en rouge) comparée avec la dépendance de la concentration des ponts S-S formés dans le DTT en fonction du même paramètre (en noir).	91
4.11	Cinétique de réduction des ponts S-S mesurée pour une solution de BSA à 1 mM et de DTT à 17 mM à pH 7,8 et à la température de 22°C	92

-
- 4.12 Cinétique de réduction des ponts S-S mesurée pour une solution de BSA à 1 mM et de DTT à 17 mM à pH 7,8 et à la température de 22°C : dépendance de la différence des surfaces des bandes 6(ho) et 7(hu) en fonction du temps. 94
- 4.13 Cinétique de réduction du GSSG par le DTT. Les mesures des spectres Raman ont été effectuées sur des solutions aqueuses contenant 15 mM de GSSG et 15 mM de DTT à pH 5,9 et à la température de 22°C. 96
- 4.14 Modèle de pont S-S dans la protéine optimisé au niveau de théorie B3LYP/6-311+G(d,p). Les angles dièdres qui caractérisent la conformation du pont sont : $\chi_1 = -61, 1^\circ$, $\chi_2 = -71, 0^\circ$, $\chi_3 = -89, 3^\circ$, $\chi'_1 = -61, 1^\circ$, $\chi'_2 = -70, 9^\circ$. 98
- 4.15 Pont Cys75-Cys91 de la HSA réoptimisé au niveau de théorie B3LYP/6-311+G(d,p) en fixant les groupements méthyle. Les angles dièdres qui caractérisent la conformation du pont sont : $\chi_1 = -63, 4^\circ$, $\chi_2 = -68, 1^\circ$, $\chi_3 = -83, 3^\circ$, $\chi'_1 = -59, 0^\circ$, $\chi'_2 = -64, 1^\circ$ 99
- 4.16 Cinétique de la réaction de réduction des ponts S-S dans la BSA par la TCEP mesurée à pH=2,0 et à la température de 22°C. La concentration des ponts S-S a été évaluée à partir des surfaces des bandes Raman S-S. Les mesures ont été effectuées sur une solution aqueuse de BSA à 0,5 mM et TCEP à 20 mM. 102
- 4.17 Cinétique de la réaction de réduction des ponts S-S dans la HSA par la TCEP mesurée à pH=2,0 et à la température de 22°C. La concentration des ponts S-S a été évaluée à partir des surfaces des bandes Raman S-S. Les mesures ont été effectuées sur une solution aqueuse de BSA à 0,5 mM et TCEP à 20 mM. 103
- 4.18 Représentation de la structure 3D de la lysozyme (HWEL). Les quatre ponts S-S sont : Cys6-Cys127, Cys30-Cys115, Cys64-Cys80 et Cys76-Cys94. 105
- 4.19 Réduction des ponts S-S dans la lysozyme par la TCEP à pH=3,2 et à la température de 22°C : (a) spectre Raman d'une solution aqueuse de lysozyme native à 4 mM et (b) spectre Raman d'une solution aqueuse de lysozyme à 4 mM et TCEP à 60 mM mesuré 8 heures après le mélange. . . 109
- 4.20 Cinétique de la réduction des ponts S-S dans la lysozyme par la TCEP à pH=3,2 et à la température de 22°C. Le nombre de ponts S-S par molécule a été évalué à partir des surfaces des bandes Raman S-S. Les concentrations en lysozyme et TCEP sont de 4 mM et 60 mM, respectivement. 110
-

4.21	Cinétique de la réduction des ponts S-S dans la lysozyme par la TCEP à pH=4,4 et à la température de 22°C. Le nombre de ponts S-S par molécule a été évalué à partir des surfaces des bandes Raman S-S. Les concentrations en lysozyme et TCEP sont de 2 mM et 20 mM, respectivement.	111
4.22	Rapport entre la surface de la bande Raman S-S centrée sur 509 cm^{-1} (S^{509}) et celle centrée sur 528 cm^{-1} (S^{528}), en fonction du temps pendant la réaction de réduction des ponts S-S dans la lysozyme.	112
4.23	Évolution de la bande Raman amide I de la lysozyme lors de la réduction d'environ 30% de ses ponts S-S par la TCEP. Les spectres ont été mesurés sur des solutions aqueuses de lysozyme à 2 mM et TCEP à 60 mM à pH 3,2.	113
4.24	Évolution de la bande Raman amide I de la lysozyme lors de la réduction de 30% à 70% de ses ponts S-S par la TCEP. Les spectres ont été mesurés sur des solutions aqueuses de lysozyme à 2 mM et TCEP à 60 mM à pH 3,2.	114
4.25	Évolution de la bande Raman amide I de la BSA lors de la réduction de 0% à 75% de ses ponts S-S par le DTT. Les spectres ont été mesurés sur des solutions aqueuses de BSA à 1 mM et DTT à 17 mM à pH 7,8.	115
4.26	Potentiel de force moyenne calculé en fonction de la distance S-S entre les thiols des ponts réduits de la lysozyme.	116
5.1	Comparaison entre les moments dipolaires ESP (noir) et NPA (bleu) du système modèle calculés au long de la trajectoire de réaction.	125
5.2	Système réel étudié au niveau MM.	127
5.3	Système réel étudié au niveau QM/MM : le pont S-S et la TCEP en présence de molécules d'eau se trouvant à moins de 10 Å de l'atome de soluté le plus proche.	128
5.4	Système modèle pour la réaction de réduction des ponts S-S par la TCEP.	129
5.5	Première étape de la réaction de réduction des ponts S-S par la TCEP. . .	129
5.6	Courbes d'énergie pour l'interaction du pont S-S avec la TCEP (système modèle) : en noir l'énergie quantique dans le vide et en rouge l'énergie libre en solution aqueuse (modèle de solvation : PCM).	130
5.7	Étude de l'interaction du pont S-S avec la TCEP dans le vide par la méthode semi-empirique PM3.	131
5.8	État de transition (ET) de la réaction de réduction d'un pont S-S par la TCEP (système modèle) en solution aqueuse. Les calculs ont été effectués au niveau de théorie B3LYP/6-311+G(d,p) en utilisant le modèle de solvation PCM.	132

5.9	Courbe d'enthalpie libre de la réaction de réduction du GSSG par la TCEP, calculée par la méthode QM-FE.	133
5.10	Courbes d'enthalpie libre de la réaction de réduction des ponts S-S dans la lysozyme par la TCEP calculées en utilisant la méthode QM-FE : cys6-cys127 (noir), cys30-cys115 (bleu), cys64-cys80 (vert) et cys76-cys94 (rose).	134
5.11	Variation de l'énergie E_{QM} en fonction de la distance d(S-P) : cys6-cys127 (noir), cys30-cys115 (bleu), cys64-cys80 (vert), cys76-cys94 (rose) et GSSG (rouge).	135
5.12	Variation de l'énergie libre d'interaction avec l'environnement en fonction de la distance d(S-P) : cys6-cys127 (noir), cys30-cys115 (bleu), cys64-cys80 (vert), cys76-cys94 (rose) et GSSG (rouge).	136
5.13	Variation de la distance entre l'atome de soufre extérieur du pont et la molécule d'eau la plus proche au long de la trajectoire simulée de la lysozyme native : cys6 (noir), cys30(bleu), cys64(vert) et cys76 (rose).	139

Liste des tableaux

3.1	Conformations expérimentales et théoriques (optimisées en DFT) de quatre ponts S-S de la lysozyme. Comparaison entre les angles dièdres impliqués dans les structures des ponts S-S.	73
4.1	Constantes apparentes de vitesse de la réaction de réduction des albumines sériques par la TCEP à pH = 2,0. Sont données aussi les fractions molaires w_1 de l'espèce active TCEPH3.	103
4.2	Constantes intrinsèques de vitesse de réaction, rapide ($k_f^{(1)}$) et lente ($k_f^{(2)}$), calculées conformément aux équations 4.28 et 4.29.	104
4.3	Enthalpie libre de dépliement local pour la réaction de réduction des ponts S-S des albumines sériques par la TCEP et par le DTT.	104
4.4	Variation d'enthalpie libre locale lors de la réduction des ponts S-S de la lysozyme par la TCEP.	107
4.5	Constantes de vitesse de réaction rapide ($k_{ap}^{(1)}$) et lente ($k_{ap}^{(2)}$) déterminées pour deux valeurs du pH différentes. Les fractions des espèces actives, neutres et ionisées, de TCEP ont été évaluées pour chaque valeur du pH.	110
4.6	Constantes intrinsèques de vitesse de réaction rapide ($k_f^{(1)}$) et lente ($k_f^{(2)}$) déterminées pour les espèces actives, neutres et ionisées, de TCEP impliquées dans la réaction de réduction des ponts S-S de la lysozyme.	111
5.1	Charges atomiques partielles du système modèle calculées par la méthode ESP (voir la Figure 5.4).	124
5.2	Charges atomiques partielles du système modèle calculées par la méthode NPA (voir la Figure 5.4).	124
5.3	Contributions à la barrière d'énergie de réaction. Sont représentées les différences ΔE_{QM} et ΔF_{env} entre $d(S-P)=2,4 \text{ \AA}$ et $d(S-P)=7,0 \text{ \AA}$	136

Liste des abréviations

Pour des raisons de lisibilité la signification d'une abréviation ou d'un acronyme n'est souvent rappelée qu'à sa première apparition dans le texte d'un chapitre. Nous présentons ci-après la liste alphabétique des abréviations les plus usuelles, rencontrées dans la thèse, avec leur traduction en anglais.

Å	angström	angström
BLG	Bêta-lactoglobuline	Bêta-lactoglobulin
BPTI	Inhibiteur de Trypsine Pancréatique Bovine	Bovine Pancreatic Trypsin Inhibitor
BSA	Albumine Sérique Bovine	Bovine Serum Albumin
CO	groupement carbonyle	carbonyl group
CCD	système à charges couplées	charge coupled device
DFT	théorie de la fonctionnelle de la densité	density functional theory
DTT	dithiothréitol	dithiothreitol
DTTox	dithiothréitol oxydé	oxidized dithiothreitol
ESP	calcul des charges par la méthode du Potentiel Electrostatique	Electrostatic Potential method for charge calculation
ET	état de transition	transition state
FWHM	largeur spectrale à mi-hauteur (sous-entendu du maximum du pic)	Full Width at Half Maximum
GAFF	champ de force AMBER général pour presque toutes les molécules organiques	general AMBER force field for almost all organic molecules

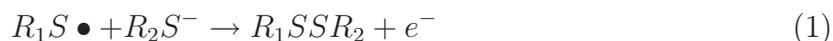
gg	gauche-gauche	gauche-gauche
gt	gauche-trans	gauche-trans
GSH	glutathion	glutathione
GSSG	glutathion oxydé	oxidized glutathione
HF	Hartree-Fock	Hartree-Fock
HSA	Albumine Sérique Humaine	Human Serum Albumin
Hz	Hertz	Hertz
IMOMO	(méthode) par orbitales moléculaires + orbitales moléculaires intégrées	integrated molecular orbital + molecular orbital
IR	infrarouge	infrared
kcal	kilocalorie	kilocalorie
kJ	kilojoule	kilojoule
LCAO	combinaison linéaire d'orbitales atomiques	linear combination of atomic orbitals
LYS	Lysozyme	Lysozyme
MM	mécanique moléculaire	molecular mechanics
mM	millimolaire	millimolar
Nd :YAG	grenat d'yttrium-aluminium dopé au néodyme	neodymium-doped yttrium aluminium garnet
nm	nanomètre	nanometer
NPA	analyse des populations (électroniques) naturelles	natural (electronic) population analysis
OA	orbitale atomique	atomic orbital
OM	orbitale moléculaire	molecular orbital
PC	complexe de produits	products complex
PCM	modèle du continuum polarisable	polarizable continuum model
PDI	protéine disulphide isomérase	protein disulfide isomerase
PED	distribution de l'énergie potentielle	potential energy distribution
PMF	Potentiel de Force Moyenne	Potential of Mean Force
QM	mécanique quantique	quantum mechanics
QM-FE	calcul de l'enthalpie libre par la méthode de la mécanique quantique	quantum mechanics free energy calculation

QM-PCM	calcul par mécanique quantique utilisant le modèle du continuum polarisable	quantum mechanics calculation using the polarizable continuum model
QM/MM	(méthode combinant) mécanique quantique et mécanique moléculaire	(method combining) quantum mechanics and molecular mechanics
RC	complexe de réactifs	reactants complex
RMSD	écart quadratique moyen	Root Mean Square Deviation
RNase A	Ribonucléase A	Ribonuclease A
RS	réactifs séparés	free reactants
SCF	champ autocohérent	self-consistent field
SPE	énergie à point fixe	single point energy
STO	orbitale de type Slater	Slater-type orbital
tt	trans-trans	trans-trans
TCEP	tris (2-carboxyéthyl)phosphine	tris(2-carboxyethyl)phosphine
TDI	méthode de l'intégration thermodynamique	thermodynamic integration method
TIP3P	modèle d'eau moléculaire	water molecular model
TRYPS	Trypsine	Trypsin
W	Watt	Watt
WHAM	Méthode d'analyse par Histogramme pondéré	Weighted Histogram Analysis Method

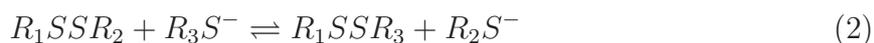
Introduction

Les ponts disulfure et leur rôle dans la structure protéique

Parmi les 22 acides aminés qui servent de briques élémentaires dans la construction des protéines, la cystéine occupe une place spéciale grâce à son groupement thiol (RSH). En plus de leurs propriétés réductrices et de leur efficacité dans la complexation des métaux, les thiols peuvent se relier entre eux par des liaisons covalentes pour constituer des ponts disulfure (S-S). Ce type de réaction implique la participation d'un thiol oxydé (radical thiyle, $RS\bullet$) et d'un thiol déprotoné (thiolate, RS^-) :



La chimie des ponts S-S est extrêmement intéressante parce que ces ponts peuvent être clivés facilement en réaction avec un antioxydant. En milieu biologique, l'antioxydant typique est le thiolate même :



La réaction 2 est connue sous le nom de *réaction d'échange thiol-disulfure* ou *réaction d'échange des ponts S-S* [1–4]. Le thiolate R_3S^- intervenant dans cette réaction pourrait être un thiolate naturel (le glutathione déprotoné, GS^-) ou bien un agent de synthèse tel le dithiothréitol (DTT). Outre le thiolate, il existe d'autres agents réducteurs efficaces pour la réduction des ponts S-S. Un cas intéressant du point de vue de l'efficacité est celui de la tris(2-carboxyethyl)phosphine (TCEP).

Les ponts S-S font souvent partie de la structure native des protéines et, dans ce cas, on les appellera des ponts natifs. À la différence des ponts natifs, les ponts mixtes sont formés entre les groupements thiols de la protéine et les agents antioxydants (à base de thiolate) ou bien entre deux groupements thiols de la protéine qui, dans la structure native, ne sont pas liés. Les organismes ont développé des mécanismes enzymatiques ayant pour but de corriger des ponts S-S mal formés. Ces mécanismes font intervenir des enzymes

telle la protéine disulfure isomérase (PDI) [5, 6].

Les ponts S-S natifs sont des éléments constitutifs essentiels de la structure native des protéines. Leur rôle et leur place dans cette structure a fait l'objet des nombreuses études. Il a été démontré que les ponts S-S stabilisent d'une manière significative la structure native des protéines [7–10]. Cet effet stabilisateur a été mis en évidence qualitativement en détectant les changements dans la conformation des protéines induits par la coupure des ponts S-S. En effet, cette coupure entraîne un dépliement global de la protéine (Figure 1). Pour plus de précision, rappelons que la structure protéique est caractérisée par quatre

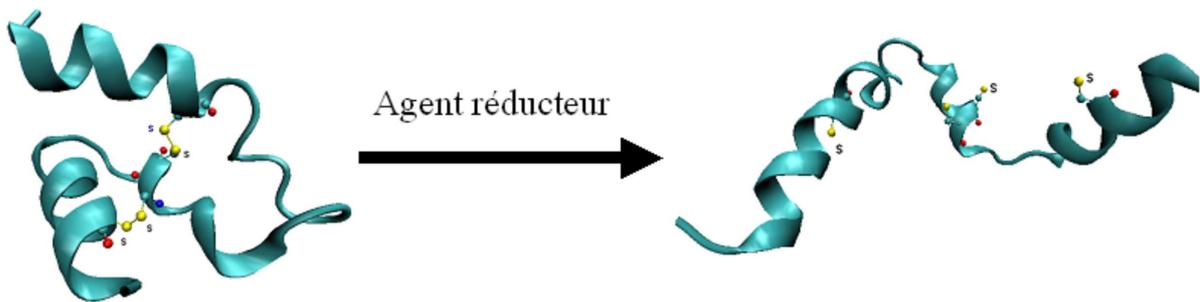


FIGURE 1 – Le dépliement de la structure protéique induit par le clivage des ponts S-S.

niveaux d'organisation :

1. La structure primaire, qui indique la séquence d'acides aminés dans la chaîne peptidique
2. La structure secondaire, qui indique l'organisation locale de la chaîne peptidique en hélices α , en brins et feuilletts β , un coudes (ou "random coils")
3. La structure tertiaire, qui correspond au repliement de la chaîne peptidique
4. La structure quaternaire, constituée par l'association d'au moins deux chaînes polypeptidiques par des liaisons, principalement non-covalentes, mais aussi par des liaisons S-S

Le dépliement de la protéine induit par la coupure des ponts S-S affecte surtout l'organisation quaternaire et tertiaire, mais la structure secondaire peut subir, elle aussi, des modifications.

À part leur rôle stabilisateur de la structure native des protéines, les ponts S-S semblent aussi guider le repliement des protéines vers leur structure native [11].

Objectifs du travail de thèse

Les réactions de coupure ou de formation des ponts disulfure trouvent une large application dans les études de dépliement et de repliement des protéines [12–16]. Dans ce type d'expériences, le dépliement de la protéine est provoqué par la réduction des ponts S-S, réalisée en utilisant des agents réducteurs (GS^- , DTT ou TCEP). Le repliement de la structure est induit par la reconstitution des ponts S-S natifs, obtenue par des réactions d'échange thiol-disulfure entre le glutathion oxydé (GSSG) ou le DTT oxydé, d'une part, et les groupes thiol de la protéine, d'autre part :



La réactivité des différents ponts S-S présents dans la structure des protéines n'est pas la même. En effet, elle semble être affectée d'une manière significative par la structure protéique environnante. Par conséquent, d'une protéine à l'autre, la constante de vitesse de réduction des ponts peut varier de plusieurs ordres de grandeur [4, 17, 18]. Comprendre les mécanismes de la réaction de réduction des ponts S-S et la manière dont la structure protéique est impliquée dans ces mécanismes présente alors un triple intérêt :

1. un intérêt pratique associé aux applications dans les études du dépliement-repliement de la structure protéique,
2. un intérêt spécifique, parce que la réactivité d'un pont est souvent liée à un certain rôle exercé par le pont dans un processus moléculaire spécifique,[5]
3. un intérêt général, puisque certains facteurs jouant un rôle dans la réaction de réduction des ponts S-S peuvent intervenir dans d'autres types de réaction en milieu biologique.

Dans cette optique, de nombreuses études ont été dédiées à la cinétique de réduction des ponts S-S modèles ou faisant partie d'une structure protéique [2–4, 17, 19–21]. Cependant, ces études souffrent d'une certaine limitation méthodologique : l'avancement de la réaction est déterminé d'une manière indirecte. La méthode est basée en principe sur l'inhibition de la réaction à un certain instant et l'évaluation des groupes thiol formés en les faisant réagir avec le réactif d'Ellman (réaction d'échange de ponts) [4]. En ce qui concerne les mécanismes proposés pour décrire le rôle de la structure protéique dans la

réaction de réduction des ponts [9], ils n'ont presque jamais été soumis au test décisif de la modélisation moléculaire. En effet, à notre connaissance, à part quelques études sur le clivage des ponts S-S par capture d'électron [22–26], il n'existe pas d'étude théorique sur la réaction de réduction des ponts S-S en milieu protéique.

Le travail de thèse présenté dans ce mémoire est une contribution à la compréhension des mécanismes de la réaction de réduction des ponts S-S dans les protéines, basée sur deux approches originales :

- a. l'utilisation de la spectroscopie Raman pour suivre l'avancement de la réaction et ses effets sur la structure de la protéine. En effet, la spectroscopie Raman est une méthode directe de détection de l'avancement de réaction basée sur la détection des bandes Raman de l'élongation S-S. Elle est aussi une méthode spécifique parce que l'intensité de ces bandes est directement liée à l'avancement de la réaction. D'autre part, la détection de la bande Raman amide I de la protéine donne des indices sur le processus de dépliement de la protéine qui accompagne la réduction des ponts S-S [27–29].
- b. la modélisation de la réaction basée sur l'utilisation des calculs moléculaires de mécanique quantique (QM) et de mécanique moléculaire (MM). Cette approche a été appliquée d'une manière systématique à l'étude de la réduction des quatre ponts S-S de la lysozyme par la TCEP.

Le développement de ce projet touche à quelques sujets épineux de la chimie-physique en milieu biologique : attribution des bandes dans les spectres Raman des protéines, calcul *ab initio* des spectres vibrationnels des molécules complexes, réactivité en milieu biologique et rôle de la structure protéique dans cette réactivité, application du calcul moléculaire à la modélisation des réactions en milieu biologique.

Attribution des bandes et interprétation des spectres Raman

Les études expérimentales effectuées sur des ponts disulfure modèles montrent que la fréquence de vibration associée à l'élongation S-S se situe dans l'intervalle spectral 505-545 cm^{-1} [30–33]. Le principal facteur qui affecte cette fréquence de vibration est la conformation relative des atomes $C_\alpha C_\beta S S' C'_\beta C'_\alpha$ autour des liaisons $C_\beta - S$ et $C'_\beta - S'$ (Figure 2).

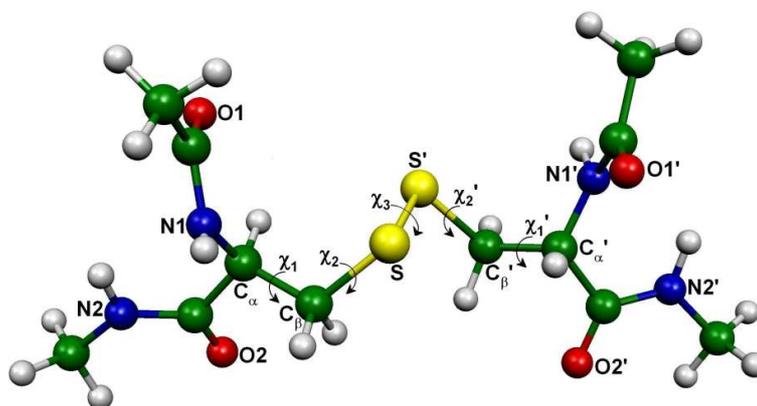


FIGURE 2 – Pont S-S modèle. La conformation du pont est caractérisée par les cinq angles dièdres : $\chi_1, \chi_2, \chi_3, \chi_1', \chi_2'$.

Ainsi, la rotation autour des liaisons $C - S$ produit des rotamères gauche (g) ou trans (t) qui, compte tenu du fait qu'il existe deux liaisons $C - S$, peuvent être classés en trois types : *gg*, *gt* et *tt*. L'analyse des données expérimentales indique que chacun de ces types est caractérisé par une fréquence d'élongation S-S spécifique : 510, 525 et 540 cm^{-1} , respectivement [31–37].

Les protéines dont la structure contient des ponts S-S présentent, elles aussi, des bandes Raman de vibration dans l'intervalle spectral 500-550 cm^{-1} . En analogie avec les ponts S-S modèles, il a été suggéré que dans ce cas aussi le facteur qui affecte la position de ces bandes est la configuration des atomes autour des liaisons $C - S$ [38]. Bien que cette hypothèse ait été acceptée par la plupart des chercheurs actifs dans ce domaine, elle n'a jamais été testée d'une manière systématique en l'absence d'une base des données significative (réunissant des structures expérimentales et des spectres Raman de bonne qualité). D'autre part, l'application de la spectroscopie Raman aux études de réactivité est évidemment basée sur les spectres Raman des protéines en solution. Il est bien connu que les structures expérimentales des protéines en solution sont beaucoup moins fiables que celles des protéines à l'état cristallin. D'une manière générale il est difficile de croire que, dans le cas d'un système si complexe que la protéine, d'autres facteurs, tels le couplage des modes ou les liaisons hydrogène, n'auraient pas d'effets importants sur la position des bandes S-S.

Le Chapitre 3 de ce mémoire de thèse présente une analyse des bandes Raman S-S de quelques protéines et systèmes modèles. Une étude détaillée a été effectuée dans le cas de la lysozyme. Cette analyse originale combine les mesures de dépolarisation des bandes, les mesures effectuées sur les formes réduites des protéines et les simulations des spectres Raman. Les conclusions de ce chapitre sont ensuite exploitées dans l'analyse des cinétiques

de réaction qui font l'objet du chapitre suivant.

Un autre élément majeur dans l'application de la spectroscopie Raman à l'étude des protéines concerne l'interprétation des modifications de la bande amide I qui accompagnent les changements de la structure protéique. La bande amide I, située dans l'intervalle spectral $1610\text{-}1700\text{ cm}^{-1}$, a son origine principalement dans les vibrations des groupements carbonyle (CO) de la chaîne polypeptidique. Les vibrations individuelles sont couplées par une interaction électrostatique dipolaire, [39–41] ce qui a pour conséquence la formation d'une bande Raman très large appelée bande amide I. Le couplage entre les vibrations CO individuelles dépend des positions relatives dans l'espace des liaisons CO qui, à leur tour, dépendent de la structure secondaire de la protéine (Figure 3).

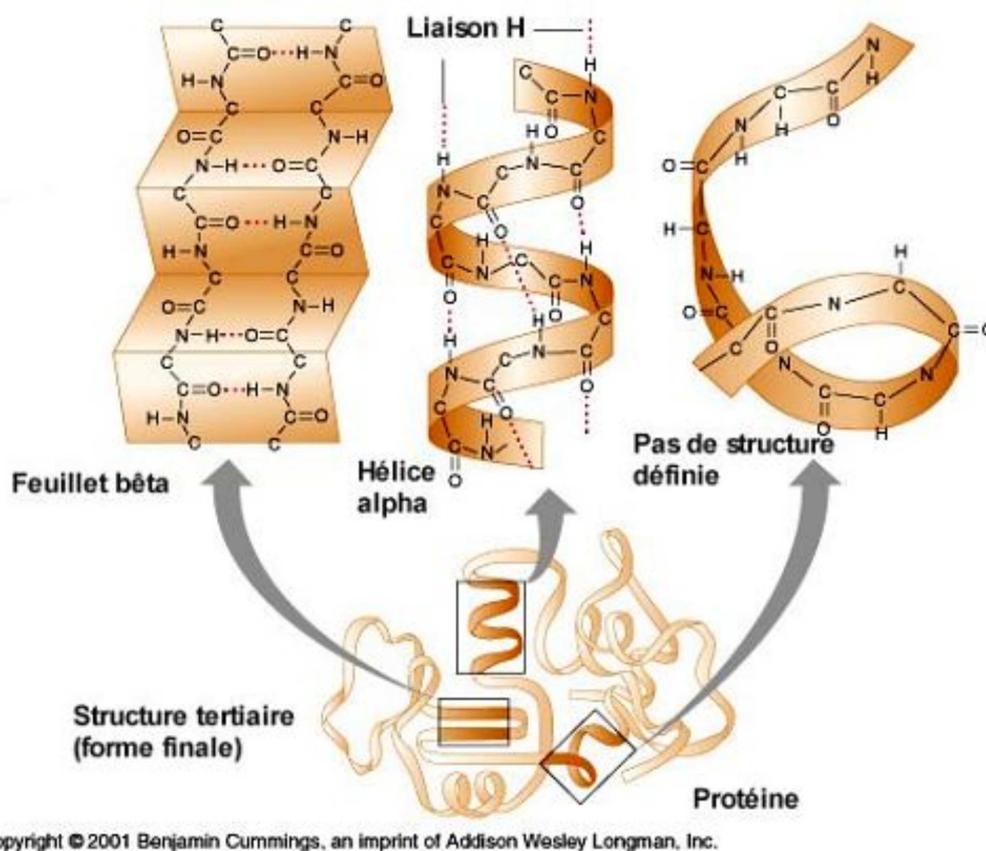


FIGURE 3 – Structure secondaire des protéines et couplage des vibrations CO.

Par conséquent, les différents types de structures secondaires (hélices α , feuillets β) sont caractérisés par des bandes amide I légèrement différentes par leur position et leur forme. L'analyse des bases de données de spectres Raman et IR et de structures de diffraction de rayons X des protéines a permis d'établir une corrélation entre les différents types de

structures secondaires et la position de la bande amide I [27, 42]. Ces résultats permettent, en principe, de suivre les modifications conformationnelles des protéines en détectant et analysant la bande Raman amide I. Cependant, à notre connaissance, il existe très peu de travaux de ce type rapportés dans la littérature de spécialité jusqu'à présent [29].

Dans le présent travail de thèse, nous avons utilisé la bande Raman amide I afin de détecter et d'interpréter les changements conformationnels qui accompagnent la coupure des ponts S-S des albumines sériques et de la lysozyme. Les résultats sont présentés dans le Chapitre 3 du mémoire.

Simulation des spectres vibrationnels des systèmes complexes

Les trois principales sources d'erreurs dans le calcul *ab initio* des spectres vibrationnels des systèmes complexes sont [43, 44] :

- a. les erreurs inhérentes à la méthode *ab initio* utilisée (qui affectent le calcul des dérivées de deuxième ordre de l'énergie potentielle)
- b. l'approximation harmonique (qui consiste à négliger le couplage des modes normaux)
- c. l'omission des effets d'environnement et le calcul effectué sur des systèmes modèles simplifiés

La première source d'erreurs ne peut être évitée qu'au prix fort, payé en temps de calcul. Étant donné que les systèmes étudiés contiennent d'habitude plusieurs dizaines d'atomes, la méthode de calcul qui s'impose presque de soi comme méthode supérieure limite est celle basée sur la théorie de la fonctionnelle de la densité (DFT). Il reste encore à choisir les potentiels d'échange-corrélation les mieux adaptés au système moléculaire étudié et, surtout, le jeu de fonctions de base [44].

L'approximation harmonique peut être corrigée en multipliant les constantes de force *ab initio* par des facteurs d'échelle soigneusement choisis [43]. Une autre méthode, radicale, consiste à abandonner complètement la méthode de calcul des fréquences de vibration basée sur l'analyse des modes normaux et d'utiliser à la place, la simulation de dynamique moléculaire [45–47]. La trajectoire ainsi générée permet de calculer la fonction d'auto-corrélation des vitesses de l'ensemble des noyaux appartenant au système moléculaire

considéré. Cette fonction globale est la somme des fonctions d'auto corrélation calculées pour chaque noyau. Le spectre de fréquences vibrationnelles (ou, précisément, la densité d'états vibrationnels) est alors obtenu en calculant la transformée de Fourier de la fonction d'auto corrélation des vitesses. Il faut préciser que les longueurs des trajectoires minima nécessaires pour effectuer un tel calcul avec une précision suffisante se situent dans le domaine des dizaines de picosecondes, ce qui paraît abordable au niveau actuel des moyens de calcul.

La troisième source d'erreurs dans le calcul des spectres vibrationnels peut être évitée en faisant appel à une méthode de calcul hybride, QM/MM. La méthode consiste à traiter le système moléculaire d'intérêt au niveau QM et son environnement au niveau MM. Cette méthode est facilement applicable dans le cas où l'environnement est simplement le solvant. Dans le cas des protéines, le système QM est entouré par la structure protéique avec laquelle il est couplé par des interactions non liées (Van der Waals et électrostatiques) mais aussi par des interactions covalentes. Malgré la complexité de ce couplage, la méthode QM/MM a été appliquée avec un certain succès dans le cas de quelques protéines [48].

Il est aussi possible de combiner la méthode QM/MM avec la simulation de dynamique moléculaire afin de s'affranchir en même temps de l'approximation harmonique. À ce jour nous n'avons pas connaissance de l'existence de travaux de ce type mais techniquement, ils sont déjà possibles. Cependant, aucune de ces méthodes n'améliorent les calculs d'intensités Raman. En effet, ces intensités ne peuvent être obtenues que par un calcul *ab initio*, puisque leur évaluation est basée sur le calcul des dérivées de la polarisabilité électronique par rapport aux coordonnées des noyaux (comme il sera montré dans le Chapitre 2 du mémoire).

Dans le présent travail de thèse nous avons simulé, en appliquant la méthode DFT, les spectres Raman des systèmes modèles construits pour les quatre ponts S-S de la lysozyme. Nous avons ensuite appliqué la méthode QM/MM pour calculer les mêmes spectres sur des systèmes modèles de taille plus grande incluant environ trente résidus autour de chaque pont S-S. Les résultats sont exposés dans le Chapitre 3 du mémoire. Ils ont été utilisés surtout dans l'interprétation de la structure des bandes Raman S-S de la lysozyme.

La structure protéique comme environnement des réactions chimiques

Les fragments moléculaires faisant partie de la structure protéique se trouvent dans

un environnement complexe qui peut influencer les processus chimiques auxquels ces fragments participent, principalement par trois facteurs :

1. les interactions électrostatiques avec les réactifs et les produits dont le rôle devient important lorsque la polarité du centre réactif change au cours de la réaction
2. l'accessibilité du site réactif qui dépend fortement de la conformation locale ou globale de la protéine
3. le changement conformationnel qui peut accompagner la réaction et qui stabilise les produits de réaction

La réaction de réduction des ponts S-S (équation 2) a pour résultat un changement important de la polarité du système réactif à cause de la formation d'une espèce chargée, le thiolate. Par conséquent, un environnement fortement polarisable stabilisera les produits de réaction. De ce point de vue, l'environnement le plus favorable à la réaction est le solvant même. La structure protéique peut affecter la constante de vitesse de la réaction d'une manière directe, lorsque des groupements chargés sont présents au voisinage du centre de réaction, ou indirecte, par l'exclusion partielle ou totale du solvant autour du centre de réaction. Dans le cas particulier d'une réaction d'échange de ponts, l'effet de la structure protéique sur la constante de vitesse de réaction peut être encore plus important : l'un des réactifs, le thiolate, étant chargé, son avancement vers le pont S-S peut être accéléré ou freiné par interaction électrostatique avec l'environnement. Des différences qui vont jusqu'à un facteur de 10^6 ont été mises en évidence entre les constantes de vitesse de cette réaction dans des environnements différents. Ces différences ont été expliquées par la présence d'interactions électrostatiques fortes [18, 49].

L'accessibilité de la cible de réaction par rapport au réactif extérieur a été reconue depuis longtemps comme un facteur décisif dans la réactivité en milieu protéique. De ce point de vue on fait la distinction entre les cibles de réaction exposées ou semi-exposées au solvant et celles cachées à l'intérieur de la structure protéique. Les cibles de réaction cachées peuvent être accessibles aux agents réactifs extérieurs par deux mécanismes :

1. par diffusion directe de l'agent à travers la structure protéique, si l'agent est une molécule de très petite taille
 2. par le dépliement local de la protéine qui a pour résultat l'exposition de la cible au solvant
-

Les antioxydants impliqués dans les réactions de réduction des ponts S-S sont des molécules de taille moyenne ou grande, ce qui rend inopérant le premier mécanisme dans ce cas. Il a été donc proposé que la réduction des ponts S-S cachés ou partiellement cachés dans la structure, passe par le dépliement de cette structure qui est le premier pas de la réaction [9, 50].

Dans le présent travail de thèse nous avons déterminé les cinétiques des réactions de réduction des ponts S-S dans les albumines sériques (par le DTT et la TCEP) et dans la lysozyme (par la TCEP). Les hétérogénéités de ces cinétiques ont été analysées afin de mettre en évidence des différences de réactivité entre les ponts S-S participants à la réaction. Les résultats, présentés dans le Chapitre 4 du mémoire, ont été interprétés par rapport aux mécanismes de réaction discutés ci-dessus.

La modélisation moléculaire des réactions dans l'environnement protéique

Les tentatives de modélisation moléculaire des réactions se déroulant dans un milieu complexe sont confrontées à deux exigences complémentaires :

1. traiter les transformations chimiques à un niveau de théorie suffisant pour une description quantitativement correcte de la réaction étudiée
2. effectuer une moyenne des résultats sur un ensemble de configurations statistiquement représentatif du système étudié

Ces deux exigences sont complémentaires dans le sens fort du mot, puisqu'augmenter la précision de traitement des interactions entraîne une réduction de l'espace de configurations accessible, en raison d'une forte augmentation du temps de calcul. Compte tenu de la puissance de calcul raisonnablement accessible, à ce jour il n'existe pas de solution complètement satisfaisante à ce problème. Plusieurs solutions de compromis sont disponibles, leur efficacité étant variable et fortement dépendante du système étudié [51] :

- a. Utilisation d'une méthode hybride QM/MM, basée sur un calcul du type QM semi-empirique.

Cette méthode permet de générer des ensembles de configurations de taille tout à fait satisfaisante. Mais une telle méthode ne peut être utilisée que dans les rares cas

où les méthodes semi-empiriques décrivent correctement la réaction en question. Les cas où cette description est suffisamment précise de point de vue quantitatif sont encore plus rares [52].

- b. Utilisation d'une méthode hybride QM/MM, où la méthode QM est une méthode de calcul *ab initio* ou DFT.

Dans ce cas on ne peut pas générer un ensemble statistique de configurations suffisamment large pour un calcul d'énergie libre. On doit alors se contenter du calcul d'une courbe d'énergie potentielle (en fonction de la coordonnée de réaction) [53–55]. Cette courbe peut avoir une certaine pertinence dans les cas où l'environnement n'est pas très flexible et l'interaction entre le système QM et cet environnement ne change pas trop d'une configuration à l'autre.

- c. Combiner les calculs QM effectués sur un système modèle et l'échantillonnage de l'espace des configurations de l'environnement réalisé au niveau MM (méthode QM-FE) [56, 57].

Dans un premier temps la trajectoire de réaction est déterminée en effectuant des calculs QM sur le système modèle. Des configurations de ce système extraites sur la trajectoire de réaction sont alors transférées dans le système réel. En fixant la coordonnée de réaction, un échantillonnage de l'espace de configurations du système réel est alors effectué et utilisé dans le calcul de la variation de l'énergie libre au long de la trajectoire de réaction. Le problème qui apparaît dans l'application de cette méthode dans sa version initiale est que, dans ces calculs, le système QM est découplé de son environnement. Les versions de cette méthode, ultérieurement proposées corrigent plus ou moins cette insuffisance.

- d. Utilisation d'une méthode hybride QM/MM avec des calculs QM effectués à deux niveaux de théorie : un niveau semi-empirique et un niveau *ab initio* ou DFT.

Dans ce cas, l'espace des configurations est échantillonné avec la méthode semi-empirique comme méthode QM mais, durant la simulation, les énergies et les forces sont corrigées par rapport à la méthode *ab initio* [58]. Évidemment, pour obtenir des résultats fiables il faut que la méthode semi-empirique utilisée décrive correctement, au moins du point de vue qualitatif, la réaction étudiée.

Dans ce travail de thèse nous avons abordé aussi la problématique de la modélisation moléculaire de la réaction de réduction des ponts S-S. Des résultats cohérents, présentés dans le Chapitre 5 du mémoire, ont été obtenus en utilisant une méthode du type QM-FE.

Ces résultats sont discutés en relation avec les données expérimentales présentées dans le Chapitre 4.

Les deux premiers chapitres du mémoire sont dédiés à la présentation générale des méthodes théoriques et expérimentales utilisées dans le présent travail. Certains détails de ces méthodes seront donnés dans les chapitres de présentation des résultats afin de faciliter le suivi du travail.

Chapitre 1

Méthode expérimentale : la spectroscopie Raman

Sommaire

1.1 Principes	14
1.1.1 États moléculaires vibrationnels	14
1.1.1.1 Modes normaux	14
1.1.1.2 Projection des modes normaux	16
1.1.2 Transitions vibrationnelles : règles de sélection	17
1.1.3 Diffusion Raman	19
1.1.3.1 Généralités	19
1.1.3.2 Diffusion Raman loin de la résonance	21
1.1.3.3 Diffusion Raman résonante	23
1.1.3.4 Calcul des intensités Raman loin de la résonance	24
1.1.3.5 Polarisation des bandes Raman	25
1.2 Méthode expérimentale	27
1.2.1 Spectromètre Raman	27
1.2.2 Traitement des données	29

1.1 Principes

1.1.1 États moléculaires vibrationnels

L'approximation adiabatique consiste à séparer le mouvement électronique du mouvement des noyaux. Par conséquent, dans cette approximation, les fonctions d'ondes Ψ des états vibrationnels stationnaires sont solutions de l'équation de Schrödinger :

$$\left[-\frac{\hbar^2}{2} \sum_{\alpha} \frac{1}{m_{\alpha}} \frac{\partial^2}{\partial X_{\alpha}^2} + U(X_1, X_2, \dots, X_n) \right] \Psi_{\text{noy}}(X_1, X_2, \dots, X_n) = E \Psi_{\text{noy}}(X_1, X_2, \dots, X_n) \quad (1.1)$$

où X_{α} et m_{α} sont les coordonnées et les masses des noyaux et $U(X_1, X_2, \dots, X_n)$ est l'énergie potentielle du système ($n=3N$, où N est le nombre d'atomes dans le système). Dans l'approximation harmonique, valable pour de faibles déplacements des noyaux autour d'une configuration d'équilibre, l'énergie potentielle s'écrit :

$$U(X_1, X_2, \dots, X_n) = U(X_1^{(0)}, X_2^{(0)}, \dots, X_n^{(0)}) + \frac{1}{2} \sum_{ij} \frac{\partial^2 U}{\partial x_i \partial x_j} x_i x_j \quad (1.2)$$

où $x_i = X_i - X_i^{(0)}$, $X_i^{(0)}$ sont les valeurs d'équilibre des coordonnées cartésiennes.

1.1.1.1 Modes normaux

Dans le traitement quantique des vibrations, le choix des coordonnées est très important. Une première simplification de l'équation 1.1 est obtenue en passant des coordonnées cartésiennes x_i aux coordonnées cartésiennes pondérées par les masses, $q_i = \sqrt{m_i} x_i$:

$$\left[-\frac{\hbar^2}{2} \sum_i \frac{\partial^2}{\partial q_i^2} + \frac{1}{2} \sum_{ij} \frac{\partial^2 U}{\partial q_i \partial q_j} q_i q_j \right] \Psi_{\text{noy}}(q_1, q_2, \dots, q_n) = E \Psi_{\text{noy}}(q_1, q_2, \dots, q_n) \quad (1.3)$$

Le Hessien $\frac{\partial^2 U}{\partial q_i \partial q_j}$ peut être diagonalisé en utilisant une transformation des coordonnées unitaire qui va préserver la forme diagonale de l'énergie cinétique. Les nouvelles coordonnées, Q_j , sont appelées coordonnées normales :

$$Q_j = \sum_i t_{ji} q_i \quad (1.4)$$

Les éléments matriciels t_{ji} satisfont le système d'équations [59] :

$$\sum_{i=1}^n t_{ji}(F_{ki} - \delta_{ki}\lambda_{jj}) = 0 \quad (1.5)$$

où $\lambda_{jj} = \frac{\partial^2 U}{\partial Q_j^2}$, $F_{ki} = \frac{\partial^2 U}{\partial q_k \partial q_i}$, $k = 1, 2, \dots, n$ et $j = 1, 2, \dots, n$. L'équation 1.3 peut donc être réécrite :

$$\left[-\frac{\hbar^2}{2} \sum_{i=1}^n \frac{\partial^2}{\partial Q_i^2} + \sum_{i=1}^n \frac{\partial^2 U}{\partial Q_i^2} Q_i^2 \right] \Psi_{noy}(Q_1, Q_2, \dots, Q_n) = E \Psi_{noy}(Q_1, Q_2, \dots, Q_n) \quad (1.6)$$

En diagonalisant le Hessien on obtient $3N$ constantes de force diagonales. Parmi celles-ci, 6 sont nulles et correspondent aux trois degrés de liberté de translation et aux trois degrés de liberté de rotation de la molécule. En conclusion le système a $3N-6$ coordonnées normales Q_i ($3N-5$ pour les molécules linéaires), leurs mouvements caractéristiques s'appelant modes normaux. Les $3N-6$ mouvements sont découplés et la fonction d'onde du système est un produit de fonctions d'onde uni-dimensionnelles, $\Psi_{v_\alpha}(Q_\alpha)$. Chacune de ces fonctions est la solution d'une équation Schrödinger du type oscillateur harmonique :

$$\left[-\frac{\hbar^2}{2} \frac{d^2}{dQ_\alpha^2} + \frac{1}{2} K_{Q_\alpha} Q_\alpha^2 \right] \Psi_{v_\alpha}(Q_\alpha) = \epsilon_{v_\alpha} \Psi_{v_\alpha}(Q_\alpha) \quad (1.7)$$

où $K_{Q_\alpha} = \frac{\partial^2 U}{\partial Q_\alpha^2}$ représente la constante de force du mode α et v_α est le nombre quantique de vibration correspondant à ce mode. L'énergie d'excitation du mode α est alors :

$$\epsilon_{v_\alpha} = \hbar \omega_\alpha \left(v_\alpha + \frac{1}{2} \right) \quad (1.8)$$

où $\omega_\alpha = \sqrt{K_{Q_\alpha}}$ est la fréquence angulaire du mode. Les modes normaux dont les fréquences sont identiques s'appellent modes dégénérés. Les fonctions d'onde des modes normaux non-dégénérés ont pour expression :

$$\Psi_\alpha(Q_\alpha) = N_\nu e^{(-\frac{1}{2} c_\alpha^2 Q_\alpha^2)} H_{\nu_\alpha}(c_\alpha Q_\alpha) \quad (1.9)$$

et celles des modes dégénérés :

$$\Psi_\alpha(Q_{\alpha_1}, Q_{\alpha_2} \dots) = N_\nu e^{(-\frac{1}{2} c_\alpha^2 \sum_{i=1}^g Q_{\alpha_i}^2)} \prod_{i=1}^g H_{\nu_{\alpha_i}}(c_\alpha Q_{\alpha_i}) \quad (1.10)$$

où N_ν est une constante de normalisation, $c_\alpha = \sqrt{\frac{\omega_\alpha}{\hbar}}$, H_{ν_α} le polynôme d'Hermite de rang ν_α et g le degré de dégénérescence du mode.

1.1.1.2 Projection des modes normaux

Afin de dégager la signification physique d'un mode normal il est utile de le décomposer, soit en déplacements cartésiennes, soit en déplacements des coordonnées internes.

Soit \vec{e}_{Q_i} le vecteur unitaire du mode normal i dans l'espace des coordonnées pondérées par les masses. La projection de ce vecteur unitaire sur le vecteur unitaire cartésien \vec{e}_j est :

$$t_{ij} = \vec{e}_{Q_i} \cdot \vec{e}_j \quad (1.11)$$

De la même manière on peut effectuer une projection dans l'espace des coordonnées non-pondérées :

$$t'_{ij} = \vec{e}'_{Q_i} \cdot \vec{e}_j \quad (1.12)$$

Ici \vec{e}'_{Q_i} est le vecteur unitaire du mode normal i dans l'espace des coordonnées non-pondérées. Le passage entre l'espace des coordonnées pondérées par les masses et l'espace des coordonnées non-pondérées se fait par la relation :

$$t_{ij} = \frac{\sqrt{m_j} t'_{ij}}{\sqrt{\sum_j m_j t'^2_{ij}}} \quad (1.13)$$

Ici m_j est la masse correspondant au déplacement le long du vecteur unitaire \vec{e}_j . Dans l'analyse d'un mode normal, les coordonnées cartésiennes et les coordonnées pondérées par les masses sont moins parlantes. On préfère un ensemble de coordonnées défini en termes de liaisons, d'angles et d'angles dièdres.

L'analyse d'un mode normal en terme de coordonnées internes dans l'espace des coordonnées non-pondérées se fait en projetant le vecteur unitaire \vec{e}'_{Q_i} sur les différents vecteurs unitaires des coordonnées internes \vec{e}_{ξ_j} :

$$l'_{ij} = \vec{e}'_{Q_i} \cdot \vec{e}_{\xi_j} \quad (1.14)$$

Alors, la projection d'un mode normal i d'amplitude Q'_i sur la direction de la coordonnée interne ξ_j sera :

$$\xi'_j = Q'_i \cdot l_{ij} \quad (1.15)$$

Pour aller encore plus loin dans la compréhension "physique" d'un mode normal il est possible d'analyser son énergie potentielle en terme de contributions des coordonnées internes (analyse s'appelant PED, *the Potential Energy Distribution*). Pour cela, on considère un ensemble complet de coordonnées internes ξ'_k . On peut définir une constante de force associée à chaque coordonnée ξ'_j :

$$K_{\xi'_j} = \left(\frac{\partial^2 U}{\partial \xi'^2_j} \right)_{[\xi'_k]} \quad (1.16)$$

où l'indice $[\xi'_k]$ signifie que les autres coordonnées de l'ensemble sont maintenues constantes lors de la dérivation. Lorsque l'on excite le mode normal i ayant l'amplitude Q'_i , son énergie potentielle maximum sera :

$$E_{p,Q'_i} = \frac{1}{2} K_{Q'_i} Q'^2_i \quad (1.17)$$

L'apport de la coordonnée interne ξ'_j à cette énergie potentielle est calculé en utilisant la constante de force respective et la projection ξ'_j :

$$E_{p,\xi'_j} = \frac{1}{2} K_{\xi'_j} (Q'_i l'_{ij})^2 \quad (1.18)$$

La contribution relative en pourcents est alors :

$$\frac{E_{p,\xi'_j}}{E_{p,Q'_i}} \times 100 = \frac{K_{\xi'_j} (l'_{ij})^2}{K_{Q'_i}} \times 100 \quad (1.19)$$

Il existe des modes ayant une contribution d'une certaine coordonnée interne qui est largement dominante. C'est le cas surtout des modes dont les fréquences sont supérieures à 2000 cm^{-1} . D'autres modes sont des mélanges plus ou moins dispersés des plusieurs coordonnées internes. Ces modes sont très sensibles à la structure moléculaire entière et ils ont une grande spécificité par rapport à cette structure. Dans les spectres vibrationnels ils donnent "l'empreinte" de la molécule.

1.1.2 Transitions vibrationnelles : règles de sélection

Propriétés de symétrie des coordonnées normales

L'application de la théorie des groupes est particulièrement utile dans l'étude des vibrations moléculaires. Soit \hat{R} une opération de symétrie appartenant au group de symétrie de la molécule. L'application de \hat{R} à la coordonnée normale Q_j doit la transformer en une combinaison des coordonnées de même fréquence. Cela est une conséquence directe du fait que l'énergie de vibration doit être invariante par rapport à \hat{R} . Autrement dit, dans toute transformation se rapportant au groupe de symétrie de la molécule, les coordonnées Q_{α_j} des modes normaux de même fréquence se transforment linéairement entre elles, et de telle façon que la somme des carrés $\sum_j (Q_{\alpha_j})^2$ reste inchangée :

$$\hat{R}Q_{\alpha_j} = \sum_i R_{ji}^\alpha Q_{\alpha_j} \quad (1.20)$$

En d'autres termes, les coordonnées Q_{α_j} réalisent une représentation irréductible du groupe de symétrie de la molécule. R^α sont les matrices de cette représentation.

Propriétés de symétrie des fonctions d'onde vibrationnelles

Ces propriétés découlent des propriétés de symétrie des coordonnées normales. Soit la fonction d'onde :

$$\Psi_{\nu_\alpha}(Q_\alpha) = N_\nu e^{(-\frac{1}{2}c_\alpha^2 Q_\alpha^2)} H_{\nu_\alpha}(c_\alpha Q_\alpha) \quad (1.21)$$

Le facteur exponentiel dans l'équation 1.21 est invariant dans toutes les transformations de symétrie (conformément au paragraphe précédent). Dans les polynômes d'Hermite, les termes de chaque degré se transforment exclusivement entre eux. Par conséquent, les propriétés de symétrie de chaque polynôme d'Hermite sont complètement déterminées par son terme de plus haut degré. La fonction Ψ_{ν_α} a les mêmes propriétés de symétrie que celles du H_{ν_α} et les propriétés de symétrie de H_{ν_α} sont les mêmes que celles du terme $Q_\alpha^{\nu_\alpha}$. La fonction d'onde vibrationnelle totale est un produit de fonctions $H_{\nu_\alpha}(\alpha = 1, 2, \dots)$ qui vont générer une représentation $\Gamma = \Gamma^\alpha \otimes \Gamma^{\alpha+1} \otimes \dots$, produit direct des représentations $\Gamma^\alpha, \Gamma^{\alpha+1}, \dots$

Règles de sélection

Théorème :

Soit $I = \int f(Q)dQ$

où Q représente l'ensemble des coordonnées normales. Supposons que f appartienne à la représentation Γ^f qui, en général, peut être réductible. Le terme I est non-nul seulement si Γ^f contient la représentation unité Γ^E .

Conséquence 1 :

Soit $f^{(\alpha)}$ et $f^{(\beta)}$ deux fonctions appartenant à deux représentations irréductibles Γ^α et Γ^β . Alors $\int f^{(\alpha)}(Q)f^{(\beta)}(Q)dQ = \Gamma^\alpha \otimes \Gamma^\beta \neq 0$ seulement si $\alpha = \beta$.

Conséquence 2 :

Soit $f^{(\alpha)}$, $f^{(\gamma)}$ et $f^{(\beta)}$ trois fonctions appartenant à trois représentations irréductibles Γ^α , Γ^γ et Γ^β .

$\int f^{(\alpha)}(Q)f^{(\gamma)}(Q)f^{(\beta)}(Q)dQ = \Gamma^\alpha \otimes \Gamma^\gamma \otimes \Gamma^\beta \neq 0$ seulement si le produit $\Gamma^\alpha \otimes \Gamma^\gamma$ contient la représentation Γ^β .

Application :

Soit \hat{h} l'hamiltonien décrivant une interaction qui peut induire une transition entre l'état

de fonction d'onde $\Psi^\alpha(Q)$ et celui de fonction d'onde $\Psi^\beta(Q)$. La probabilité de transition est proportionnelle au carré de l'élément de matrice respectif de \hat{h} :

$$P \approx \left| \int \Psi^{\alpha*}(Q) \hat{h} \Psi^\beta(Q) dQ \right|^2 \quad (1.22)$$

La probabilité P sera non-nulle seulement si le produit $\Gamma^\alpha \otimes \Gamma^h$ contient la représentation Γ^β . C'est la formulation la plus générale des règles de sélection pour les transitions vibrationnelles.

1.1.3 Diffusion Raman

1.1.3.1 Généralités

La diffusion inélastique de la lumière par les molécules peut être traitée à trois niveaux de théorie : le niveau classique, où les molécules et le rayonnement sont décrits dans le cadre de la physique classique, le niveau semi-classique, où le rayonnement est décrit par la théorie classique et les molécules par la théorie quantique, et le niveau quantique où le rayonnement et les états moléculaires sont quantifiés.

Le traitement classique offre une description intuitive du processus de diffusion mais il ne permet pas d'expliquer ses règles de sélection et les propriétés de polarisation du rayonnement diffusé. Ces inconvénients disparaissent dans le traitement semi-classique qui offre un cadre adéquat pour la simulation numérique des spectres Raman et pour la discussion des résultats présentés dans ce mémoire de thèse.

Ce traitement, proposé par A.C. Albrecht [60] a comme point de départ l'expression de l'intensité Raman diffusée par une molécule dans un angle solide de 4π . Cette expression a été obtenue après avoir effectué une moyenne par rapport aux orientations possibles de cette molécule :

$$I_{mn} = \frac{2^7 \pi^5}{3^2 c^4 \epsilon_0^2} I_0 (\nu_0 + \nu_{mn})^4 \sum_{\rho, \sigma} |(\alpha_{\rho\sigma})_{mn}|^2 \quad (1.23)$$

Ici n est l'état initial de la molécule, m est l'état final, $\nu_{mn} = (E_m - E_n)/h$, h la constante de Planck, I_0 et ν_0 sont l'intensité et la fréquence laser, respectivement. $(\alpha_{\rho\sigma})_{mn}$ est l'élément de matrice du tenseur polarisabilité de la molécule avec $\rho = x, y, z$ et $\sigma = x, y, z$, dans un référentiel solide à la molécule.

L'expression quantique de $(\alpha_{\rho\sigma})_{mn}$ est :

$$(\alpha_{\rho\sigma})_{mn} = \frac{1}{\hbar} \sum_r \left(\frac{(M_\rho)_{rn} (M_\sigma)_{mr}}{\nu_{rm} - \nu_0} + \frac{(M_\rho)_{rn} (M_\sigma)_{mr}}{\nu_{rn} + \nu_0} \right) \quad (1.24)$$

La somme est effectuée sur tous les états vibroniques de la molécule. Ici $(M_\rho)_{rn}$ et $(M_\sigma)_{mr}$ sont les amplitudes (scalaires) des moments dipolaires de transition respectifs :

$$(M_\rho)_{rn} = \int \Psi_r^* M_\rho \Psi_m dr \quad (1.25)$$

où Ψ_r et Ψ_m sont les fonctions d'onde vibroniques et M_ρ est la composante ρ du vecteur du moment dipolaire. Les fréquences ν_{rm} et ν_{rn} dans l'équation 1.24 sont calculées à partir des différences entre les énergies des états vibroniques respectives.

Dans l'approximation de Born-Oppenheimer d'ordre zéro, une fonction d'onde vibronique est le produit entre la fonction d'onde électronique et celle vibrationnelle. En utilisant la notation $m = gi$ (g pour l'état électronique fondamental et i pour l'état vibrationnel) on peut écrire :

$$\Psi_m = \Theta_g(\xi, Q) \Phi_i^g(Q) \quad (1.26)$$

où $\Theta_g(\xi, Q)$ est la fonction d'onde électronique de l'état g et $\Phi_i^g(Q)$ celle du i -ème état vibrationnel associé à ce même état électronique. ξ et Q représentent la totalité de coordonnées électroniques et de vibration, respectivement. L'équation 1.24 devient maintenant :

$$(\alpha_{\rho\sigma})_{gi,gj} = \frac{1}{h} \sum_{ev} \left(\frac{(M_\rho)_{ev,gj} (M_\sigma)_{gi,ev}}{\nu_{ev,gi} - \nu_0} + \frac{(M_\rho)_{gi,ev} (M_\sigma)_{ev,gj}}{\nu_{ev,gj} + \nu_0} \right) \quad (1.27)$$

Nous avons utilisé ici les notations ev pour l'état vibronique r (e pour l'état électronique, v pour l'état vibrationnel), gi pour l'état m et gj pour l'état n . Après la substitution de Ψ_r et Ψ_m , l'équation 1.25 devient :

$$(M_\rho)_{gi,ev} = \int (\Phi_v^e)^* (\mathcal{M}_\rho(Q))_{g,e} (\Phi_i^g) dQ \quad (1.28)$$

où

$$(\mathcal{M}_\rho(Q))_{g,e} = \int \Theta_e^* M_\rho \Theta_g d\xi \quad (1.29)$$

L'équation 1.29 donne le moment dipolaire de transition électronique en fonction des coordonnées de vibration Q . Afin d'effectuer l'intégrale 1.28 par rapport aux coordonnées de vibration Q , il faut trouver une forme explicite de cette dépendance. Le développement de $(M_\rho(Q))_{g,e}$ en série de Taylor en fonction de Q donne :

$$\mathcal{M}_{g,e} \cong \mathcal{M}_{g,e}^0 + \sum_s \lambda_{es}(Q) \mathcal{M}_{g,s}^0 \quad (1.30)$$

On montre, dans le cadre de la théorie Herzberg-Teller sur les transitions vibrationnelles, que :

$$\lambda_{es}(Q) = \left(\sum_a h_{es}^a Q_a \right) (\Delta E_{es}^0)^{-1} \quad (1.31)$$

Ici h_{es}^a est le facteur de couplage entre les états électroniques e et s réalisé par la coordonnée Q_a :

$$h_{es}^a = \int \Psi_e^* \frac{\partial \hat{H}}{\partial Q_a} \Psi_s d\xi \quad (1.32)$$

Compte tenu des équations 1.30 et 1.31, l'équation 1.28 devient :

$$(M_\rho)_{gi,ev} = (\mathcal{M}_\rho)_{g,e}^0 \langle ev|gi \rangle + \sum_{s,a} h_{es}^a (\Delta E_{es}^0)^{-1} (\mathcal{M}_\rho)_{g,s}^0 \langle ev|Q_a|gi \rangle \quad (1.33)$$

où $\langle ev|gi \rangle$ est le produit scalaire des fonctions d'onde vibrationnelles $\Phi_v^e(Q)$ et $\Phi_i^g(Q)$:

$$\langle ev|gi \rangle = \int \Phi_v^{e*}(Q) \Phi_i^g(Q) dQ \quad (1.34)$$

Après la substitution de l'équation 1.33 dans l'équation 1.27, l'expression finale du tenseur polarisabilité $(\alpha_{\rho\sigma})_{gi,gj}$, est :

$$(\alpha_{\rho\sigma})_{gi,gj} = A + B + C \quad (1.35)$$

où

$$A = \frac{1}{h} \sum_{ev} \left(\frac{1}{\nu_{ev,gi} - \nu_0} + \frac{1}{\nu_{ev,gj} + \nu_0} \right) [(\mathcal{M}_\rho)_{g,e}^0 (\mathcal{M}_\sigma)_{g,e}^0 (gi|ev)(ev|gj)] \quad (1.36)$$

$$B = \frac{1}{h} \sum_{ev} \left(\frac{1}{\nu_{ev,gi} - \nu_0} \right) \sum_{s,a} \frac{h_{es}^a}{\Delta E_{es}^0} \{ (\mathcal{M}_\rho)_{g,e}^0 (\mathcal{M}_\sigma)_{g,s}^0 (gj|ev)(ev|Q_a|gi) + (\mathcal{M}_\sigma)_{g,e}^0 (\mathcal{M}_\rho)_{g,s}^0 (gi|ev)(ev|Q_a|gj) \} \quad (1.37)$$

$$C = \frac{1}{h} \sum_{ev} \left(\frac{1}{\nu_{ev,gj} + \nu_0} \right) \sum_{s,a} \frac{h_{es}^a}{\Delta E_{es}^0} \{ (\mathcal{M}_\rho)_{g,e}^0 (\mathcal{M}_\sigma)_{g,s}^0 (gi|ev)(ev|Q_a|gj) + (\mathcal{M}_\sigma)_{g,e}^0 (\mathcal{M}_\rho)_{g,s}^0 (gj|ev)(ev|Q_a|gi) \} \quad (1.38)$$

Cette expression générale de $(\alpha_{\rho\sigma})_{gi,gj}$ conduit à des expressions simplifiées dans le cas de la diffusion Raman loin de la résonance ($\nu_{ev,gi} \gg \nu_0$) et, respectivement, à la résonance ($\nu_0 \cong \nu_{ev,gi}$).

1.1.3.2 Diffusion Raman loin de la résonance

En partant de la condition $\nu_{ev,gi} \gg \nu_0$ on peut remplacer $\nu_{ev,gi}$ par :

$$\nu_{e,g} \cong \frac{1}{h} (E_e^0 - E_g^0) = \nu_e \quad (1.39)$$

Cela revient à négliger la contribution vibrationnelle à $\nu_{ev,gi}$. La substitution de l'équation 1.39 dans l'équation 1.36 permet de séparer la somme sur v appliquée pour un état e donné

au produit $(gi|ev)(ev|gj)$. Cette somme est ensuite effectuée en utilisant la règle de Van Vleck :

$$\sum_v \langle gi|ev \rangle \langle ev|gj \rangle = \langle gi|gj \rangle = \delta_{ij} \quad (1.40)$$

L'équation 1.40 est valable lorsque la somme est effectuée sur tous les états vibrationnels. Pour les termes B et C la règle de Van Vleck s'écrit :

$$\sum_v \langle gi|ev \rangle \langle ev|Q_a|gj \rangle = \sum_v \langle gj|ev \rangle \langle ev|Q_a|gi \rangle = \langle gi|Q_a|gj \rangle \quad (1.41)$$

Les règles de sélection des transitions vibrationnelles décrites dans un paragraphe précédent donnent :

$$\langle gi|Q_a|gj \rangle = \begin{cases} 0 & \text{si } v_a^j \neq v_a^i \pm 1 \\ [(v_a^i + 1)/2\gamma_a]^{1/2} & \text{si } v_a^j = v_a^i + 1 \\ (v_a^i/2\gamma_a)^{1/2} & \text{si } v_a^j = v_a^i - 1 \end{cases} \quad (1.42)$$

Les valeurs pour les intégrales non-nulles ont été obtenues en utilisant les fonctions d'onde de l'oscillateur quantique harmonique. Ici v_a^i représente le nombre quantique du mode normal a excité dans son i -ème état vibrationnel. D'autre part, $\gamma_a = 4\pi^2\nu_a/h$, où ν_a est la fréquence caractéristique du mode normal a . En introduisant les équations 1.40 et 1.41 dans l'équation 1.35 et en remplaçant ΔE_{es}^0 par $-h(\nu_s - \nu_e)$, on obtient :

$$(\alpha_{\rho\sigma})_{gi,gj} = A'' + B'' \quad (1.43)$$

où

$$A'' = \frac{1}{h} \sum_e \left(\frac{2\nu_e}{\nu_e^2 - \nu_0^2} \right) (\mathcal{M}_\rho)_{g,e}^0 (\mathcal{M}_\sigma)_{g,e}^0 (gi|gj) \quad (1.44)$$

et

$$B'' = \frac{-1}{h} \sum_{e,s;s>e} \sum_a \frac{(\nu_e\nu_s + \nu_0^2)h_{es}^a}{(\nu_e^2 - \nu_0^2)(\nu_s^2 - \nu_0^2)} [(\mathcal{M}_\rho)_{g,e}^0 (\mathcal{M}_\sigma)_{g,s}^0 + (\mathcal{M}_\rho)_{g,s}^0 (\mathcal{M}_\sigma)_{g,e}^0] (gi|Q_a|gj) \quad (1.45)$$

Interprétation physique

Conformément à l'équation 1.42, le terme A'' est nul pour $i \neq j$ et il est responsable de la diffusion Rayleigh (diffusion élastique). D'autre part, le terme B'' est non-nul seulement pour les transitions pour lesquelles $v_a^j = v_a^i \pm 1$. Il s'agit donc d'une diffusion inélastique, la diffusion Raman (avec ses deux composantes, Stokes et anti-Stokes). On remarque aussi qu'à la différence des spectres vibrationnels infra rouges, les harmoniques

ne seront pas présentes dans les spectres Raman non-résonants. L'équation 1.45 montre aussi que la condition $\nu_a^j = \nu_a^i \pm 1$ est nécessaire mais pas suffisante pour que B'' soit non-nul. Il faut aussi qu'il existe deux états électroniques excités, e et s , couplés par la coordonnée vibrationnelle Q_a , ce qui signifie que $h_{es}^a \neq 0$. De plus, au moins l'un des produits $(\mathcal{M}_\rho)_{g,e}^0(\mathcal{M}_\sigma)_{g,s}^0$ et $(\mathcal{M}_\rho)_{g,s}^0(\mathcal{M}_\sigma)_{g,e}^0$ doit être non-nul. Cela implique que les deux transitions électroniques, $g \rightarrow e$ et $g \rightarrow s$, doivent être permises.

1.1.3.3 Diffusion Raman résonante

Lorsque $\nu_{ev,gi} - \nu_0$ est très faible, le facteur d'amortissement γ_e ne peut plus être négligé dans l'expression de $\alpha_{\rho\sigma}$. L'équation 1.27 devient alors :

$$(\alpha_{\rho\sigma})_{gi,gj} = \frac{1}{\hbar} \sum_{ev} \left(\frac{(\mathcal{M}_\rho)_{ev,gj}(\mathcal{M}_\sigma)_{gi,ev}}{\nu_{ev,gi} - \nu_0 + i\gamma_e} + \frac{(\mathcal{M}_\rho)_{gi,ev}(\mathcal{M}_\sigma)_{ev,gj}}{\nu_{ev,gj} + \nu_0 + i\gamma_e} \right) \quad (1.46)$$

où γ_e est la constante d'amortissement pour l'état électronique e .

Lorsque l'on est très proche de la résonance on prend seulement en compte dans la somme précédente l'état e pour lequel la condition de résonance est satisfaite. De plus, le deuxième terme dans l'expression de $\alpha_{\rho\sigma}$ est négligeable par rapport au premier parce que $\nu_{ev,gj} + \nu_0 \gg \nu_{ev,gj} - \nu_0$. Par conséquent, après l'utilisation du développement de $(\mathcal{M}_\rho)_{g,e}$ en fonction de Q on obtient :

$$(\alpha_{\rho\sigma})_{gi,gj} = A''' + B''' \quad (1.47)$$

où

$$A''' = \frac{1}{\hbar} \sum_v \left(\frac{(\mathcal{M}_\rho)_{g,e}^0(\mathcal{M}_\sigma)_{g,e}^0(gi|ev)(ev|gj)}{\nu_{ev,gj} + \nu_0 + i\gamma_e} \right) \quad (1.48)$$

$$B''' = \frac{-1}{\hbar^2} \sum_v \sum_s \sum_a h_{es}^a \left\{ [(\mathcal{M}_\rho)_{g,e}^0(\mathcal{M}_\sigma)_{g,e}^0(gj|ev)(ev|Q_a|gi) + (\mathcal{M}_\rho)_{g,e}^0(\mathcal{M}_\sigma)_{g,e}^0(gi|ev)(ev|Q_a|gj)] / (\nu_{ev,gj} + \nu_0 + i\gamma_e)(\nu_s - \nu_e) \right\} \quad (1.49)$$

Interprétation physique

Dans le cas où la conformation d'équilibre de la molécule dans l'état électronique e n'est pas identique à celle dans l'état g , les fonctions (ev) et (gi) où (ev) et (gj) ne sont pas orthogonales ce qui fait que la contribution A''' est non-nulle. D'autre part, l'absence de restrictions sur les valeurs de i et de j conduit à un spectre ayant des harmoniques correspondant aux transitions $(g0) \rightarrow (g1), (g0) \rightarrow (g2), (g0) \rightarrow (g3)$, etc.

La contribution de B''' est beaucoup plus faible que la contribution de A''' et elle doit

être prise en compte seulement quand $A''' = 0$.

Les conditions nécessaires pour que B''' soit non-nul sont :

- a. $h_{es}^a \neq 0$, ce qui signifie qu'il existe un autre état électronique excité s qui est couplé à l'état e par la coordonnée Q_a (la coordonnée de la transition vibrationnelle)
- b. $(\mathcal{M}_\rho)_{g,e}^0 \neq 0$ et $(\mathcal{M}_\sigma)_{g,s}^0 \neq 0$, ce qui veut dire que les transitions électroniques $g \rightarrow e$ et $g \rightarrow s$ sont permises toutes les deux
- c. si $A''' = 0$ cela implique que les géométries de la molécule dans l'état g et l'état e sont identiques. Par conséquent $(gi|ev) \neq 0$ seulement si $v_a^v = v_a^i \pm 1$ et $(gj|ev) \neq 0$ seulement si $v_a^v = v_a^j \pm 1$. La seule transition possible est alors la transition $v_a^j = v_a^i + 1$. Dans ce cas il n'y aura pas d'harmonique dans le spectre Raman résonant.

1.1.3.4 Calcul des intensités Raman loin de la résonance

Loin de la résonance, l'expression de l'intensité Raman totale est obtenue en substituant les équations 1.43 et 1.45 dans l'équation 1.23. En pratique, cette expression est difficile à utiliser parce qu'elle exige le calcul des états excités de la molécule. Une expression plus simple mais moins explicite est obtenue en utilisant le développement en série limitée de la polarisabilité $\alpha_{\sigma\rho}$ en fonction des coordonnées Q_a .

Dans l'approximation du premier ordre en Q_a :

$$\alpha_{\sigma\rho} = \alpha_{\sigma\rho}^0 + \sum_a \left(\frac{\partial \alpha_{\sigma\rho}}{\partial Q_a} \right) Q_a \quad (1.50)$$

et donc :

$$(\alpha_{\sigma\rho})_{mn} = \alpha_{\sigma\rho}^0 \langle m|n \rangle + \sum_a \left(\frac{\partial \alpha_{\sigma\rho}}{\partial Q_a} \right) \langle m|Q_a|n \rangle \quad (1.51)$$

Afin d'établir une analogie, on peut identifier le premier et le deuxième terme de l'équation 1.51 avec le premier et, respectivement, le deuxième terme de l'équation 1.43. Dans ces conditions, la contribution de la polarisabilité à l'effet Raman est donnée par l'expression :

$$(\alpha_{\sigma\rho})_{gi,gj}^R = \sum_a \left(\frac{\partial \alpha_{\sigma\rho}}{\partial Q_a} \right) \langle gi|Q_a|gj \rangle \quad (1.52)$$

Conformément à l'équation 1.42, pour une transition $v_a^i \rightarrow v_a^i + 1$ on peut écrire :

$$\langle gi|Q_a|gj \rangle = [(v_a^i + 1)/2\gamma_a]^{1/2} \quad (1.53)$$

La substitution des équations 1.52 et 1.53 dans l'équation 1.23 donne :

$$I_{gi,gj} = \frac{2^4 \pi^3}{3^2 c^4 \epsilon_0^2} I_0 (\nu_0 - \nu_a)^4 (v_a^i + 1) h / v_a \sum_{\rho, \sigma} \left(\frac{\partial \alpha_{\sigma\rho}}{\partial Q_a} \right)^2 \quad (1.54)$$

L'équation 1.54 est utilisée par la plupart de logiciels de calcul moléculaire *ab initio* pour le calcul d'intensités Raman. Dans ce cas, les dérivées $\frac{\partial \alpha_{\sigma\rho}}{\partial Q_a}$ ne sont pas évaluées à partir de l'équation 1.45 (ce qui exigerait un calcul *ab initio* d'états excités) mais d'une manière semi-classique : en dérivant, numériquement, par rapport à Q_a , les dérivées de deuxième ordre de l'énergie par rapport au champ électrique extérieur appliqué au système quantique.

1.1.3.5 Polarisation des bandes Raman

Afin d'analyser les propriétés de polarisation du rayonnement diffusé Raman, on considère une configuration expérimentale dans laquelle le rayonnement incident se propage le long de l'axe Z. La direction d'observation de la diffusion est la direction Y (observation perpendiculaire). Soit X la direction de polarisation du vecteur électrique dans le faisceau incident. L'intensité I_{\parallel} est l'intensité du rayonnement diffusé avec son vecteur électrique polarisé le long de l'axe X tandis que I_{\perp} est celle du rayonnement dont la polarisation est dirigée le long de l'axe Z (voir la Figure 1.1). Le rapport entre les deux intensités (I_{\perp}) et (I_{\parallel}) représente le facteur de dépolarisation ρ :

$$\rho = \frac{I_{\perp}}{I_{\parallel}} \quad (1.55)$$

Après avoir effectué une moyenne par rapport à toutes les orientations possibles de la molécule, les intensités I_{\perp} et I_{\parallel} ont les expressions suivantes [61] :

$$I_{\perp} = 3\beta_s(\alpha)^2 + 5\beta_a(\alpha)^2 \quad (1.56)$$

$$I_{\parallel} = 45(\alpha)^2 + 4\beta_s(\alpha)^2 \quad (1.57)$$

où :

$$(\alpha)^2 = \frac{1}{9} \text{Re} \left(\sum_{\sigma\rho} \alpha_{\sigma\sigma}^s \alpha_{\rho\rho}^{s*} \right) \quad (1.58)$$

$$\beta_s(\alpha)^2 = \frac{1}{2} \text{Re} \left(\sum_{\sigma\rho} 3\alpha_{\sigma\rho}^s \alpha_{\sigma\rho}^{s*} - \sum_{\sigma\rho} \alpha_{\sigma\sigma}^s \alpha_{\rho\rho}^{s*} \right) \quad (1.59)$$

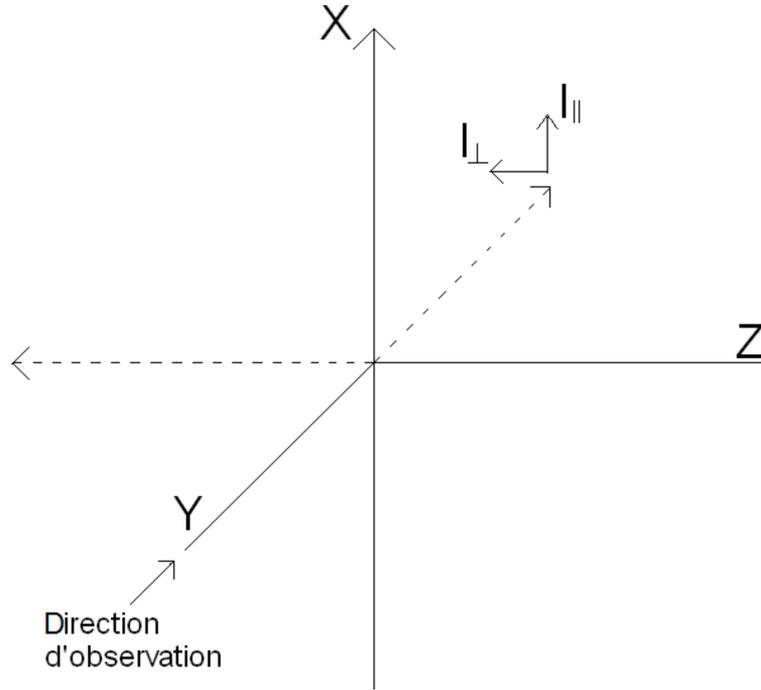


FIGURE 1.1 – Représentation schématique de conditions expérimentales pour définir les directions de polarisation de la diffusion Raman.

$$\beta_a(\alpha)^2 = \frac{1}{2} \text{Re} \left(\sum_{\sigma\rho} 3\alpha_{\sigma\rho}^a \alpha_{\sigma\rho}^{a*} \right) \quad (1.60)$$

$$\alpha_{\sigma\rho}^s = \frac{1}{2} (\alpha_{\sigma\rho} + \alpha_{\rho\sigma}) \quad (1.61)$$

$$\alpha_{\sigma\rho}^a = \frac{1}{2} (\alpha_{\sigma\rho} - \alpha_{\rho\sigma}) \quad (1.62)$$

L'équation 1.45 nous permet de déduire les conditions dans lesquelles le terme $\alpha_{\sigma\rho}$ est non-nul. D'une part, pour que la somme $(\mathcal{M}_\rho)_{g,e}^0 (\mathcal{M}_\sigma)_{g,s}^0 + (\mathcal{M}_\rho)_{g,s}^0 (\mathcal{M}_\sigma)_{g,e}^0$ soit non-nulle, Ψ_e doit appartenir à la représentation irréductible Γ_ρ et Ψ_s à la représentation irréductible Γ_σ , où l'inverse. D'autre part :

$$h_{es}^a = \int \Psi_e^* \frac{\partial \hat{\mathcal{H}}}{\partial Q_a} \Psi_s d\xi \quad (1.63)$$

(ayant donc la symétrie $\Gamma_\sigma \otimes \Gamma_a \otimes \Gamma_\rho$) où Γ_a est la représentation irréductible à laquelle appartient la coordonnée Q_a . Soit $\sigma \equiv \rho$. Alors h_{es}^a sera non-nul seulement pour $\Gamma_a \equiv \Gamma_E$, autrement dit, si Q_a est une coordonnée totalement symétrique. Pour tous les autres modes : $\alpha_{\rho\rho} \equiv 0, \forall \rho$. Cela implique $(\alpha)^2 = 0$ et, conformément aux équations 1.55, 1.56 et 1.57 : $\rho \geq 3/4$ [59].

En conclusion, tous les modes qui ne sont pas totalement symétriques ont un degré

de dépolarisation supérieur ou égal à $3/4$. Cela est un critère pratique pour identifier les modes totalement symétriques dans les spectres Raman des molécules possédant un certain degré de symétrie.

Dans le cas des molécules sans symétrie on peut discuter les propriétés de polarisation du rayonnement Raman en considérant la symétrie locale. Supposons que parmi les coordonnées internes qui apparaissent dans le développement de la coordonnée normale Q_a , il y en a seulement une (l_i) qui réalise un couplage entre deux états électroniques excités. Alors :

$$h_{es}^a = \sum_j \left(\int \Psi_e^* \frac{\partial \hat{\mathcal{H}}}{\partial l_j} \Psi_s d\xi \right) \frac{\partial l_j}{\partial Q_a} \cong \left(\int \Psi_e^* \frac{\partial \hat{\mathcal{H}}}{\partial l_j} \Psi_s d\xi \right) \frac{\partial l_j}{\partial Q_a} \quad (1.64)$$

Si les états e et s sont des états σ localisés sur les atomes A et B, l_i l'élongation de la liaison A-B et O_z l'axe dirigé le long de la liaison A-B, alors l'élément α_{zz} sera non-nul en raison de l'invariance de l'intégrant dans l'équation 1.45 par rapport aux opérations de symétrie locale. De plus, tous les autres éléments du tenseur α seront nuls. Dans ce cas, les équations 1.55-1.59 donnent $\rho = 1/3$. On tire la conclusion que dans le cas des molécules sans symétrie, le degré de dépolarisation d'un mode peut être inférieur à $3/4$ s'il est rendu Raman actif par une seule coordonnée interne de type élongation qui réalise un couplage entre deux états excités du type σ . Il est possible de montrer que lorsque les états excités couplés par une coordonnée interne du type élongation sont du type n (doublet non liant), le degré de dépolarisation peut être encore plus faible.

1.2 Méthode expérimentale

1.2.1 Spectromètre Raman

Le spectromètre Raman utilisé dans ce travail de thèse est un système non-conventionnel qui a été configuré dans le laboratoire. Les deux composantes essentielles du système sont les suivantes (Figure 1.2) :

- a. un laser Nd :YAG Spectra Physics, modèle LAB-170-10 délivrant des impulsions laser de 5 ns à une fréquence de répétition de 10 Hz.
- b. un système de détection Roper Scientifique constitué d'un monochromateur Spectra Pro 2500i et d'une camera CCD PIMAX-1024-RB de Princeton Instruments

La spécificité de cette configuration tient en quelques caractéristiques de fonctionnement :

- fonctionnement en régime impulsionnel. Cela permet de réduire le rapport entre le signal de fluorescence et le signal Raman. En effet, à partir d'un certain niveau d'intensité du faisceau laser incident, l'intensité de fluorescence est saturée tandis que l'intensité Raman continue à augmenter linéairement par rapport à l'intensité incidente.
- synchronisation de l'intensificateur de la camera CCD avec l'impulsion laser, l'intensificateur Q étant actif seulement pendant la durée de 5 ns de l'impulsion laser. Cette synchronisation permet de réduire d'une manière considérable le bruit d'obscurité du système de détection.
- accès à des fluences d'irradiation très élevées (à 532 nm la puissance maximum équivalente en continue du faisceau laser est de 4,5 W)

Ces caractéristiques ont pour conséquence une augmentation considérable de la sensibilité de détection du système. Cela permet l'analyse des échantillons de concentration relativement faible, de l'ordre d'un mM. Les spectres ont été enregistrés en utilisant la longueur d'onde d'émission de la deuxième harmonique (532 nm). Les échantillons (P) ont été placés dans une cellule en quartz de dimensions 1cm×1cm et ils ont été irradiés par un faisceau laser de 10 mm de diamètre à une densité de puissance équivalente de 35 mWmm^{-2} . La lumière diffusée a été détectée à 90° . Cette lumière est recueillie par un système optique constitué de deux lentilles (L1 et L2) en montage télescopique, et d'un filtre Notch (F) qui a pour fonction de rejeter le rayonnement laser à 532 nm. Les distances focales des lentilles L1 et L2 sont de 120 mm et de 100 mm, respectivement. La bande passante du filtre Notch est de 300 cm^{-1} ce qui impose une limite inférieure de 150 cm^{-1} aux fréquences Raman détectables par ce système. La sensibilité de détection de la lumière diffusée a été augmentée en plaçant un miroir sphérique M3 du côté de l'échantillon opposé par rapport à la fente (α) du monochromateur (MC). Le miroir plan M4 renvoie le faisceau pour une deuxième fois à travers l'échantillon ce qui a pour résultat le doublement de l'intensité du signal détecté. L'acquisition des spectres a été effectuée sur un nombre de 4000 à 8000 coups laser. La stabilité de l'échantillon par rapport à la durée d'irradiation a été testée en répétant l'acquisition plusieurs fois et en comparant les spectres obtenus. La largeur de bande instrumentale, testée en détectant la bande Raman de l'azote dans l'air à 2331 cm^{-1} , est de 8 cm^{-1} . D'autre part, la position des maxima des bandes localisées par le traitement des données est reproductible avec une erreur de $\pm 1 \text{ cm}^{-1}$.

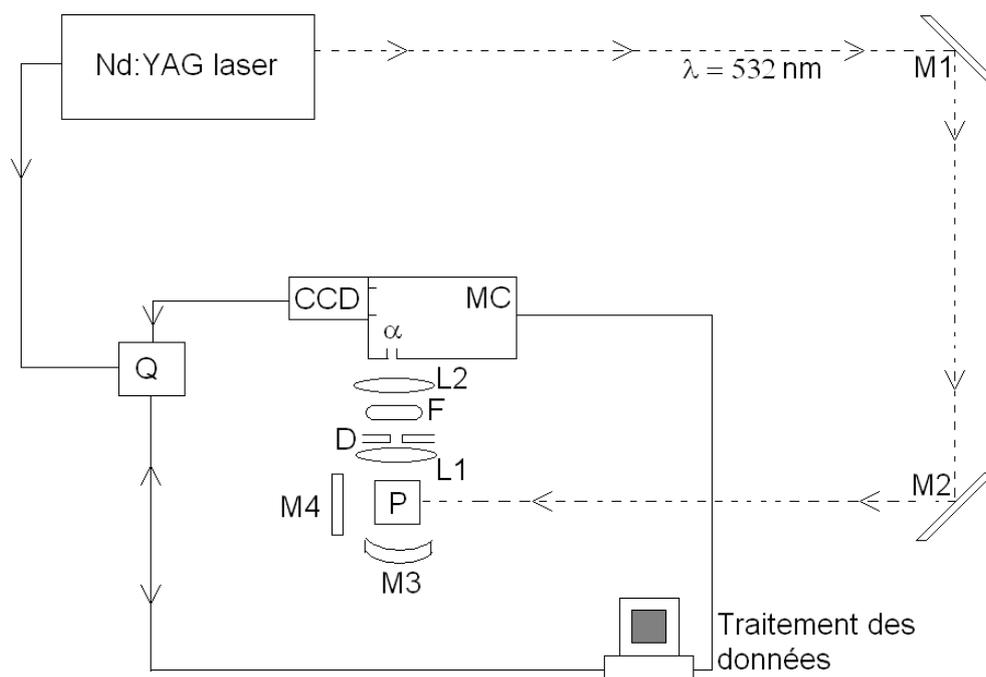


FIGURE 1.2 – Schéma du spectromètre Raman.

1.2.2 Traitement des données

Analyse de la bande Raman S-S

Dans la gamme spectrale où se trouvent les bandes S-S ($450-600\text{ cm}^{-1}$), le spectre Raman expérimental des protéines en solution est complexe et comporte plusieurs contributions : à part celle des bandes S-S, on y trouve la contribution du solvant, celle d'autres bandes de la protéine, celles des réactifs éventuellement présents dans la solution et, dans le cas des certaines protéines, une contribution de fluorescence. Dans ce spectre complexe, la séparation des bandes S-S peut, en principe, être obtenue par deux opérations : une soustraction physique des autres contributions préalablement déterminées en mesurant des solutions de référence, ou une séparation par traitement numérique des courbes expérimentales. La première opération a été effectuée systématiquement pour enlever la contribution du solvant et celle des réactifs présents en solution. Cependant, cette soustraction peut être imparfaite et une erreur de 1-2 % dans l'amplitude des spectres de référence peut contaminer le spectre corrigé d'une manière significative. C'est pour cela que, dans certains cas, la fonction modèle utilisée pour analyser le spectre corrigé doit inclure des faibles contributions provenant du solvant, des autres réactifs et, éventuellement, de la fluorescence de la protéine.

L'analyse des bandes Raman S-S des protéines a été effectuée en utilisant une fonction modèle qui consiste en une somme des fonctions lorentziennes dont les amplitudes, les largeurs et les fréquences des maxima sont des paramètres de modélisation. La modélisation a été effectuée avec le logiciel Sigmaplot. Dans le cas du suivi des cinétiques, seules les amplitudes des lorentziennes ont été maintenues comme paramètres ajustables dans la fonction modèle. Les largeurs des lorentziennes et les positions de leurs maxima ont été fixées par une modélisation préalable des bandes de la protéine native.

Analyse de la bande Raman amide I

La bande Raman amide I des protéines révèle des informations spéciales sur leurs structures secondaires [27, 39]. L'accès à ces informations requiert une décomposition de la bande en plusieurs contributions attribuées à des différents types de structures secondaires. Parmi les méthodes d'attribution proposées dans la littérature, nous avons choisi celle de Przybycien et al.[27]. Cette attribution est basée sur une comparaison entre les structures de rayons X et de les spectres IR et Raman de plusieurs dizaines de protéines. Le spectre Raman expérimental ($1500-1800\text{ cm}^{-1}$) a été modélisé par une somme de huit fonctions gaussiennes directement liées à la contribution de la protéine et une lorentzienne accompagnée par une droite, représentant l'importante contribution de l'eau. Comme dans le cas de l'analyse des bandes S-S, l'avancement de la réaction de réduction de ponts a été suivi dans la bande amide I en fixant les largeurs et les positions des gaussiennes aux valeurs trouvées pour la protéine en état natif. Seules les amplitudes relatives des gaussiennes restent comme variables d'ajustement.

Chapitre 2

Méthodes de modélisation moléculaire

Sommaire

2.1	Calculs <i>ab initio</i>	32
2.1.1	Théorie d'Hartree-Fock	32
2.1.2	Théorie de la fonctionnelle de la densité (DFT)	34
2.1.3	Jeux de fonctions de base	35
2.1.4	Calcul d'enthalpie libre	37
2.1.5	Théorie de l'état de transition	39
2.2	Calculs MM	40
2.2.1	Simulation de dynamique moléculaire	40
2.2.2	Champ de forces	41
2.2.3	Évaluation de l'énergie libre d'un système classique	43
2.2.4	Le potentiel de force moyenne (PMF)	45
2.3	Calculs QM/MM	47

2.1 Calculs *ab initio*

Les méthodes de calcul moléculaire *ab initio* sont basées sur la résolution de l'équation de Schrödinger :

$$\hat{H}\Psi(\vec{r}_i; \vec{R}_\alpha) = E\Psi(\vec{r}_i; \vec{R}_\alpha) \quad (2.1)$$

où \hat{H} est l'opérateur Hamiltonien de la molécule, $\Psi(\vec{r}_i; \vec{R}_\alpha)$ est la fonction d'onde et E l'énergie. Ici \vec{r}_i représente l'ensemble des vecteurs de position des électrons et \vec{R}_α celui des vecteurs de position des noyaux. Cette équation permet de déterminer les états stationnaires du système et leur énergies. L'hamiltonien \hat{H} est composé de cinq contributions : l'énergie cinétique des électrons \hat{T}_{el} , l'énergie cinétique des noyaux \hat{T}_{noy} , l'énergie potentielle de répulsion entre électrons \hat{V}_{el-el} , l'énergie potentielle de répulsion entre noyaux $\hat{V}_{noy-noy}$ et l'énergie potentielle d'attraction entre les noyaux et les électrons \hat{V}_{noy-el} :

$$\hat{H} = \hat{T}_{el} + \hat{T}_{noy} + \hat{V}_{el-el} + \hat{V}_{noy-noy} + \hat{V}_{noy-el} \quad (2.2)$$

En pratique, l'équation de Schrödinger ne peut pas être résolue de manière exacte et de ce fait, des approximations sont nécessaires. Afin de séparer le mouvement électronique de celui des noyaux, Born et Oppenheimer ont proposé en 1927 [62]. une approximation qui consiste à considérer que le déplacement des noyaux est beaucoup plus lent que celui des électrons. Par conséquent, l'état stationnaire des électrons s'établit pour chaque position des noyaux comme s'ils étaient fixes. Cela revient à négliger dans l'Hamiltonien le terme d'énergie cinétique des noyaux. La fonction d'onde électronique est alors la solution de l'équation Schrödinger :

$$\hat{H}_{el}\Psi_{el} = E\Psi_{el} \quad (2.3)$$

où $\hat{H}_{el} = \hat{T}_{el} + \hat{V}_{el-el} + \hat{V}_{noy-noy} + \hat{V}_{noy-el}$. Dans cette équation, les coordonnées des noyaux sont des simples paramètres. La séparation des deux mouvements, électronique et des noyaux, est acceptable surtout pour les états proches de l'état fondamental.

2.1.1 Théorie d'Hartree-Fock

Malgré l'approximation de Born-Oppenheimer, une solution exacte de l'équation de Schrödinger n'est pas trouvable pour des systèmes contenant 2 électrons ou plus. D'autres approximations sont nécessaires. La théorie d'Hartree-Fock (HF) est basée sur le principe variationnel qui postule que si E_1 est l'énergie de l'état fondamental d'un système, alors

pour toutes les fonctions d'onde possibles du système la valeur moyenne de l'énergie est supérieure ou égale à E_1 :

$$W = \frac{\int \Psi^* \hat{H} \Psi d\tau}{\int \Psi^* \Psi d\tau} \geq E_1 \quad (2.4)$$

Une solution approximative de l'équation de Schrödinger est alors obtenue en cherchant la fonction d'onde qui minimise W . Dans le cadre de la théorie d'Hartree-Fock on fait appel à une approximation spécifique aux systèmes multi-électroniques : l'approximation du champ auto cohérent (SCF). Elle consiste à considérer que chaque électron se déplace dans le champ moyen des autres. En conséquence, la fonction d'onde peut être écrite comme un produit de fonctions monoélectroniques (normalisées et orthogonales) appelées orbitales moléculaires (OM).

$$\Psi_{el}(\vec{r}) = \Phi_1(\vec{r}_1)\Phi_2(\vec{r}_2) \dots \Phi_n(\vec{r}_n) \quad (2.5)$$

Les fonctions d'onde monoélectroniques Φ_i sont les solutions de l'équation :

$$\hat{F}\Phi_i = E_i\Phi_i \quad (2.6)$$

où \hat{F} joue le rôle d'un Hamiltonien monoélectronique effectif (opérateur de Fock) dont la forme est issue de l'application du principe variationnel. Dans le cas d'une molécule à couches électroniques complètes, chaque orbitale est doublement occupée et la fonction d'onde totale est représentée par le déterminant :

$$\Psi_{el} = \frac{1}{\sqrt{(2n!)}} |\Phi_1(\vec{r}_1)\alpha(s_1)\Phi_1(\vec{r}_2)\beta(s_2) \dots \Phi_n(\vec{r}_{2n-1})\alpha(s_{2n-1})\Phi_n(\vec{r}_{2n})\beta(s_{2n})| \quad (2.7)$$

où α et β sont les fonctions de spin. La forme de déterminant de la fonction d'onde est exigée par le principe de Pauli : la fonction d'onde totale d'un système d'électrons doit être antisymétrique par rapport à la permutation des électrons.

Limites de la méthode d'Hartree-Fock

Considérer que chaque électron se déplace dans un champ moyen créé par les autres revient à négliger les corrélations entre les mouvements des électrons. Cet inconvénient provoque dans la plupart des cas des erreurs non négligeables dans les calculs d'énergie. Afin d'améliorer la fonction d'onde et l'énergie, on peut appliquer d'autres méthodes qui utilisent la fonction d'onde HF (et les fonctions dérivées d'elle) comme approximation de départ. Les méthodes de ce type sont appelées méthodes post-HF. Une alternative à ces méthodes est la méthode de la fonctionnelle de la densité.

2.1.2 Théorie de la fonctionnelle de la densité (DFT)

La DFT est une méthode très populaire qui prend en compte la corrélation des électrons mais qui n'est pas une méthode post-Hatree-Fock. Elle est fondée sur le principe que la vrai observable du système quantique (la molécule) est la densité électronique [63, 64]. Par conséquent, si on connaît la densité électronique (ρ_0) d'une molécule supposée dans l'état fondamental, on peut calculer l'énergie de cet état (E_0) et sa fonction d'onde (Ψ_0). Cette correspondance densité-énergie permet d'écrire l'énergie comme une fonctionnelle de ρ :

$$E_0 = E(\rho_0) = T(\rho_0) + V_{el-noy}(\rho_0) + V_{el-el}(\rho_0) + E_{XC} \quad (2.8)$$

où E_{XC} est un terme d'échange-corrélation qui inclut l'énergie d'échange (X) entre électrons et leur corrélation dynamique (C). Afin de déterminer la densité $\rho(r)$, l'algorithme fait appel à un système fictif d'électrons qui n'interagissent pas entre eux mais dont la densité électronique est égale à la densité du système réel. Ce système peut être décrit d'une manière exacte par un seul déterminant formé par des OM indépendantes. Dans ce cas, la densité électronique réelle est donnée par l'expression :

$$\rho = \sum_i |\Psi_i|^2 \quad (2.9)$$

où Ψ_i est l'orbitale fictive occupée par l'électron i . En appliquant la méthode variationnelle qui consiste à minimiser E par rapport à ρ , on obtient un système d'équations pour les orbitales fictives Ψ_i . L'équation satisfaite par ces fonctions d'onde ressemble à une équation de Schrödinger :

$$-\frac{\hbar^2}{2m_e^2} \nabla^2 \Psi_i + (V_{el-noy} + V_{el-el} + E_{XC}) \Psi_i = \epsilon_i \Psi_i \quad (2.10)$$

Tandis que V_{el-noy} et V_{el-el} sont calculés à partir des expressions exactes des opérateurs respectifs, E_{XC} est calculé à partir d'une expression semi-empirique. C'est la spécificité de la DFT qui fait que cette méthode n'est pas considérée comme étant strictement *ab initio*. Cependant, dans un langage simplifié, elle est appelée, elle aussi, méthode *ab initio*. Cela sera souvent le cas au long de ce mémoire.

Il existe différentes versions d'applications de la DFT et leur spécificité est donnée par l'expression utilisée pour E_{XC} . La méthode DFT a été largement utilisée dans notre étude avec les potentiels E_{XC} hybrides B3LYP [65–68].

Performances et limitations de la DFT

La DFT est surtout appliquée pour corriger les effets de corrélation électronique dans le cas des systèmes de taille moyenne et grande (50-200 atomes). Utilisée avec des ensembles de fonctions de base du type valence séparée triple-zeta, cette méthode donne une erreur moyenne d'énergie d'atomisation des petites molécules d'environ 5 kcal.mol^{-1} . Le problème est que l'erreur peut varier d'une manière imprévisible d'un système à l'autre. Malgré cela, la méthode DFT reste très populaire car pour des systèmes dont la taille dépasse les 50 atomes, elle n'a pratiquement pas de concurrent.

2.1.3 Jeux de fonctions de base

Afin d'obtenir des représentations mathématiques explicites des orbitales moléculaires, Roothaan propose en 1951 [69] de les écrire comme des combinaisons linéaires d'orbitales atomiques (approximation "LCAO") :

$$\Psi_i = \sum_{\mu=1}^N c_{\mu i} \chi_{\mu} \quad (2.11)$$

où χ_{μ} sont des orbitales atomiques et les $c_{\mu i}$ les coefficients d'expansion. Les méthodes d'orbitales moléculaires implémentées dans les logiciels des calculs *ab initio* sont basées sur cette représentation. Le choix des orbitales atomiques χ_{μ} (fonctions de base) est essentiel pour l'efficacité et la précision de la méthode.

En ce qui concerne la nature des orbitales atomiques, plusieurs choix sont possibles.

Orbitales de Slater (e^{-r})

Les orbitales de type "Slater" (STO) semblent être du point de vue physique le choix naturel des fonctions de base : elles possèdent un point de rebroussement à l'origine, et une décroissance exponentielle à l'infini :

$$\chi^{STO} = N r_A^{n-1} e^{-\frac{\zeta r_A}{a_0}} Y_l^m(\Theta_A, \varphi_A) \quad (2.12)$$

où Y_l^m est l'harmonique sphérique, n, m, l sont les trois nombres quantiques, a_0 le rayon de Bohr, N la constante de normalisation et ζ l'exposant orbital toujours positif. Cependant, dans ce type de fonction, l'exponentielle pose de grandes difficultés dans le calcul des intégrales lorsque plus de deux atomes sont présents.

Gaussiennes (e^{-r^2})

On peut remplacer la fonction de base *STO* par une gaussienne de nature "cartésienne" (*GTF*) :

$$g_{ijk} = N x_A^i y_A^j z_A^k e^{-\zeta r_A^2} \quad (2.13)$$

où i, j, k sont des entiers positifs ou nuls, ζ l'exposant orbital (positif), et x_A, y_A et z_A les coordonnées cartésiennes du noyau de l'atome A. Si $i + j + k = 0$, l'orbitale sera de type s . Si $i + j + k = 1$ l'orbitale sera de type p , etc. Cette gaussienne présente des différences avec l'exponentielle : elle décroît plus vite quand on s'éloigne du noyau, et, de plus, elle n'a pas le même comportement pour $r = 0$ (par exemple, sa dérivée est nulle). Pour décrire les régions proches du noyau, il est nécessaire d'utiliser une combinaison linéaire de gaussiennes afin d'obtenir la qualité d'une fonction de Slater :

$$\chi_\mu = \sum_p d_{\mu p} g_p \quad (2.14)$$

où χ_μ est la gaussienne contractée, $d_{\mu p}$ sont des coefficients de contraction (fixés pour un jeu de fonctions de base donné) et g_p la gaussienne primitive de type g_{ijk} . La précision d'un calcul d'OM dépend d'une manière critique de la taille du jeu de fonctions de base choisi. Le choix le plus économique est celui du jeu de fonctions de base minimal. Le jeu de fonctions de base minimal attribue à chaque sous-couche atomique ($1s, 2s, 2p, \dots$) des g_{ijk} ayant le même exposant (un exposant par sous-couche). Les jeux étendus apportent des améliorations à la description des orbitales moléculaire. Ils peuvent être constitués de :

- Fonctions de base double-zeta (DZ), chaque sous-couche étant décrite par deux GTF.
- Fonctions de base triple-zeta (TZ), chaque sous-couche étant décrite par trois GTF.
- Fonctions de base valence séparée double-zeta (VDZ), les couches intérieures étant représentées par une seule fonction par sous-couche, celles de valence par deux fonctions par sous-couche.

Fonctions de polarisation

Dans la molécule, les orbitales atomiques sont déformées par rapport aux orbitales atomiques des atomes libres. Pour prendre en compte cet aspect il est possible d'ajouter au jeu de fonctions de base des fonctions ayant le nombre quantique azimutal l supérieur

à celui de la sous-couche de valence. Cela permet un modelage angulaire de la densité électronique. Ce type de fonction est généralement indispensable pour une bonne description de la liaison chimique.

Fonctions diffuses

Afin de décrire des régions très éloignées par rapport aux noyaux centraux, on ajoute des fonctions de base (généralement *p* ou *d*) d'exposant ζ très faible. Elles sont importantes pour décrire les interactions intermoléculaires faibles comme par exemple les liaisons hydrogènes entre deux molécules.

Les calculs *ab initio* présentés dans ce travail de thèse ont été effectués avec le logiciel Gaussian 03 [70].

2.1.4 Calcul d'enthalpie libre

Les prédictions sur l'évolution d'un système moléculaire à température et pression constantes exigent le calcul de l'enthalpie libre :

$$G = U + PV - TS \quad (2.15)$$

où S est l'entropie du système, T la température, U est l'énergie interne, P la pression et V le volume du système.

La fonction de partition d'un système se trouvant dans un état macroscopique M s'exprime de la façon suivante :

$$Q_M = \sum_i e^{-\epsilon_i/k_B T} \quad (2.16)$$

où ϵ_i est l'énergie du micro-état i , k_B est la constante de Boltzmann, et la somme est effectuée sur tous les micro-états appartenant à l'état M . L'enthalpie libre d'un gaz se trouvant dans un état macroscopique M associé à un minimum local (sur la surface d'énergie potentielle) est :

$$g_M = u_{p,M} + PV - k_B T \ln Q_M \quad (2.17)$$

Ici $u_{p,M}$ est l'énergie potentielle obtenue par un calcul quantique, le minimum d'énergie étant choisi comme origine de ϵ_i dans l'équation 2.16. La fonction de partition Q_M d'un

système de N molécules occupant un volume V s'exprime en fonction de la fonction de partition q d'une seule molécule en tenant compte aussi de l'indiscernabilité des molécules :

$$Q_M = \frac{q_M^N}{N!} \quad (2.18)$$

où q_M est le produit des fonctions de partition qui correspondent aux quatre mouvements indépendants de la molécule (électronique, de translation, de vibration et de rotation) :

$$q_M = q_M^e q_M^T q_M^V q_M^R \quad (2.19)$$

Le terme q_M^e est égal à la multiplicité de l'état électronique (1 pour les singulets). Le terme q_M^T est donné par :

$$q_M^T = \frac{V}{\Lambda^3} \quad (2.20)$$

avec $\Lambda = h \left(\frac{1}{2\pi k_B T} \right)^{1/2}$ où h est la constante de Plack, k_B est la constante de Boltzmann et T la température. Le terme q_M^R a pour expression :

$$q_M^R \cong \left(\frac{k_B T}{hc} \right)^{3/2} \left(\frac{\pi}{ABC} \right)^{1/2} \quad (2.21)$$

où A, B, C sont les trois constantes rotationnelle et c la vitesse de la lumière dans le vide. Le terme q_M^V est le produit des fonctions de partition de tous les modes normaux de la molécule. Pour un seul mode on peut écrire :

$$q_M^{V_i} = \frac{e^{-h\nu_i/2k_B T}}{1 - e^{-h\nu_i/k_B T}} \quad (2.22)$$

La fonction de partition vibrationnelle totale sera :

$$q_M^V = \prod_i q_M^{V_i} \quad (2.23)$$

Après avoir appliqué l'approximation de Stirling dans l'évaluation de $\ln N!$, l'équation 2.17 peut être réécrite :

$$g_M = n \left(U_{p,M} - RT \ln \frac{q_M}{N_A} \right) \quad (2.24)$$

où n est le nombre de moles, N_A le nombre d'Avogadro, $U_{p,M}$ l'énergie potentielle molaire et R la constante des gaz. L'enthalpie libre molaire est alors :

$$G_M = \frac{g_M}{n} = U_{p,M} - RT \ln \frac{q_M}{N_A} \quad (2.25)$$

L'évaluation *ab initio* de G_M consiste à calculer l'énergie potentielle $U_{p,M}$ du système et la correction thermique :

$$\Delta G^{corr} = RT \ln \frac{q_M}{N_A} \quad (2.26)$$

Dans le cas de l'utilisation du logiciel Gaussian 03, cette dernière correction est évaluée lors d'un calcul de fréquences. Pour une espèce solvatée se trouvant dans l'état M, l'enthalpie libre molaire peut être exprimée de la manière suivante :

$$G_{aq}(M) = G_g(M) + \Delta G_{solv}(M) \quad (2.27)$$

où $G_g(M)$ est l'enthalpie libre en phase gazeuse et $\Delta G_{solv}(M)$ l'enthalpie libre de solvation du système. Pour le calcul de $\Delta G_{solv}(M)$ nous avons utilisé le modèle du continuum polarisable (*polarizable continuum model, PCM*) [71]. Le principe de cette méthode est le suivant. La molécule (ou le système), entourée par un milieu diélectrique continu, se trouve dans une cavité moléculaire qui résulte de l'assemblage de cavités atomiques sphériques. Le champ électrique de la molécule (du système) polarise le diélectrique qui produit donc un champ électrique de réaction agissant sur la molécule. La fonction d'onde et l'énergie de la molécule sont calculées en prenant en compte ce champ électrique du solvant. A cette énergie potentielle on ajoute deux contributions semi-empiriques : une contribution de dispersion et une autre due à la formation de la cavité moléculaire. La somme des trois termes est appelée « énergie libre en solution ». L'énergie de solvation est calculée comme la différence entre « l'énergie libre en solution » et l'énergie potentielle de la molécule en phase gazeuse. On peut admettre que la méthode PCM donne des énergies de solvation pertinentes tant qu'il n'existe pas une interaction spécifique entre le soluté et les molécules du solvant.

2.1.5 Théorie de l'état de transition

L'évolution typique d'un système moléculaire au cours d'une réaction chimique peut être représentée par la courbe d'énergie potentielle en fonction de la coordonnée de réaction ($U(q_c)$). On parle de trajectoire de réaction. Au long de la trajectoire de réaction on distingue, en général, quatre états : réactifs séparés RS, réactifs complexés RC, l'état de transition ET, les produits P.

Les réactifs complexés et les produits sont représentés par des minima sur la courbe $U(q_c)$. L'état de transition (ET) est un point en « selle » : son énergie est un minimum par rapport à toutes les coordonnées, sauf q_c pour laquelle elle est un maximum. Dans la théorie de l'état de transition, il est supposé qu'il existe une surface séparant le domaine des réactifs de celui des produits, telle qu'une trajectoire de réaction ne franchit qu'une fois cette surface [72]. Autrement dit, une fois cette surface franchie, le système ne peut plus revenir en arrière, vers les réactifs. Dans cette approche, la constante de vitesse pour

le passage de l'état RC à l'état ET est donnée par l'équation d'Eyring :

$$k_a = \frac{k_B T}{h} e^{-\frac{(G(ET)-G(RC))}{RT}} \quad (2.28)$$

Dans le cas d'une réaction contrôlée par l'activation, la constante de vitesse de réaction globale sera :

$$k = c_0 \frac{k_B T}{h} e^{-\frac{(G(ET)-G(RS))}{RT}} \quad (2.29)$$

Ici c_0 est la concentration molaire standard en solution (1 M). L'expression 2.28 est obtenue en prenant en compte les deux étapes de la réaction : la formation du complexe RC suivie par le franchissement de l'état ET.

2.2 Calculs MM

2.2.1 Simulation de dynamique moléculaire

La minimisation de l'énergie d'un système moléculaire permet de calculer les positions des atomes et des molécules à 0 K. Cette description du système n'est pas suffisante à des températures non-nulles. C'est la méthode de dynamique moléculaire qui tente de simuler le mouvement des atomes et des molécules dans ces cas. Cette simulation est basée sur les lois de Newton de la mécanique classique :

$$\vec{F}_i = m_i \vec{a}_i = m_i \frac{d^2 \vec{r}_i(t)}{dt^2} \quad (2.30)$$

où \vec{F}_i est le vecteur force agissant sur l'atome i ayant la masse m_i soumis à l'accélération \vec{a}_i . La force \vec{F}_i est déterminée par différentiation de la fonction d'énergie potentielle par rapport à la position $\vec{r}_i(t)$:

$$\vec{F}_i = -\frac{dE(\vec{r}_i, \dots, \vec{r}_n)}{d\vec{r}_i(t)} \quad (2.31)$$

L'énergie potentielle E est calculée à partir des paramètres du champ de forces de la mécanique classique.

Afin de générer la trajectoire du système moléculaire, il suffit d'intégrer les équations de mouvement de manière itérative pour une durée choisie. Le pas d'intégration Δt est très faible, en général de 1 à 10 fs. Cet intervalle de temps doit être plus petit que la période du mouvement de la plus haute fréquence qui correspond aux mouvements des atomes d'hydrogène. Afin d'augmenter le pas temporel il est d'usage de restreindre les

degrés de liberté internes du système en fixant les longueurs des liaisons impliquant les atomes d'hydrogène. C'est la méthodologie SHAKE [73, 74] qui sera utilisée.

Le protocole typique de simulation de dynamique moléculaire contient :

- La création de la structure moléculaire de départ
- La minimisation de l'énergie
- Le chauffage du système
- L'équilibration du système à la température d'étude
- La production de la trajectoire
- L'analyse de la trajectoire

Typiquement, avant d'effectuer la simulation de dynamique moléculaire, les systèmes ont été chauffés de 0 K à 300 K pendant 20 ps et équilibrés à la même température pendant encore 20 ps. Les simulations ont été effectuées à température et pression constantes (300 K et 1 atm) dans des boîtes de solvant rectangulaires de type TIP3P [57] avec des conditions périodiques aux frontières. Des contre-ions ont été ajoutés afin de préserver la neutralité du système. Le pas temporel utilisé dans les simulations a été de 2 fs, toutes les liaisons covalentes des atomes d'hydrogène étant fixées avec l'algorithme SHAKE. Les interactions non-liées ont été négligées pour les distances supérieures à 10 Å.

Les calculs de dynamique moléculaire rapportés dans cette travail de thèse ont été effectués avec l'ensemble de logiciels AMBER9 [75].

2.2.2 Champ de forces

AMBER (*Assisted Model Building with Energy Refinement*) contient plusieurs champs de force dont celui implémenté par le fichier *parm99.dat* [76], largement utilisé dans notre travail de recherche. Les interactions entre les atomes consistent en une composante intramoléculaire (déformations harmoniques des liaisons, des angles et des dièdres : les termes liés) et une composante intermoléculaire décrivant les interactions entre les atomes "non

liés" (c'est à dire, séparés par plus de deux liaisons : les termes non-liés). L'énergie totale du système s'écrit alors :

$$E_{tot} = E_{stretching} + E_{bending} + E_{torsion} + E_{non-bonded} \quad (2.32)$$

où $E_{stretching}$ est l'énergie de déformation des liaisons, $E_{bending}$ l'énergie de déformation des angles, $E_{torsion}$ l'énergie de déformation des angles dièdres et $E_{non-bonded}$ l'énergie liée aux interactions non-liées. E_{tot} dépend de $3N$ variables (N étant le nombre d'atomes de la molécule).

Énergies des interactions liées

Les énergies de déformation des liaisons et des angles sont représentées par des termes harmoniques. La somme est effectuée sur toutes les liaisons covalentes ou sur tous les angles formés entre les atomes liés par liaisons "covalentes" :

1. Déformation des liaisons

$\sum_{liaisons} k_b(r - r_0)^2$ où r est la longueur de liaison instantanée, r_0 la longueur de liaison à l'équilibre et k_b la constante de force

2. Déformation des angles

$\sum_{angles} k_\theta(\theta - \theta_0)^2$ où θ est l'angle instantané, θ_0 l'angle à l'équilibre et k_θ la constante de force associée

3. Déformation des dièdres

$\sum_{dièdres} \sum_n V_n [1 + \cos(n\phi - \gamma)]$ où V_n est la valeur de la barrière de rotation d'ordre n , ϕ la phase et γ la valeur instantanée du dièdre

Interactions intermoléculaires non-liées

Toutes ces interactions concernent les atomes qui appartiennent à des molécules différentes ou qui sont séparés par plus de deux liaisons covalentes. Les interactions sont décrites par deux termes : un terme de Lennard-Jones et un terme électrostatique.

1. Le terme de Lennard-Jones décrit un potentiel en 6-12 dont l'expression est la suivante :

$$\sum_{i < j} \left[-2\epsilon_{ij} \left(\frac{R_{ij}^*}{R_{ij}} \right)^6 + \epsilon_{ij} \left(\frac{R_{ij}^*}{R_{ij}} \right)^{12} \right] \quad (2.33)$$

où ϵ_{ij} représente la profondeur du puits de potentiel à la distance d'équilibre R_{ij}^* . Le terme en R_{ij}^{-6} représente la dispersion attractive de London entre deux atomes et celui en R_{ij}^{-12} la répulsion due à l'exclusion de Pauli.

2. Les interactions électrostatiques

Les interactions coulombiennes sont représentées par un potentiel en R_{ij}^{-1} :

$$\sum_{i < j} \frac{q_i q_j}{R_{ij}} \quad (2.34)$$

où q_i et q_j sont les charges partielles des atomes i et j . La distribution des électrons, considérée comme implicite, est localisée sur les liaisons et représentée par des charges partielles atomiques. Dans le cas du champ de force AMBER ces charges ont été ajustées de façon à reproduire les surfaces de potentiel électrostatique obtenues par des calculs *ab initio* d'OM (charges ESP) [77].

2.2.3 Évaluation de l'énergie libre d'un système classique

L'énergie libre d'un système classique s'exprime à l'aide de la fonction de partition configurationnelle Z :

$$F = -k_B T \ln Z \quad (2.35)$$

où

$$Z = \int \dots \int e^{-\beta E(X)} dX \quad (2.36)$$

et $\beta = k_B T$. Dans l'équation 2.36, X signifie l'ensemble des coordonnées et l'intégrale est effectuée dans l'espace de configurations du système. D'autre part, dans l'équation 2.36 la contributions des vitesses à la fonction de partition a été négligée. En effet, pour des processus se déroulant à température constante, dans le calcul des différences d'énergie libre les contributions cinétiques s'annulent réciproquement. La différence d'énergie libre entre deux états du système, 1 et 0, a donc l'expression :

$$F_1 - F_0 = -k_B T \ln \frac{Z_1}{Z_0} \quad (2.37)$$

En pratique, l'intégrale dans l'équation 2.36 est remplacée par une somme sur un ensemble de configurations représentatif du système. Cet ensemble peut être généré par une simulation de dynamique moléculaire. Afin de limiter les erreurs dues à un échantillonnage du système qui est toujours partiel, les fonctions de partition Z_1 et Z_0 intervenant dans l'équation 2.37 devraient être calculées en utilisant le même ensemble de configurations.

C'est l'idée sur laquelle sont basées deux méthodes pratiques de calcul des différences d'énergie libre, la méthode de l'intégration thermodynamique et la méthode des perturbations.

La méthode de l'intégration thermodynamique (TDI)

La méthode de l'intégration thermodynamique peut être appliquée même dans le cas où les deux états, 1 et 0, seraient très différents. Dans cette méthode, le calcul est basé sur des simulations de dynamique moléculaire effectuées pour des systèmes hybrides dont l'énergie est une combinaison des énergies des deux états, 1 et 0. Le mélange des deux états est dosé à l'aide d'un paramètre λ qui peut varier entre 0 et 1. Dans le cas le plus simple d'un mélange linéaire on peut écrire :

$$E(X, \lambda) = (1 - \lambda)E_0(X) + \lambda E_1(X) \quad (2.38)$$

Formellement, la différence d'énergie libre entre les deux états s'exprime alors de la manière suivante :

$$F_1 - F_0 = F(1) - F(0) = \int_0^1 \frac{\partial F(\lambda)}{\partial \lambda} d\lambda \quad (2.39)$$

Les équations 2.35, 2.36 et 2.38 conduisent à la relation :

$$\frac{\partial F(\lambda)}{\partial \lambda} = \left\langle \frac{\partial E(X, \lambda)}{\partial \lambda} \right\rangle_\lambda \quad (2.40)$$

La notation $\langle \rangle_\lambda$ représente une moyenne effectuée sur l'ensemble statistique des configurations d'un système hybride de paramètre λ donné.

Afin d'évaluer par voie numérique l'intégrale intervenant dans l'équation 2.39, plusieurs simulations de dynamique moléculaire sont effectuées pour différentes valeurs de λ situées dans l'intervalle 0-1. La variation de l'énergie libre est alors obtenue en utilisant la relation :

$$F_1 - F_0 = \sum_i \left\langle \frac{\partial F(\lambda)}{\partial \lambda} \right\rangle_{\lambda_i} \Delta \lambda_i \quad (2.41)$$

Dans le cas d'une combinaison linéaire d'énergies (équation 2.38) on peut écrire :

$$\frac{\partial F(\lambda)}{\partial \lambda} = \langle E_1(X) - E_0(X) \rangle_\lambda \quad (2.42)$$

Donc :

$$F_1 - F_0 = \sum_i \langle E_1 - E_0 \rangle_\lambda \Delta \lambda_i \quad (2.43)$$

La méthode des perturbations

En partant de l'équation 2.36 on obtient :

$$\frac{Z_1}{Z_0} = \frac{\int \cdots \int e^{-\beta E_1(X)} dX}{\int \cdots \int e^{-\beta E_0(X)} dX} \quad (2.44)$$

En multipliant l'intégrand du numérateur par le facteur unité $1 = e^{+\beta E_0(X)} e^{-\beta E_0(X)}$ on peut écrire :

$$\frac{Z_1}{Z_0} = \frac{\int \cdots \int e^{-\beta E_1(X)} e^{+\beta E_0(X)} e^{-\beta E_0(X)} dX}{\int \cdots \int e^{-\beta E_0(X)} dX} \quad (2.45)$$

d'où résulte :

$$\frac{Z_1}{Z_0} = \int \cdots \int e^{-\beta [E_1(X) - E_0(X)]} P_0(X) dX \quad (2.46)$$

où P_0 est la fonction de probabilité de Boltzmann. Alors l'équation 2.45 devient :

$$\frac{Z_1}{Z_0} = \langle e^{-\beta \Delta E(X)} \rangle_0 \quad (2.47)$$

Ici l'indice 0 indique que la moyenne est calculée dans l'état initial du système. En conclusion :

$$\Delta F = -k_B T \ln \langle e^{-\beta \Delta E(X)} \rangle_0 \quad (2.48)$$

et

$$\Delta F = -k_B T \ln \langle e^{+\beta \Delta E(X)} \rangle_1 \quad (2.49)$$

Cette approche sera précise seulement si les états 0 et 1 sont très proches l'un par rapport à l'autre. Dans le cas contraire, le calcul peut être effectué en plusieurs pas au long de la coordonnée λ de dosage des deux états. La variation d'énergie libre est obtenue ensuite par l'addition des contributions de chaque pas :

$$\Delta F = \sum_{i=0}^{n-1} \Delta F_i(\lambda_i \rightarrow \lambda_{i+1}) \quad (2.50)$$

2.2.4 Le potentiel de force moyenne (PMF)

Le profil d'énergie (enthalpie) libre au long d'une coordonnée donnée est appelé potentiel de force moyenne. Il présente un intérêt particulier lorsque la coordonnée choisie est la coordonnée de réaction : dans ce cas le profil obtenu peut être vu comme une moyenne effectuée sur tous les chemins de réaction possibles. Cela ne doit pas être confondu avec le chemin de réaction le plus probable, qui correspond à la barrière d'énergie minimum.

Pour être plus précis, on définira une énergie libre partielle qui correspond à une valeur fixe, ξ_c , de la coordonnée de réaction :

$$F(\xi_c) = -k_B T \ln Z(\xi_c) \quad (2.51)$$

Ici $Z(\xi_c)$ à l'expression :

$$Z(\xi_c) = \int \delta(\xi - \xi_c) e^{-\beta E(X)} d\xi dX' \quad (2.52)$$

où $\delta(\xi - \xi_c)$ est la fonction de Dirac et X' signifie l'ensemble des coordonnées du système moins la coordonnée de réaction. L'enthalpie libre partielle sera alors :

$$G(\xi_c) = F(\xi_c) + RT \quad (2.53)$$

Le PMF [78] peut être calculé à partir d'un ensemble statistique de configurations du système généré par une simulation de dynamique moléculaire. Le principe est le suivant. Les fréquences d'apparition, au long de la trajectoire simulée, d'une valeur donnée de la coordonnée choisie sont liées à l'enthalpie libre du sous-espace respectif par la relation :

$$n(\xi_c) \approx e^{-G(\xi_c)/RT} \quad (2.54)$$

Si on considère deux valeurs ξ_1 et ξ_2 de la variable ξ , on peut écrire :

$$\frac{n(\xi_1)}{n(\xi_2)} = e^{-[G(\xi_1) - G(\xi_2)]/RT} \quad (2.55)$$

et :

$$G(\xi_2) - G(\xi_1) = RT \ln \frac{n(\xi_1)}{n(\xi_2)} \quad (2.56)$$

L'équation 2.56 permet le calcul de la différence d'enthalpie libre $\Delta G(\xi)$ par rapport à une valeur de référence $G(\xi_0)$.

L'efficacité de l'échantillonnage peut être augmentée en utilisant une technique appelée *Umbrella Sampling* [79]. Dans ce cas, le système est forcé d'évoluer au voisinage d'une certaine valeur, ξ_c , de la coordonnée de réaction par l'ajout d'une contrainte harmonique :

$$U(\xi) = \frac{1}{2} K (\xi - \xi_c)^2 \quad (2.57)$$

Plusieurs fenêtres d'échantillonnage sont alors générées en modifiant la valeur de la contrainte ξ_c . Les effets des contraintes sont corrigés dans le traitement final des résultats. Dans le présent travail de thèse, ce traitement a été effectué avec le logiciel WHAM (Weighted Histogram Analysis Method) [80, 81].

2.3 Calculs QM/MM

Les méthodes de calcul basées sur la mécanique quantique sont difficilement applicables aux systèmes moléculaires contenant plus de 100 atomes. Dans ce cas, l'utilisation d'une méthode hybride mécanique quantique-mécanique moléculaire est envisageable. Warshel et Levitt [82] ont ouvert la voie vers ce type d'études. La méthode implique la division du système en deux parties : une petite région qui est traitée avec une méthode QM et le reste du système traité avec des méthodes classiques, moins coûteux en temps de calcul (Figure 2.1). La région quantique contient tous les atomes qui sont directement impliqués dans la réaction chimique. Le reste du système, qui subit peu de changements au cours de la réaction, est traité avec un champ de forces de mécanique moléculaire. La connexion entre les deux systèmes, QM et MM, joue un rôle très important dans la précision et l'efficacité de la méthode. Il existe deux approches pour le traitement des liaisons de frontière :

- L'approche utilisant des atomes de saturation ("link-atom") [83, 84]. Des atomes d'hydrogène sont ajoutés pour saturer la valence des atomes QM de frontière. Ces atomes ajoutés n'interagissent pas avec les atomes MM mais ils sont inclus dans le calcul QM.
- L'approche dite "boundary-region" (pseudo-liaison, atomes de connexion, orbitales hybrides gelées). Il s'agit d'une région située entre le système QM et le système MM et qui sera prise en compte dans les deux calculs.

Pour le calcul d'énergie du système, il existe deux schémas QM/MM : un schéma additif et un autre soustractif.

Le schéma additif

L'énergie totale du système est obtenue par l'addition des énergies QM et MM en appliquant aussi des corrections et des termes mixtes.

Dans l'approche du type "atomes de saturation" l'énergie totale s'exprime de la façon suivante :

$$E_{QM/MM}^T = E_{MM}^{R-M} + E_{QM}^{M,L} + E_{QM-MM}^{M,R-M} + E_{corr}^{M,L} \quad (2.58)$$

où E_{MM}^{R-M} est l'énergie du système extérieur calculée en MM, $E_{QM}^{M,L}$ l'énergie du système modèle et des atomes de saturation calculée en QM, $E_{QM-MM}^{M,R-M}$ le terme mixte d'interaction entre le système QM et le système MM. $E_{corr}^{M,L}$ représente le terme de correction qui réduit

la dépendance d'énergie totale par rapport aux atomes de saturation.

Dans l'approche du type "boundary-region" on retrouve une énergie totale de la forme :

$$E_{QM/MM}^T = E_{MM}^{R-M,B} + E_{QM}^{M,B} + E_{QM-MM}^{M,R-M,B} \quad (2.59)$$

Le schéma soustractif

Ce schéma est appliqué en général dans le cas de l'approche du type "atomes de saturation". L'énergie totale peut être écrite :

$$E_{QM/MM}^T = E_{MM}^R + E_{QM}^{M,L} - E_{MM}^{M,L} \quad (2.60)$$

L'énergie du système réel est calculée au niveau MM (E_{MM}^R) et celle du système modèle et des atomes de saturation au niveau QM ($E_{QM}^{M,L}$). L'énergie du système modèle calculée au niveau MM ($E_{MM}^{M,L}$) est ensuite soustraite afin d'éviter le double comptage des atomes.

Le couplage électrostatique entre le système modèle et son environnement peut être traité selon les schémas suivants [51, 53, 85] :

- Assemblage ("embedding") mécanique
L'énergie du système modèle est calculée au niveau QM sans tenir compte de la présence des charges extérieures. Les interactions électrostatiques entre le système modèle et le système extérieur sont traitées seulement au niveau MM.
- Assemblage ("embedding") électrostatique
Le calcul QM sur le système modèle est effectué en incluant les interactions électrostatiques avec les charges extérieures. Ces charges sont les charges partielles des atomes appartenant au système extérieur.

Dans ce travail nous avons utilisé la méthode ONIOM[86–88] qui est l'implémentation du schéma IMOMO (*Integrated Molecular Orbital + Molecular Orbital*)[89, 90] dans le logiciel Gaussian03 [70]. Le calcul d'énergie est fait selon un schéma QM/MM soustractif. La frontière entre le système modèle et le système extérieur est représentée par des atomes de saturation. Donc :

$$E^{ONIOM} = E^{QM,mod} + E^{MM,real} - E^{MM,mod} \quad (2.61)$$

Le système modèle (Figure 2.1) a été traité au niveau DFT et le système extérieur en mécanique moléculaire avec le champ de force AMBER.

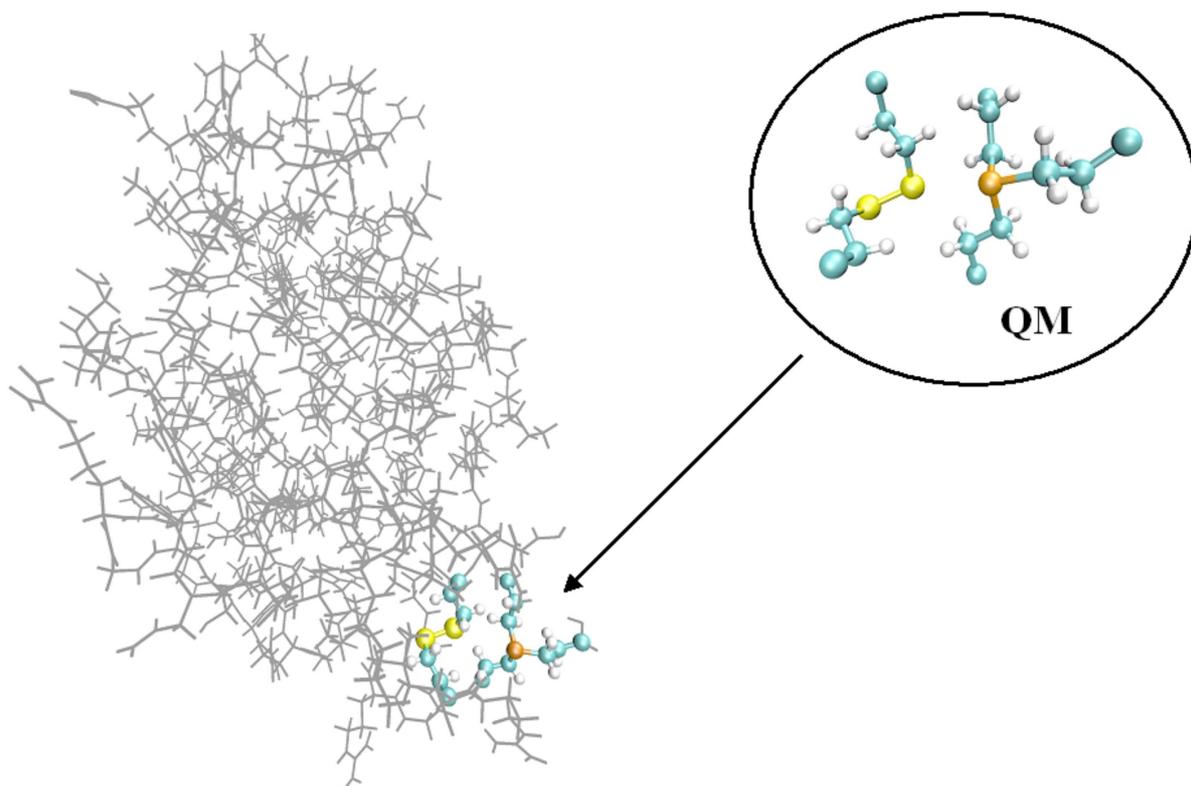


FIGURE 2.1 – Exemple de traitement QM/MM sur le système lysozyme+TCEP : le système modèle est représenté par des billes et le reste du système par des traits.

Chapitre 3

Attribution des bandes S-S dans les spectres Raman des protéines

Sommaire

3.1	Analyse des bandes Raman S-S de systèmes moléculaires simples	54
3.2	Analyse des bandes Raman S-S des albumines sériques	59
3.3	Analyse des bandes Raman S-S de la lysozyme	63
3.4	Les bandes S-S dans les spectres Raman d'autres protéines	66
3.5	Simulation des bandes Raman des quatre ponts S-S de la lysozyme	69

Afin d'effectuer une analyse quantitative de la réduction des ponts S-S des protéines basée sur la détection des bandes Raman, il est nécessaire d'identifier correctement ces bandes. Généralement, dans la littérature, l'attribution des bandes S-S est basée presque exclusivement sur les valeurs des fréquences de ces bandes. Des études expérimentales [37, 91–99] et théoriques [38, 99] indiquent que les bandes Raman appartenant à l'élongation S-S ont des fréquences situées entre 500 et 560 cm^{-1} . Les différences entre les valeurs de ces fréquences sont couramment attribuées aux différentes conformations des angles dièdres $-C_\alpha - C_\beta - S - S-$ [30, 38, 91, 92]. Cependant, il est connu que même dans le cas des ponts S-S des molécules modèles (cystine [100], GSSG [98]) la bande S-S a une structure complexe. Cela suggère que l'environnement joue un rôle significatif dans la forme et la position des bandes S-S [19, 34]. De plus, le spectre Raman des protéines étant riche, il est très probable que dans l'intervalle spectral caractéristique des bandes S-S, se trouvent des bandes ayant une autre nature [101]. Ces éléments nous amènent à supposer qu'une analyse des bandes S-S qui prend en compte seulement les conformations possibles des ponts autour de la liaison $C_\beta - S$ est trop simpliste. Afin d'avancer dans cette analyse nous avons fait appel à trois méthodes :

1. le suivi des modifications des bandes des protéines qui accompagnent la réduction des ponts S-S
2. la mesure du degré de dépolarisation des bandes Raman
3. la simulation par modélisation moléculaire des spectres Raman des ponts S-S modèles dans l'intervalle 500-560 cm^{-1}

Ces méthodes permettent en principe d'apporter des contributions intéressantes à l'analyse des bandes, mais il convient d'abord de discuter leurs limitations. Ainsi, la première méthode est basée sur l'idée simple que, lorsqu'on réduit les ponts S-S, les seules bandes se trouvant dans l'intervalle spectral 500-560 cm^{-1} qui vont disparaître sont celles appartenant à l'élongation S-S. Cependant, en spectroscopie Raman les choses sont un peu plus complexes. Supposons qu'il existe un mode normal de vibration qui contient une contribution, même mineure, de l'élongation S-S. Supposons aussi que toutes les autres coordonnées internes participant à ce mode ne contribuent pas à son activité Raman. Cela veut dire qu'à la différence de l'élongation S-S, ces coordonnées ne réalisent pas de couplage significatif entre deux états excités. Par conséquent, ce mode va disparaître dans le spectre Raman après la coupure de la liaison S-S. La seule conclusion que l'on peut tirer d'une telle disparition est que le mode respectif contient une contribution d'une élongation

S-S et qu'il doit son activité Raman seulement à cette contribution.

En ce qui concerne l'utilisation des mesures de degré de dépolarisation, il a été précisé dans le deuxième chapitre que les modes totalement symétriques ont un degré de dépolarisation inférieur à $3/4$. Par ailleurs, un mode qui contiendrait une contribution largement dominante de l'élongation S-S serait en première approximation un mode totalement symétrique, facile donc à identifier par une mesure de degré de dépolarisation. Mais dans le même chapitre 2 nous avons montré que dans le cas des molécules à faible symétrie, il existe des modes qui ne sont pas totalement symétriques mais qui ont un faible degré de dépolarisation : ce sont, par exemple, les modes qui sont activés Raman par un seul couplage entre deux états excités de type σ . Cette réalité ressort directement si on compare les deux spectres Raman de la lysozyme mesurés, un en polarisation parallèle, l'autre en polarisation perpendiculaire (par rapport à la polarisation du faisceau incident)(Figure 3.1)

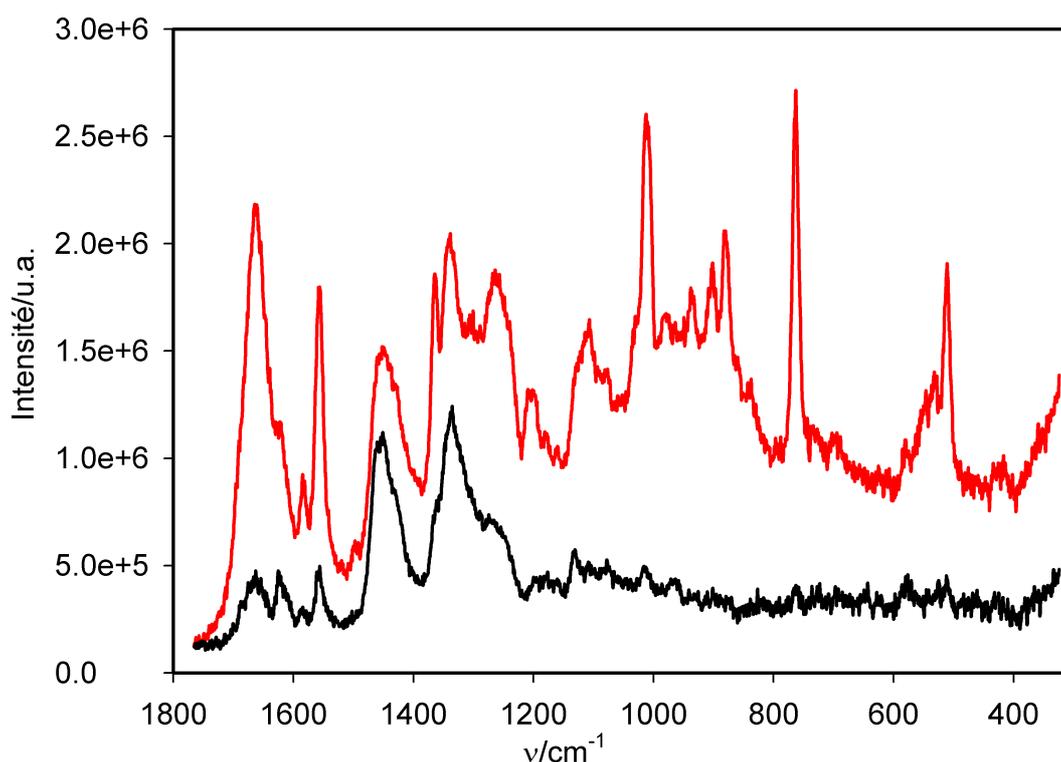


FIGURE 3.1 – Spectre Raman de la lysozyme en poudre mesuré en polarisation parallèle (rouge) et perpendiculaire (noir).

On constate alors que la plupart des bandes Raman de ce spectre ont un très faible degré de dépolarisation. Évidemment, toutes n'appartiennent pas à des modes totalement symétriques, mais elles peuvent appartenir aussi à des modes activés Raman par un seul couplage $\sigma - \sigma$. Ce couplage est très probablement réalisé par une élongation de liaison.

Une analyse élémentaire de l'équation 1.45 nous montre que des modes activés Raman par des couplages entre les états excités du type n peuvent aussi présenter des faibles degrés de dépolarisation qui sont faibles.

3.1 Analyse des bandes Raman S-S de systèmes moléculaires simples

La cystine

Formée par la réaction d'oxydation de deux cysteines, la cystine représente le plus simple modèle de pont S-S ayant une importance biologique.

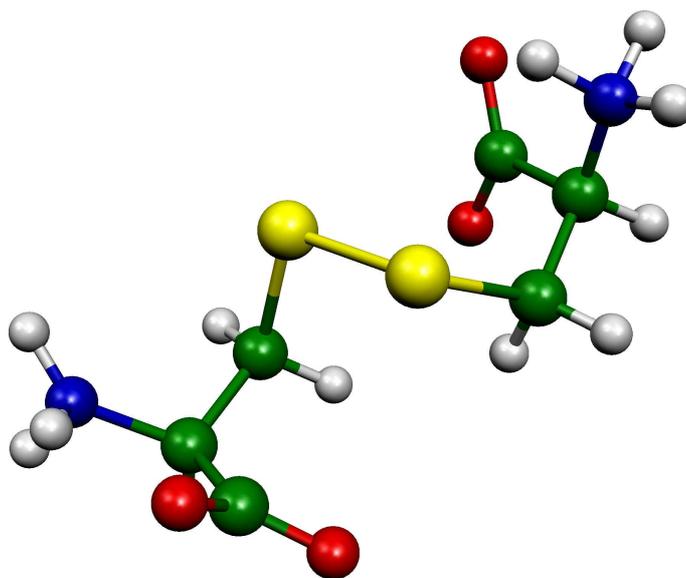


FIGURE 3.2 – La L-cystine.

En solution aqueuse, sa bande Raman S-S présente une forme asymétrique. Elle a été décomposée en utilisant deux fonctions de Lorentz (Figure 3.3). Contrairement à ce résultat, la bande Raman S-S de la cystine à l'état solide (poudre) présente une seule composante (Figure 3.4). Les deux bandes S-S mesurées en solution aqueuse sont plus larges, 25 cm^{-1} de largeur à mi-hauteur, par rapport à la bande mesurée à l'état solide qui ne fait que 11 cm^{-1} de largeur à mi-hauteur. Elles sont déplacées vers des hautes fréquences, 511 et 524 cm^{-1} respectivement, à comparer avec 502 cm^{-1} à l'état solide. On peut interpréter

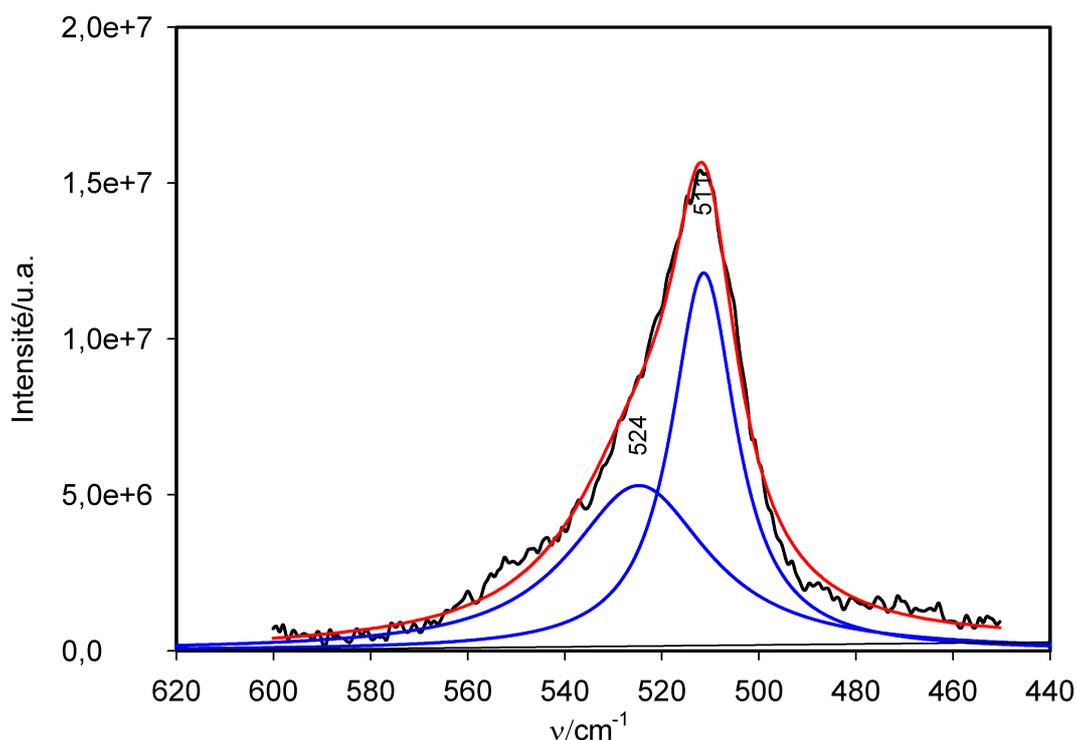


FIGURE 3.3 – Bande Raman S-S de la cystine mesurée en solution aqueuse à une concentration de 16 mM et modélisée par deux fonctions de Lorentz.

l'apparition d'une deuxième bande en solution aqueuse par l'existence de deux rotamères de la cystine : un du type *gg* (ce qui signifie que les deux configurations autour des liaisons $C_{\beta} - S$ et $C'_{\beta} - S'$ sont du type *cis*) et un autre du type *gt* (où une de ces configurations serait du type *trans*). Cependant le déplacement de la bande principale de 502 cm^{-1} (à l'état solide) à 511 cm^{-1} (en solution) ne peut évidemment pas être expliqué de la même manière. On peut attribuer ce déplacement à un changement de couplage entre les coordonnées internes de la cystine ou entre ces coordonnées et celles du solvant. En effet, en solution aqueuse la cystine est plus flexible ce qui rend le couplage des coordonnées internes plus efficace. Ce couplage renforcé pourrait aussi expliquer l'élargissement des bandes.

Le DTT oxydé

Le DTT (dithiothréitol) oxydé est une molécule cyclique refermée par un pont S-S (Figure 3.5). Sa bande Raman S-S mesurée à l'état solide (poudre) a été modélisée par une seule fonction de Lorentz centrée sur 516 cm^{-1} et ayant une largeur à mi-hauteur de 20 cm^{-1} (Figure 3.6). La structure de la même bande mesurée en solution aqueuse

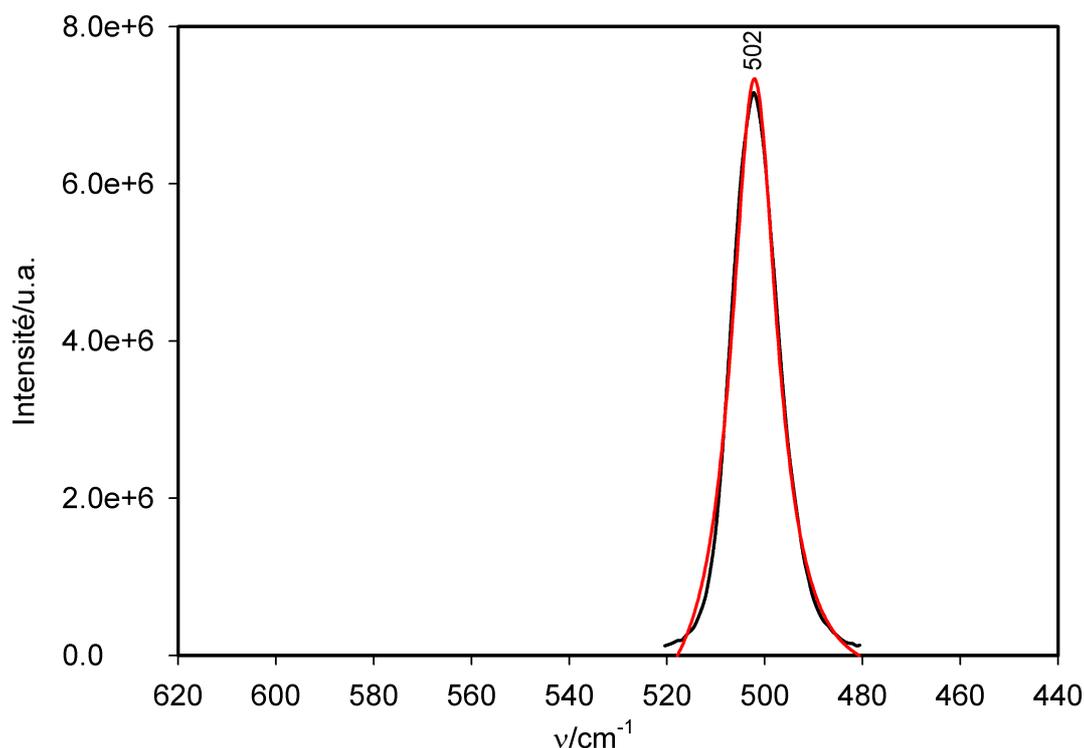


FIGURE 3.4 – Bande Raman S-S de la cystine mesurée en poudre et modélisée par une seule fonction de Lorentz (courbe rouge).

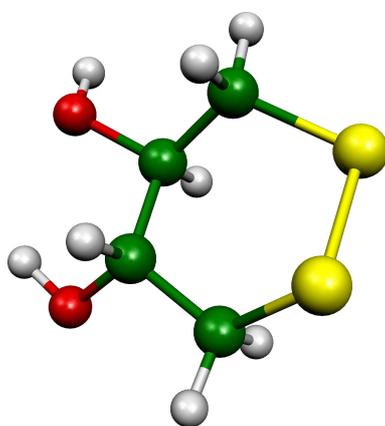


FIGURE 3.5 – DTT oxydé.

est complexe : elle a été modélisée par une somme de trois fonctions Lorentz centrées sur 509, 517 et 522 cm^{-1} et ayant la même largeur à la mi-hauteur, 11 cm^{-1} (Figure 3.7). Dans le cas du DTT, les contraintes structurales ne permettent pas l'existence des conformations *gt* ou *tt*. Cela indique que la complexité de la structure de la bande S-S en solution aqueuse doit être attribuée soit au couplage des coordonnées internes, soit à l'existence des conformères autres que ceux constitués par la rotation autour des liaisons *C-S*.

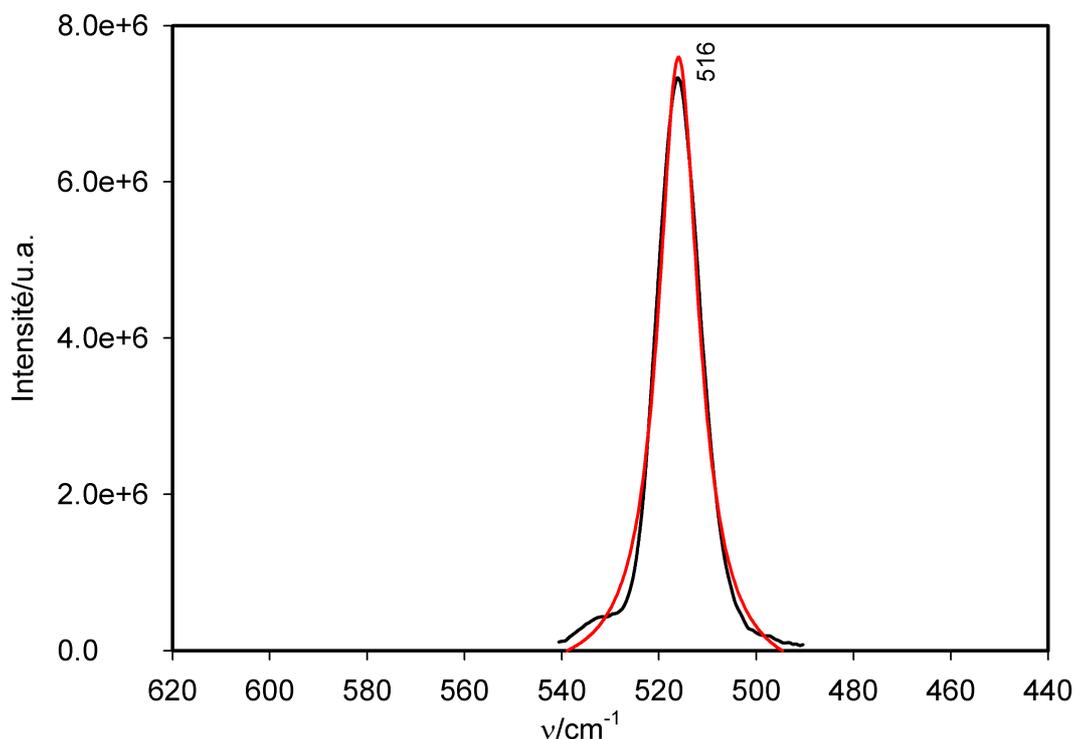


FIGURE 3.6 – Bande Raman S-S du DTT oxydé mesurée en poudre et modélisée par une seule fonction de Lorentz (courbe rouge).

La flexibilité de la structure moléculaire en solution aqueuse favoriserait les deux facteurs.

Le glutathion oxydé

Un modèle de ponts S-S encore plus proche de celui des ponts S-S dans les protéines est représenté par le glutathion oxydé (GSSG) (Figure 3.8). Les bandes S-S du GSSG en solution aqueuse et à l'état cristallin (poudre) sont remarquablement similaires (Figures 3.9 et 3.10). Dans les deux cas, le spectre expérimental peut être représenté par une superposition de deux lorentziennes, l'une centrée à 514 cm^{-1} , l'autre autour de 530 cm^{-1} . Les largeurs à mi-hauteur de ces bandes sont d'environ 20 et 40 cm^{-1} respectivement. L'attribution de ces deux bandes aux rotamères autour des liaisons $C_{\beta} - S$ pose problème parce qu'il est difficile de croire que le changement d'environnement au passage entre l'état solide et l'état aqueuse n'entraîne pas des changements significatifs dans la stabilité de ces rotamères. Une explication alternative serait que dans le cas du GSSG, on ait à faire à deux modes de vibrations distincts pour une même conformation, les deux modes possédant une participation significative de l'élongation S-S.

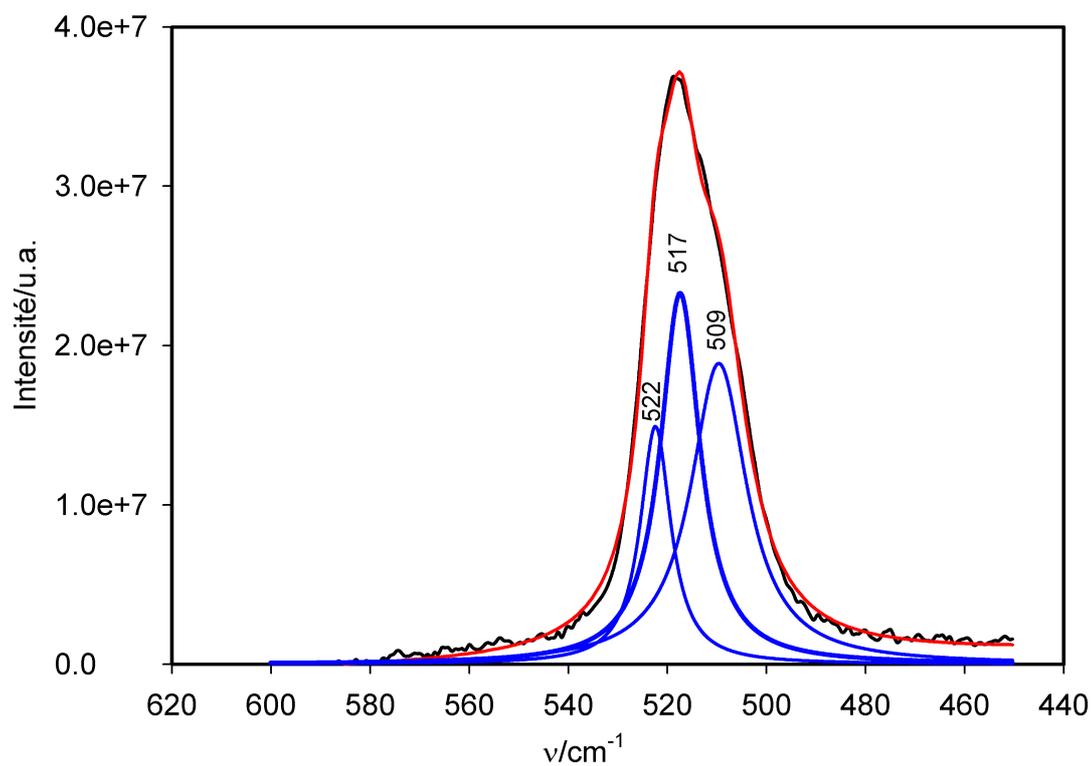


FIGURE 3.7 – Bande Raman S-S du DTT oxydé mesurée en solution aqueuse à une concentration de 16 mM et modélisée par une somme de trois fonctions de Lorentz.

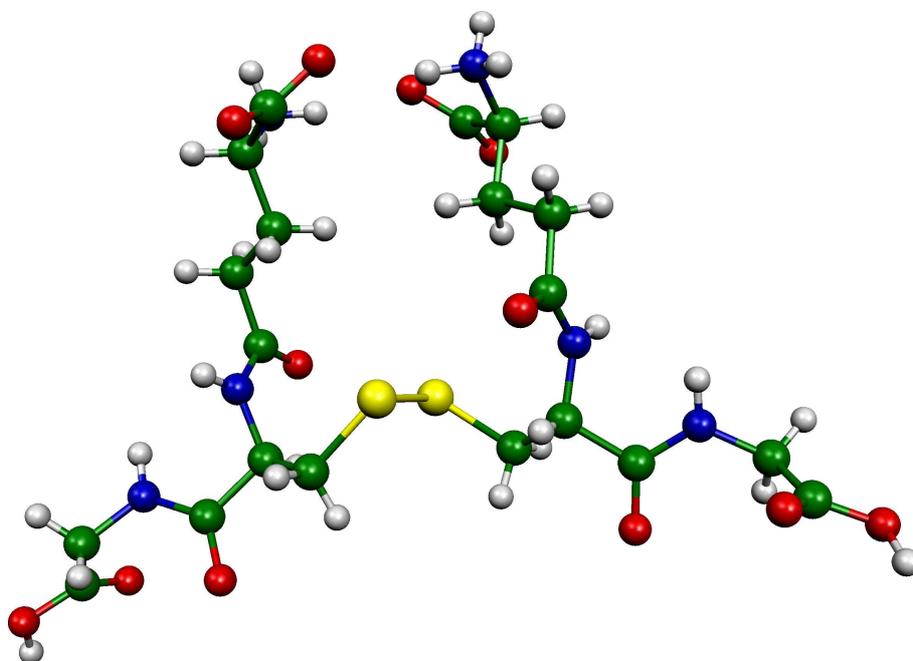


FIGURE 3.8 – Le glutathion oxydé.

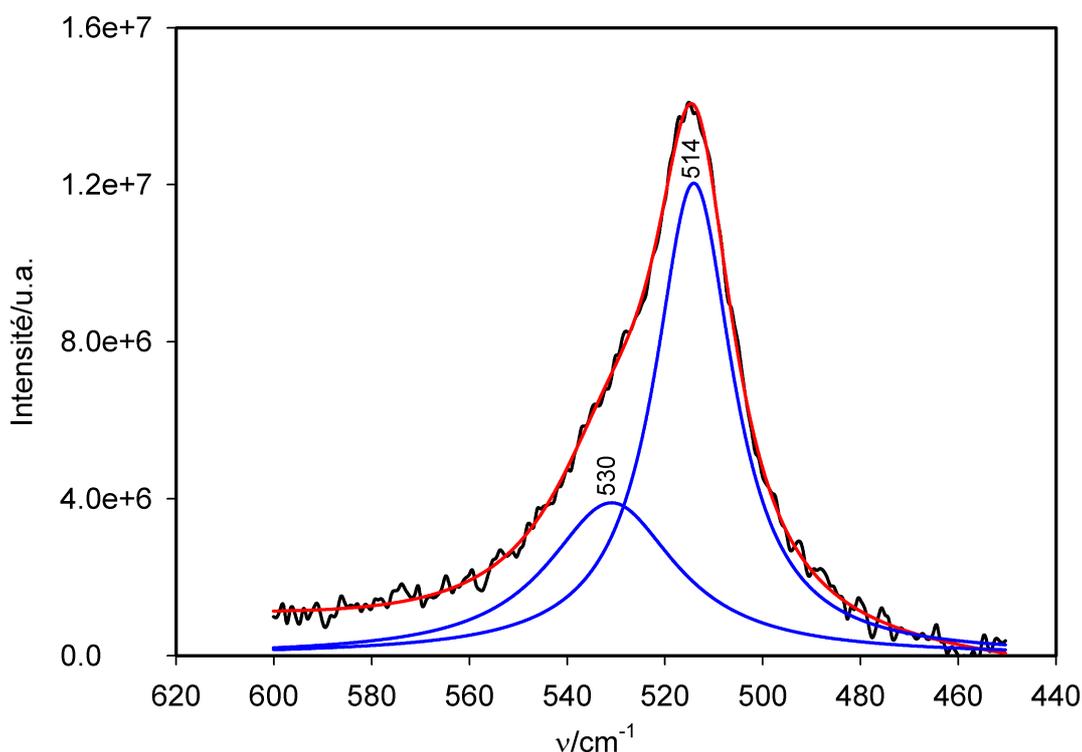


FIGURE 3.9 – Bande Raman S-S du GSSG mesurée en solution aqueuse à une concentration de 16 mM et modélisée par une somme de deux fonctions de Lorentz.

L'analyse des bandes S-S des systèmes moléculaires simples montre que, même dans ces cas, le modèle des rotamères est loin de fournir une explication suffisante de la structure de ces bandes. Nous avons proposé ici quelques facteurs alternatifs qui pourraient jouer un rôle dans la structure des bandes. Cependant, la validation de ces hypothèses exige des études structurales détaillées sur ces systèmes modèles qui restent hors de l'objectif principal de ce travail de thèse.

3.2 Analyse des bandes Raman S-S des albumines sériques

L'albumine sérique bovine (BSA) et l'albumine sérique humaine (HSA) représentent des modèles intéressants dans l'étude du mécanisme de réduction des ponts S-S dans les protéines, en raison de leur contenu élevé en ponts S-S. Constituants organiques du plasma sanguin dans lequel elles représentent 55% des protéines totales, les albumines sériques

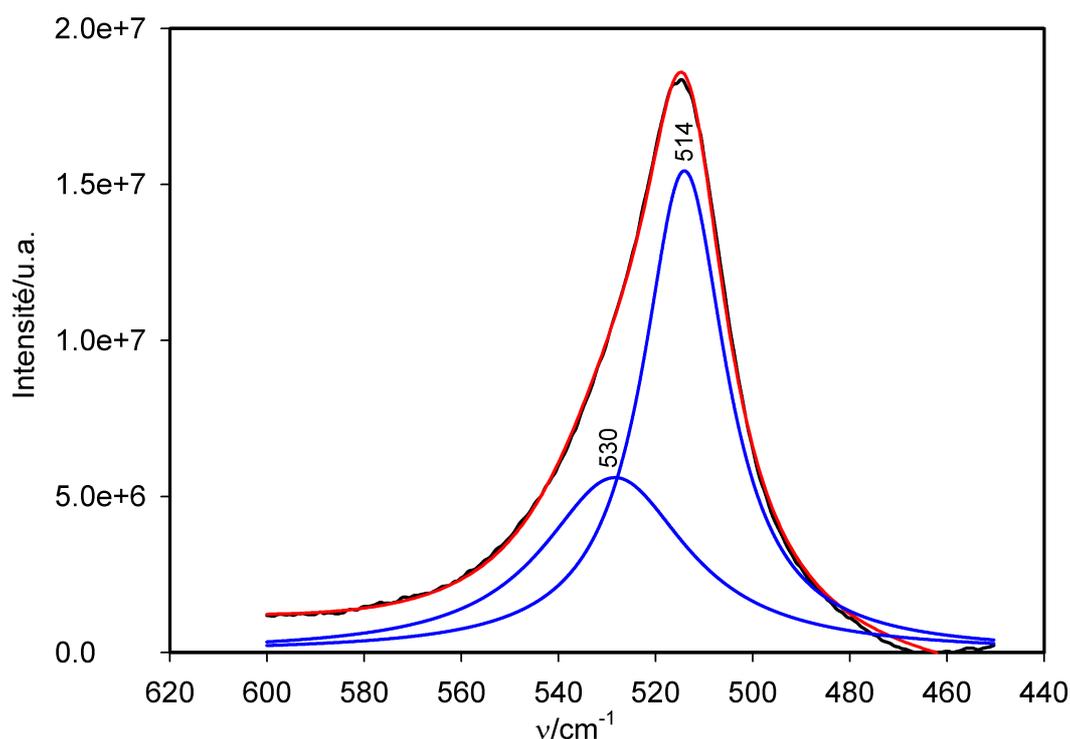


FIGURE 3.10 – Bande Raman S-S du GSSG mesurée en poudre et modélisée par une somme de deux fonctions de Lorentz.

ont pour fonction principale de contribuer à maintenir la pression osmotique à un niveau convenable dans les vaisseaux sanguins vis-à-vis des tissus. La structure secondaire de ces protéines est en grande partie hélicoïdale (70%). Toutefois, une structure tertiaire a été déterminée par la méthode de diffraction des rayons X seulement pour la HSA [102]. La comparaison des séquences des acides aminés de ces deux albumines sériques montre une remarquable homologie (environ 80%). La BSA et la HSA comportent environ 585 acides aminés et un poids moléculaire de 66,4 kDa chacune. Le nombre des ponts S-S inclus dans la structure protéique est de 17.

La structure de diffraction de rayons X de la HSA [102] montre que parmi ses 17 ponts S-S, un seul pont S-S pourrait être considéré comme se trouvant en configuration *gt*. À notre connaissance, il n'existe pas de résultats expérimentaux fiables sur la conformation de ces ponts en solution aqueuse. Le spectre Raman de la HSA en solution aqueuse mesuré dans l'intervalle 450-600 cm^{-1} est présenté sur la Figure 3.11. Ce spectre complexe a été décomposé en utilisant trois lorentziennes centrées respectivement sur 507, 524 et 562 cm^{-1} . Ces trois bandes disparaissent lors de la réduction des ponts S-S par la TCEP (Figure 3.12). En outre, toutes les trois ont de très faibles facteurs de dépolarisation (Figure 3.13).

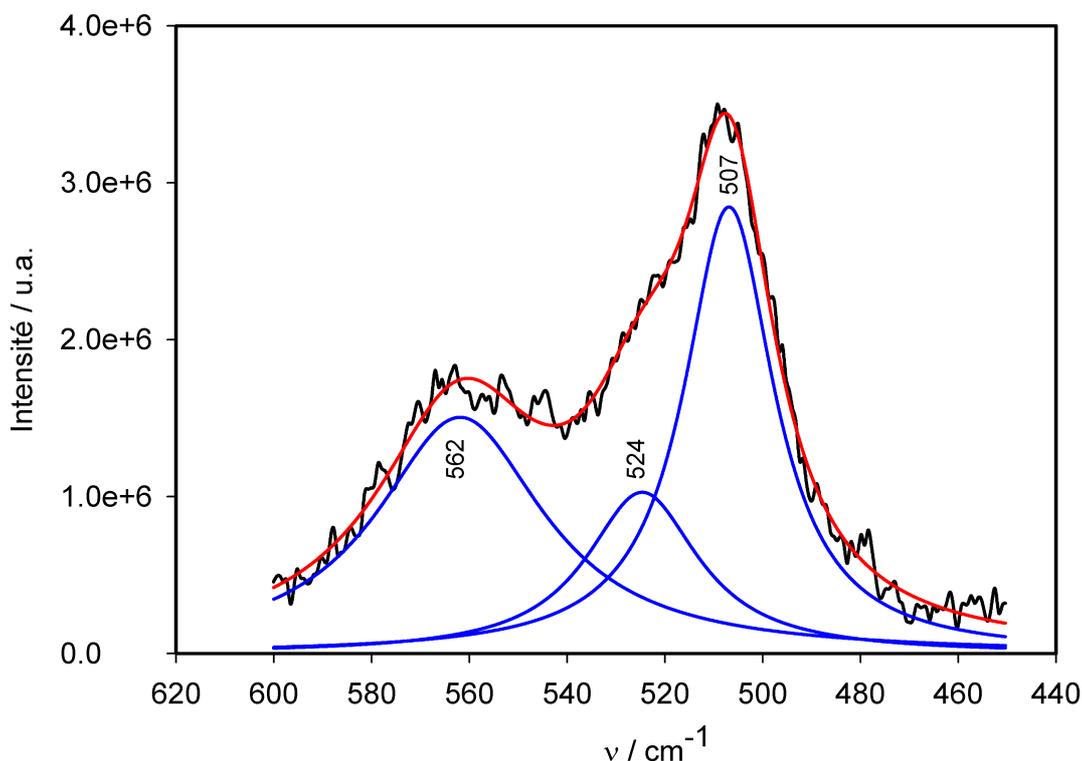


FIGURE 3.11 – Spectre Raman d’une solution aqueuse de HSA native à 1 mM et pH=7,0.

Étant donné la grande ressemblance entre les deux molécules, il n’est pas étonnant que les spectres Raman de la BSA (Figures 3.14 et 3.15) soient très similaires à ceux de la HSA.

Ces résultats nous permettraient d’attribuer à la vibration S-S les trois bandes dans le spectre Raman des albumines sériques, ayant leurs maxima à 507, 524 et 562 cm^{-1} respectivement. Alternativement, la bande centrée sur 562 cm^{-1} dont la fréquence est un peu trop élevée pourrait être attribuée à un mode (ou un groupe de modes) qui ne contient qu’une contribution S-S minoritaire, mais qui est activé Raman par cette contribution. Il faut préciser que jusqu’à présent, dans la littérature, on a attribué à la vibration S-S seulement les deux bandes situées aux fréquences plus basses [91, 92]. La raison réside dans la difficulté d’interprétation de la bande à 562 cm^{-1} . En effet, on peut imaginer que la bande à 507 cm^{-1} est formée par la contribution des ponts en conformation *gg* tandis que celle à 524 cm^{-1} est formée par la contribution des ponts en conformation *gt*. Il faudrait aussi admettre qu’une partie des ponts qui sont du type *gg* à l’état cristallin changent de conformation en solution aqueuse. En suivant le même raisonnement, la bande à 562 cm^{-1} devrait être attribuée à des ponts en configuration *tt*. Or il est difficile d’admettre qu’une fraction si importante des ponts change de conformation en solution aqueuse pour

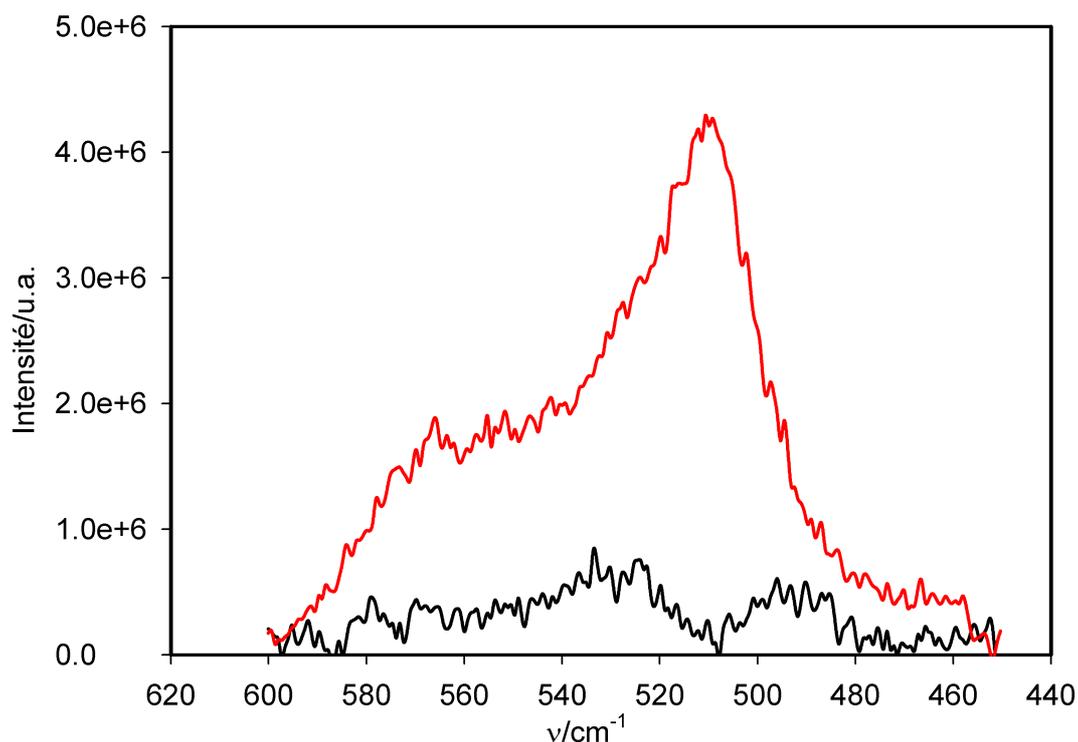


FIGURE 3.12 – Spectre Raman d’une solution aqueuse de HSA native à 0,5 mM (rouge) et en présence de TCEP à 20 mM (noir) ; les mesures ont été effectuées à pH=2,0.

devenir des ponts *tt*.

Il a été rapporté que la forme des bandes Raman S-S de la BSA est affectée par le changement du pH et que l’évolution des bandes S-S de la BSA en fonction de pH pourrait être expliquée dans le cadre du modèle des rotamères autour des liaisons $C_{\beta} - S$ [92]. Ces changements seraient plus nets pour les deux bandes à fréquences plus basses et ils seraient liés aux transitions conformationnelles de la BSA : la transition $N \rightarrow F$, qui se produit dans l’intervalle de pH 4,40-3,75 et la transition $F \rightarrow E$ qui a lieu dans l’intervalle de pH 3,6-2,8 [92, 103]. Nos mesures des spectres Raman de la BSA aux différentes valeurs de pH infirment clairement ces résultats : nous n’avons constaté aucune modification des bandes Raman S-S de la BSA lors du changement de pH (Figure 3.16). On peut aussi se demander si les bandes à 507 et 524 cm^{-1} ont une signification physique quelconque ou bien elles ne représentent qu’une décomposition conventionnelle d’une distribution quasi-continue des contributions venant des 17 ponts S-S des albumines sériques. Cette position nous semble plus raisonnable.

Dans l’étude des cinétiques expérimentales présentée dans le chapitre suivant, nous avons attribué les bandes à 507 et 524 cm^{-1} à l’élongation S-S tandis que la bande à 562

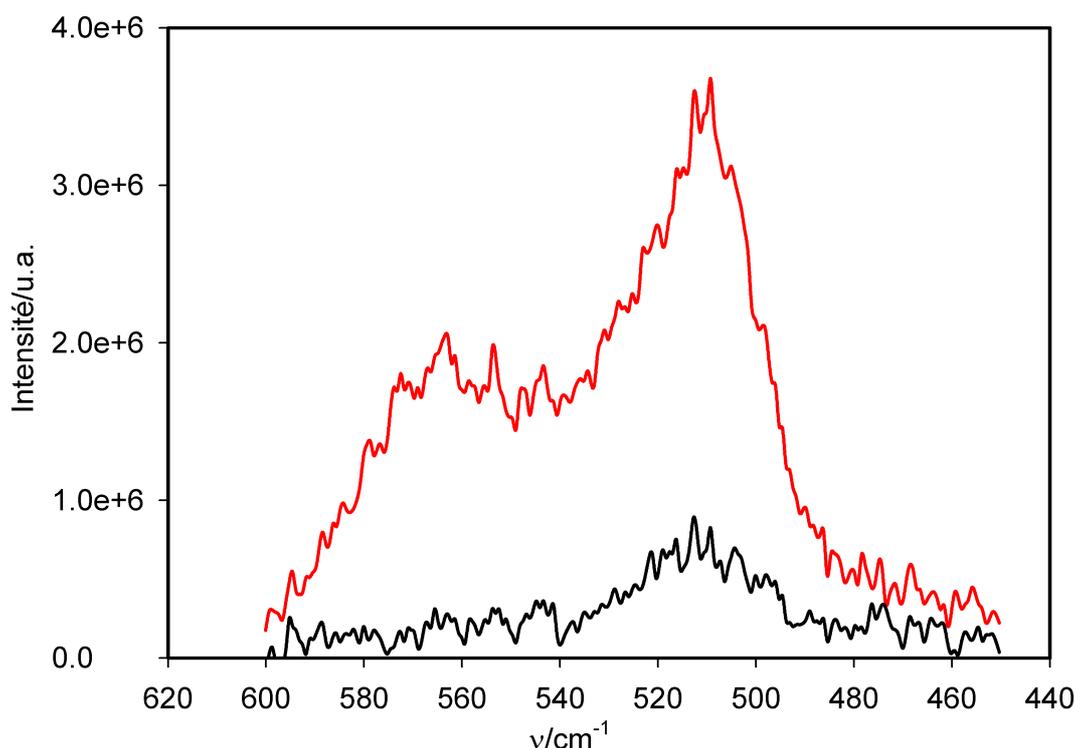


FIGURE 3.13 – Spectre Raman d’une solution aqueuse de HSA native à 1 mM et pH=7,0 mesurée en polarisation parallèle (rouge) et en polarisation perpendiculaire (noir).

cm^{-1} a été attribuée à un mode contenant une contribution minoritaire de l’élongation S-S.

3.3 Analyse des bandes Raman S-S de la lysozyme

L’intérêt présenté par la lysozyme dans l’étude du mécanisme de réduction des ponts S-S dans les protéines vient du fait que sa structure et ses propriétés sont très bien connues [104]. La lysozyme provenant du blanc d’oeuf (HWEL) est constitué d’une seule chaîne de 129 acides aminés formant sept hélices alpha, trois feuillets β antiparallèles et plusieurs régions désordonnées. Étant une protéine avec une structure globulaire, sa forme est ellipsoïdale et cache le site actif dans une grande "poche". La molécule contient quatre ponts S-S dont un est plus exposé au solvant que les trois autres. Parmi ces quatre ponts, trois se trouvent en conformation *gg* et un en conformation *gt*.

Le spectre Raman expérimental de la lysozyme, dans l’intervalle spécifique de la vibration S-S est présenté sur la Figure 3.17. Nous l’avons modélisé par une somme de cinq

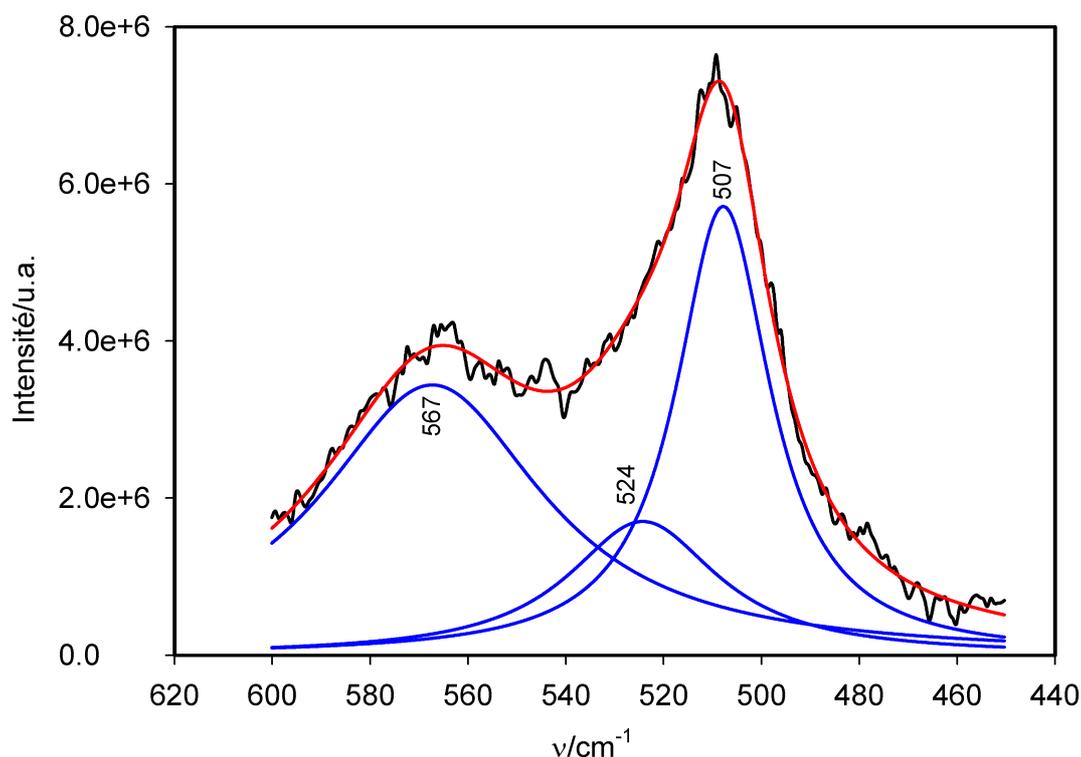


FIGURE 3.14 – Spectre Raman d’une solution aqueuse de BSA native à 1 mM et pH=7,0.

lorentziennes. Compte tenu des fréquences, trois de ces bandes pourraient être attribuées à la vibration S-S : celle à 509 cm^{-1} , celle à 528 cm^{-1} et celle à 545 cm^{-1} . La réduction partielle des ponts S-S par la TCEP (Figure 3.18) fait diminuer ces trois bandes dans la même proportion. Les mesures de spectres Raman polarisés mettent en évidence un très faible degré de dépolarisation de ces trois bandes (Figure 3.19).

Dans le modèle des rotamères autour des liaisons $C_{\beta} - S$, la bande à 509 cm^{-1} serait attribuée aux trois ponts en conformation *gg* et celle à 528 cm^{-1} au pont en configuration *gt* [105]. Nous remarquons que le rapport des intensités des deux bandes est cohérent avec cette attribution. En ce qui concerne la bande à 545 cm^{-1} , trois alternatives sont possibles :

1. l’attribuer à la vibration S-S et admettre qu’en solution aqueuse une fraction importante des ponts S-S de la lysozyme seraient en conformation *tt*. Cette alternative est à exclure parce que la bande à 545 cm^{-1} est présente aussi dans le spectre Raman de la lysozyme à l’état cristallin [105] qui ne contient pas des ponts S-S en configuration *tt*.

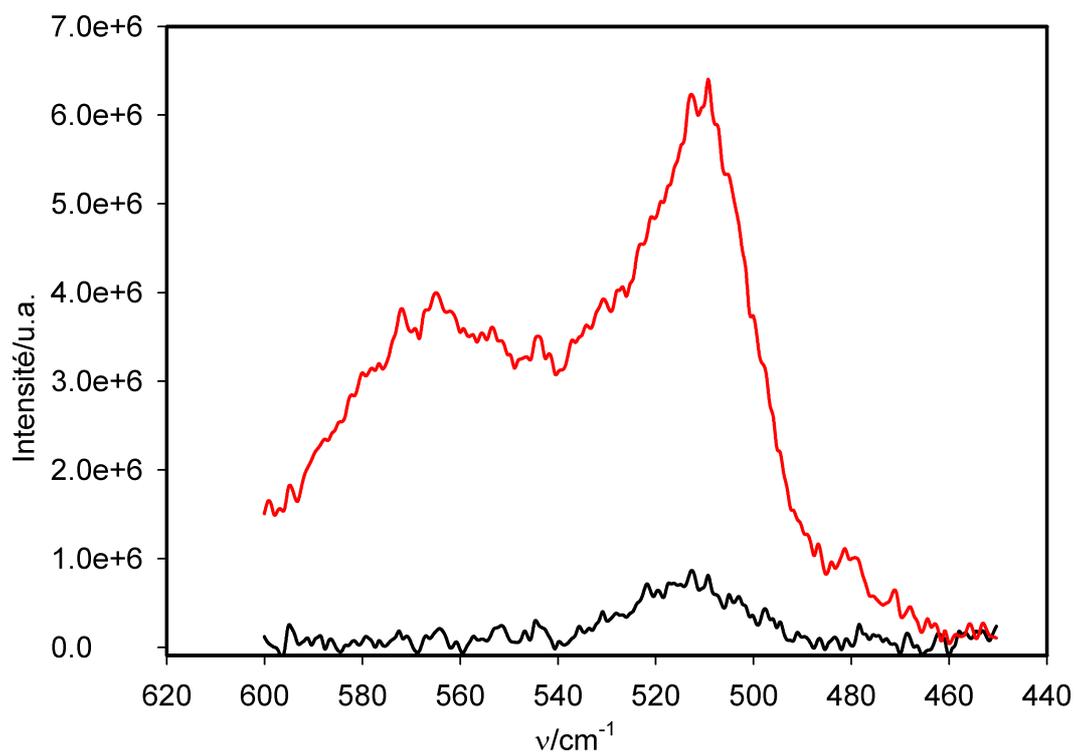


FIGURE 3.15 – Spectre Raman d’une solution aqueuse de BSA native à 1 mM et pH=7,0 mesurée en polarisation parallèle (rouge) et en polarisation perpendiculaire (noir).

2. l’attribuer toujours à la vibration S-S mais considérer en même temps que le déplacement vers les hautes fréquences de cette bande est dû à d’autres facteurs que le changement de conformation des ponts.
3. attribuer cette bande à un mode qui n’est pas à dominance S-S, mais qui est actif Raman seulement grâce à une contribution de l’elongation S-S. Cette interprétation nous paraît la plus plausible.

Il est évident que plus des travaux supplémentaires sont nécessaires pour trancher entre ces alternatives. Les études de cinétique de réduction des ponts S-S (voir le chapitre suivant) montrent que les cinétiques des trois bandes sont quasiment identiques, ce qui rend encore plus difficile leur analyse en terme de contributions venant des différents ponts S-S.

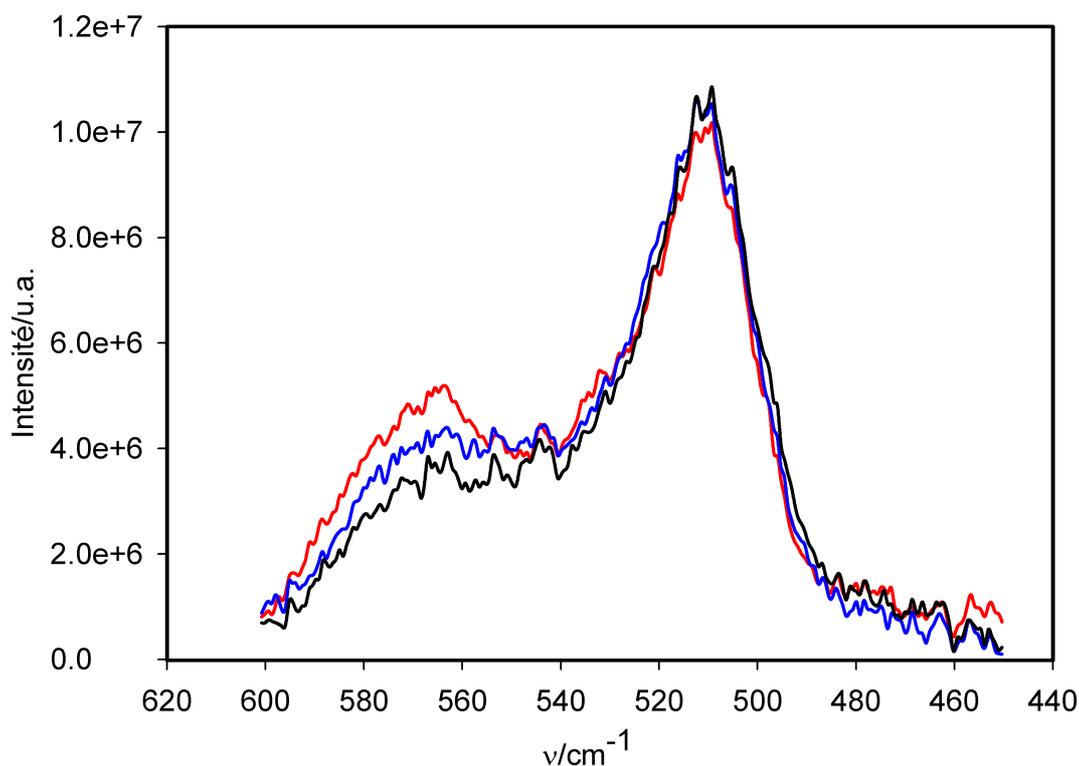


FIGURE 3.16 – Spectres Raman d’une solution aqueuse de BSA native à 1 mM et pH=7,0 (rouge), pH=3,6 (bleu), pH=2,6 (noir).

3.4 Les bandes S-S dans les spectres Raman d’autres protéines

Dans la littérature on trouve très peu d’analyses détaillées des bandes Raman S-S des protéines. Afin d’élargir notre base d’analyse, nous avons entrepris de mesurer les bandes Raman S-S de trois autres protéines. Dans tous les cas, nous avons établi avec soin la ligne de base des spectres et nous avons essayé de modéliser les bandes par des sommes de lorentziennes.

La bêta-lactoglobuline

La bêta-lactoglobuline (BLG) est une protéine globulaire qui contient 162 acides aminés et qui enveloppe dans sa structure deux ponts S-S. Le spectre Raman de cette protéine présente une bande très nette à 509 cm^{-1} (Figure 3.20) qui a pu être modélisée par une seule lorentzienne d’une largeur à mi-hauteur de 15 cm^{-1} . Une autre bande, cette fois-ci

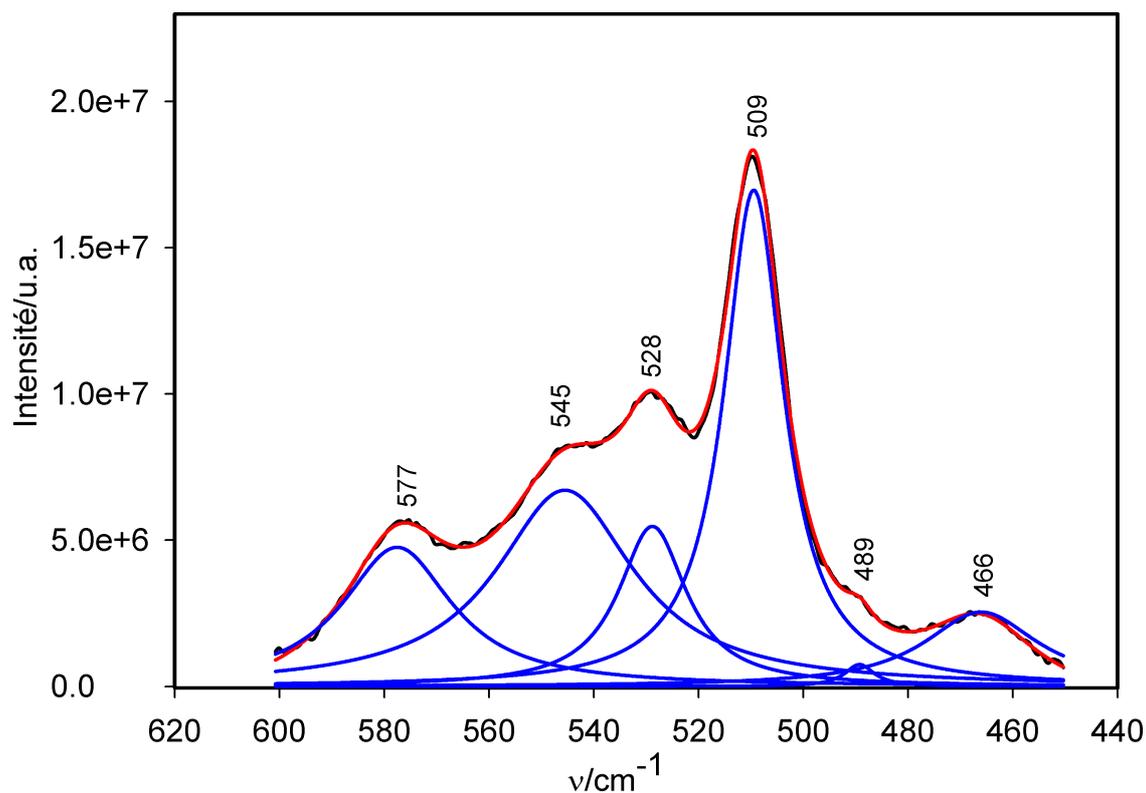


FIGURE 3.17 – Spectre Raman d'une solution aqueuse de lysozyme native à 2 mM et pH=3,2.

très faible, apparaît à 544 cm^{-1} . La précipitation de la protéine en présence de TCEP nous a empêché de vérifier si cette bande faible disparaît lors de la réduction des ponts S-S.

Ribonucléase A

La ribonucléase A (RNase A), la plus petite des enzymes connues (124 acides aminés), contient dans sa structure protéique quatre ponts S-S. Ces quatre ponts donnent une bande S-S très nette à 518 cm^{-1} d'une largeur à mi-hauteur de 18 cm^{-1} (Figure 3.21). La présence du TCEP en solution n'affecte pas cette bande ce qui indique que les quatre ponts ne sont pas accessibles à l'interaction avec cet antioxydant.

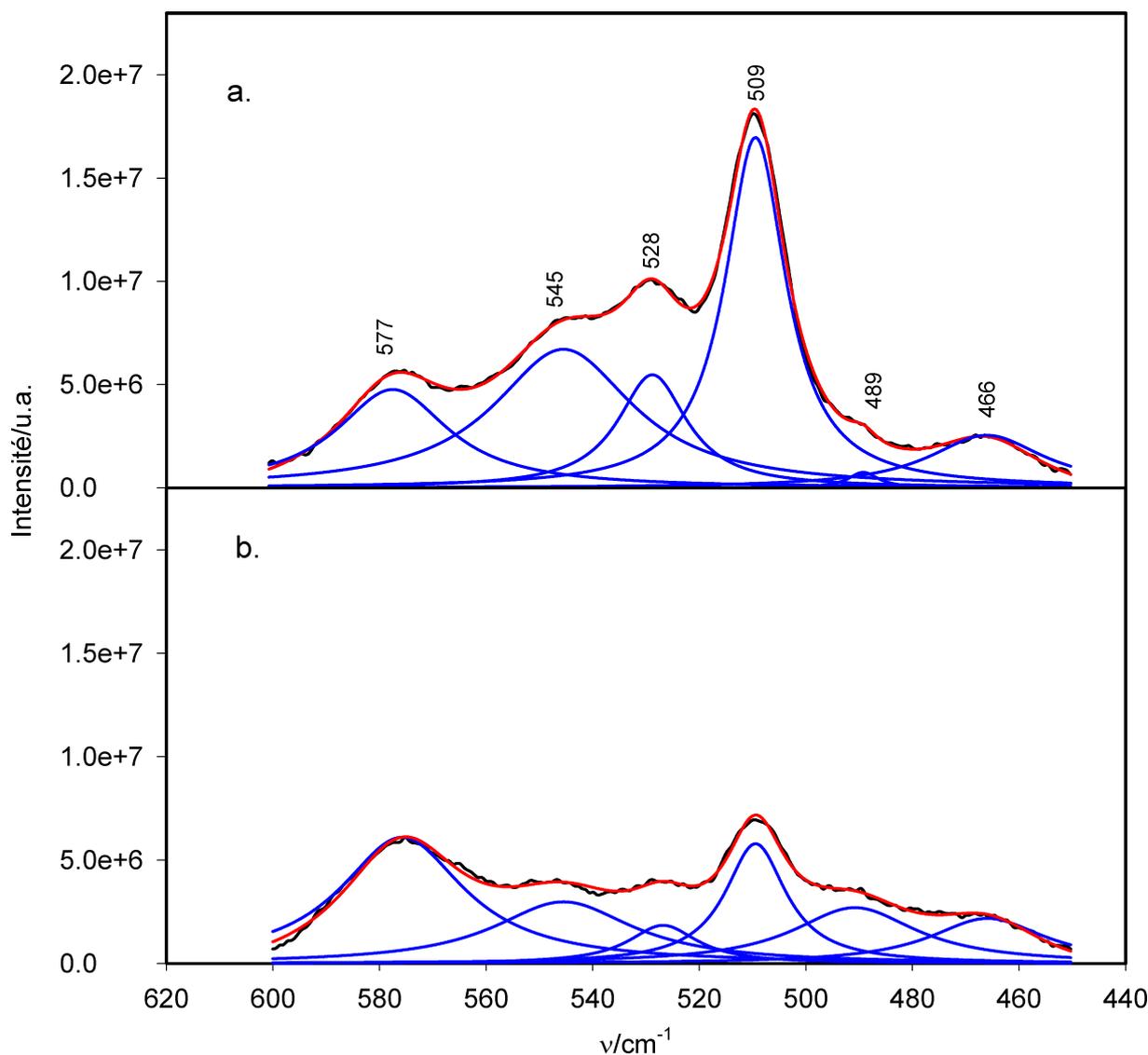


FIGURE 3.18 – (a) Spectre Raman S-S d’une solution aqueuse de lysozyme native à 2 mM et (b) en présence de TCEP à 60 mM, après 8 heures de réaction ; les mesures ont été effectuées à pH=3,2. La concentration de la protéine dans les deux cas est la même.

La trypsine

La trypsine (TRYPS), une enzyme composée de 201 acides aminés, contient six ponts S-S. La bande S-S associée à ces ponts est large (environ 35 cm^{-1} largeur à mi-hauteur) et centrée sur 514 cm^{-1} (Figure 3.22). Elle a une structure complexe qui n’a pas pu être décomposée en contributions de lorentziennes simples. D’autre part, en présence du TCEP, l’amplitude de cette bande diminue d’une manière significative. Il est intéressant de noter

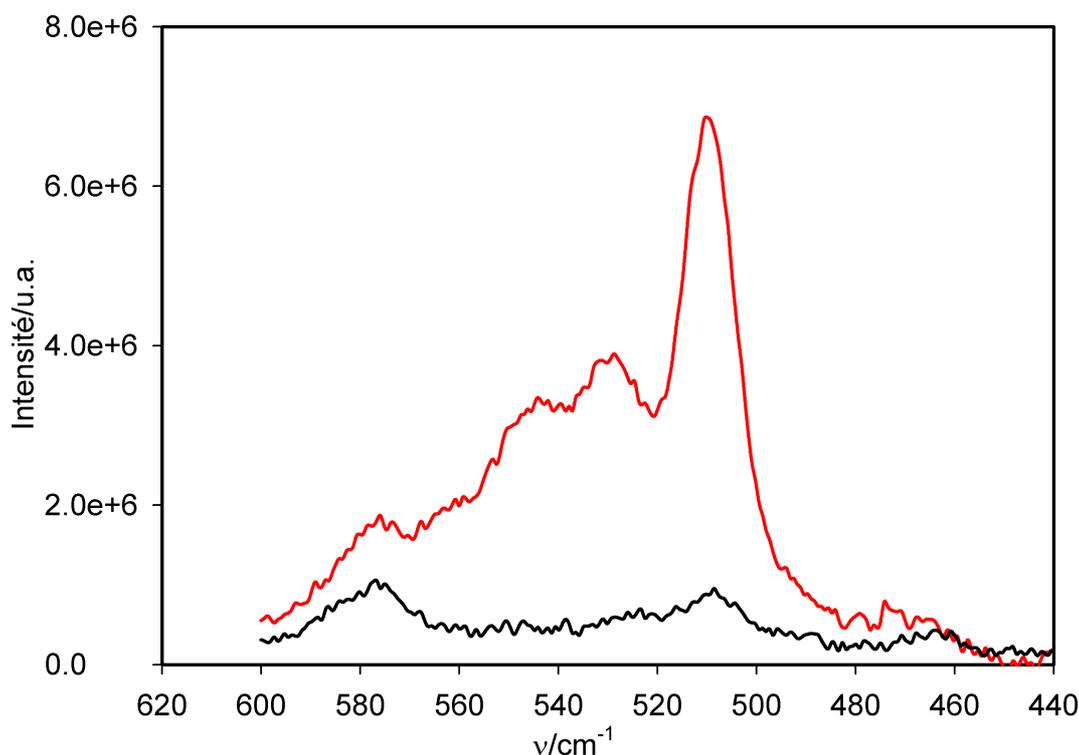


FIGURE 3.19 – Spectre Raman d’une solution aqueuse de lysozyme native à 2 mM mesuré en polarisation parallèle (rouge) et perpendiculaire (noir).

aussi que la forme de la bande change après cette réduction partielle, ce qui indique des différences spectrales et des différences de réactivité entre les 6 ponts.

3.5 Simulation des bandes Raman des quatre ponts S-S de la lysozyme

Les limitations des méthodes de modélisation moléculaire actuelles ne permettent pas de viser une simulation très fidèle des spectres vibrationnels, et surtout des spectres Raman. Les principaux facteurs qui affectent la précision des spectres calculés ont été déjà mentionnés dans l’introduction : le traitement insatisfaisant des corrélations électroniques, l’approximation harmonique, la "troncature" du système à traiter et l’omission des effets d’environnement [43, 44]. Dans le cas des spectres Raman, le calcul des intensités des bandes exige le calcul des dérivées de troisième ordre de l’énergie, ce qui amplifie les conséquences de ces approximations.

Malgré ces limitations, la modélisation moléculaire peut donner des repères, au moins

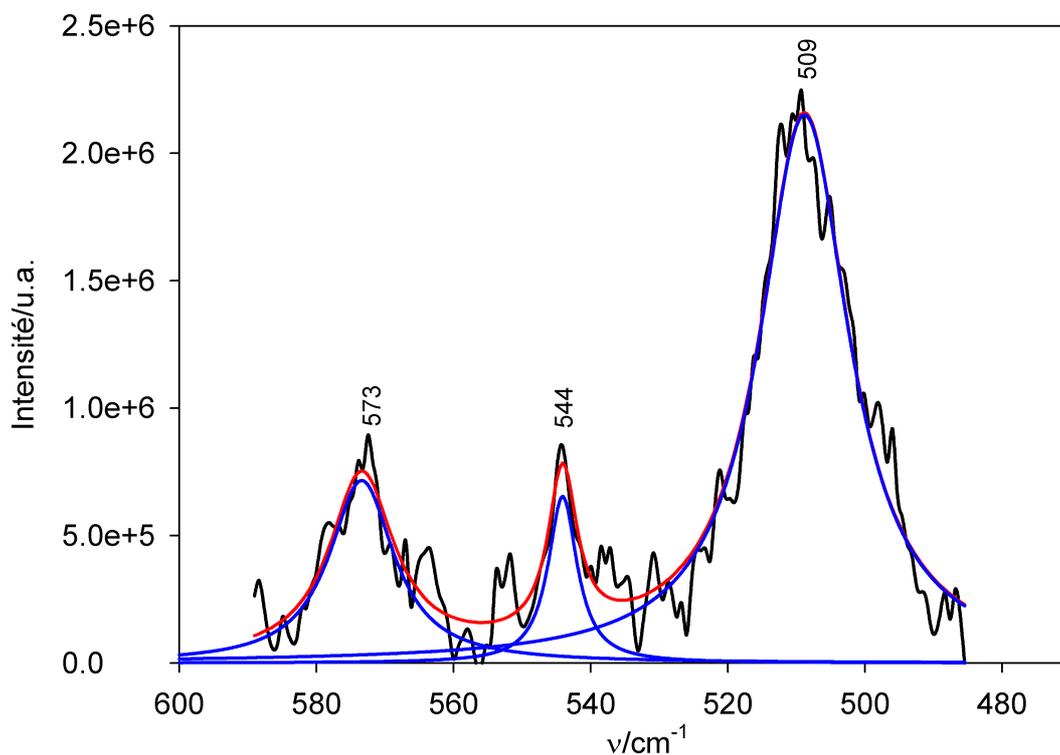


FIGURE 3.20 – Spectre Raman d’une solution aqueuse de BLG native à 2 mM et pH=7,0.

qualitatifs, dans l’interprétation des spectres Raman expérimentaux. Lorsqu’ils sont utilisés d’une manière systématique, ces repères sont évidemment plus fiables que les considérations intuitives dont on fait souvent usage dans ce type d’interprétations.

Nous avons abordé le problème de la simulation des spectres Raman des ponts S-S de la lysozyme en deux étapes. Dans une première étape nous avons effectué des calculs de type QM purs sur des modèles de taille réduite des quatre ponts. Dans l’étape suivante nous avons élargi la taille des systèmes étudiés et nous les avons traités par une méthode hybride QM/MM.

Calcul QM sur les modèles des quatre ponts S-S de la lysozyme

Le spectre Raman de la lysozyme a été simulé dans l’intervalle $450\text{-}600\text{ cm}^{-1}$ en effectuant des calculs *ab initio* sur des fragments moléculaires représentant les quatre ponts S-S de la lysozyme. Un tel fragment est présenté dans la Figure 2. Les structures de départ de ces fragments ont été extraites directement de la structure de diffraction de rayons X de la lysozyme [104] et réoptimisées au niveau de la théorie B3LYP/6-311+G(d,p). Le

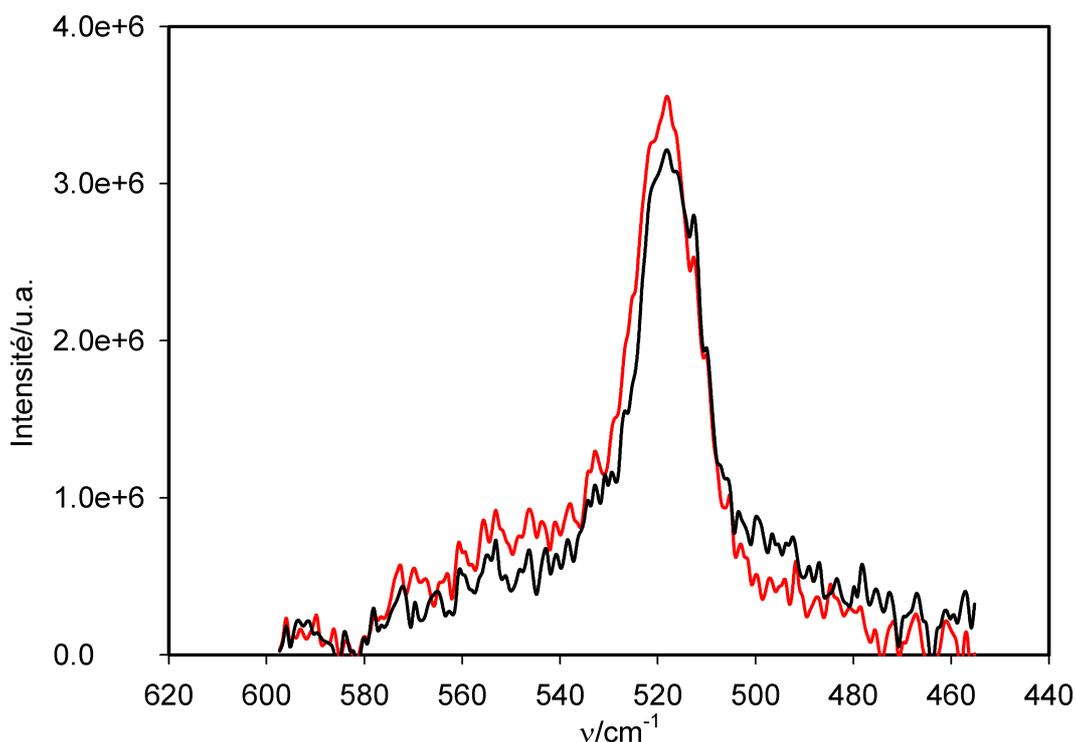


FIGURE 3.21 – Spectre Raman d’une solution aqueuse de RNase A native à 2 mM (rouge) et en présence de TCEP à 10 mM (noir), à pH=8,13. La concentration de la protéine dans les deux cas est la même.

tableau 3.1 regroupe les valeurs calculées des angles dièdres autour de la liaison S-S de chaque pont disulfure ainsi que leurs valeurs expérimentales. L’écart standard (RMSD) des valeurs théoriques par rapport aux valeurs expérimentales a été d’environ 12° , ce qui représente un écart relativement faible. Pour chaque structure optimisée, une analyse de modes normaux a ensuite été effectuée et complétée par le calcul des intensités Raman des bandes. Afin de comparer les résultats théoriques aux spectres expérimentaux, nous avons simulé des spectres théoriques en utilisant les fréquences et les intensités Raman calculées. Dans les spectres simulés, chaque mode normal a été représenté par une fonction de Lorentz ayant une largeur de 15 cm^{-1} (qui correspond à la FWHM instrumentale) et une amplitude proportionnelle à l’intensité Raman calculée. Les valeurs théoriques des fréquences calculées ont été multipliées par un facteur d’échelle de 1,08 choisi, afin d’optimiser l’accord avec les spectres expérimentaux.

Les spectres Raman simulés des quatre ponts S-S de la lysozyme sont présentés sur la Figure 3.23. Dans chaque cas la structure des bandes est complexe. Nos calculs montrent que parmi les modes normaux de chaque pont S-S, il y en a un où l’élongation S-S est prédominante (environ 40% en énergie). Les fréquences calculées de ces modes à prédomi-

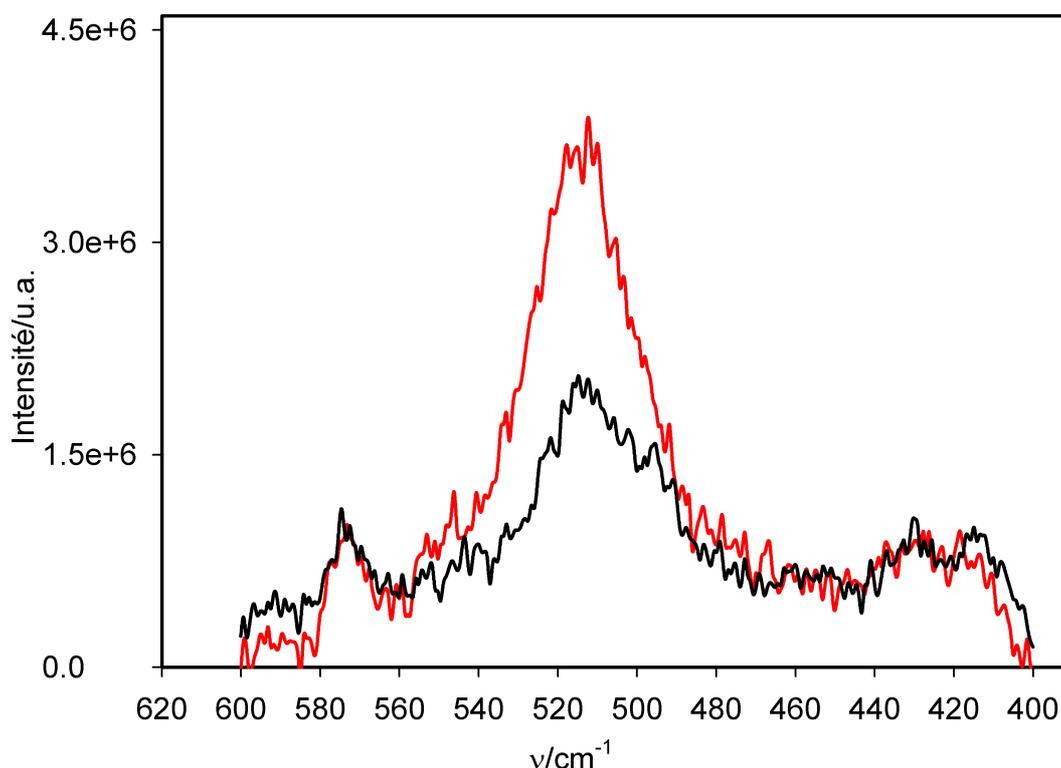


FIGURE 3.22 – Spectre Raman d’une solution aqueuse de TRYPS native à 2 mM (rouge) et en présence de TCEP à 16 mM (noir), à pH=3,0. La concentration de la protéine dans les deux cas est la même.

nance S-S ont été : 532 cm^{-1} pour le pont cys76-cys94, 514 cm^{-1} pour le pont cys30-cys115, et 507 cm^{-1} pour les ponts cys6-cys127 et cys64-cys80, respectivement. Dans le cas du pont cys76-cys94, le mode à dominance S-S est accompagné d’un mode à 517 cm^{-1} ayant une contribution de l’élongation S-S de 7% en énergie. D’une manière similaire, pour le pont cys64-cys80 une bande à 487 cm^{-1} ayant une contribution S-S de 5% en énergie a été observée.

Les résultats confirment, d’une manière qualitative, l’existence d’un déplacement des bandes S-S vers les hautes fréquences dans le cas des configurations *trans* autour de la liaison $C - S$. Ils confirment aussi l’existence des bandes qui ne sont pas à dominance S-S mais qui contiennent des contributions significatives de l’élongation S-S.

Le spectre Raman total simulé, obtenu en additionnant les quatre contributions, est présenté sur la Figure 3.24. Il a été comparé à la somme des spectres simulés des ponts S-S réduits. Afin de caractériser le degré de dépolarisation des bandes Raman, nous avons simulé aussi le spectre polarisé en direction perpendiculaire par rapport à la direction de polarisation du rayonnement incident (Figure 3.24). On constate que toutes les bandes dans l’intervalle spectral considéré, sauf peut être la bande à 577 cm^{-1} ont un faible

$\nu_{calc}/cm^{-1}/QM$		cys6-cys127	cys30-cys115	cys64-cys80	cys76-cys94
		507	514	507	523
$C_{\beta} - S - S - C_{\beta'}$	Expérimental	-84,8	-93,6	95,4	84,6
	QM	-87,4	-90,8	89,9	84,8
	QM/MM	-84,9	-106,6	96,6	82,9
$C_{\alpha} - C_{\beta} - S - S$	Expérimental	-54,2	-90,5	82,7	173,2
	QM	-70,8	-91,2	101,4	177,1
	QM/MM	-35,1	-94,3	81,2	161,8
$C_{\alpha'} - C_{\beta'} - S - S$	Expérimental	-46,1	-78,8	-60,7	52,5
	QM	-64,3	-64,1	-85,6	76,2
	QM/MM	-69,6	-87,0	-67,1	55,4
$C_{\gamma} - C_{\alpha} - C_{\beta} - S$	Expérimental	165,8	63,6	-59,3	164,7
	QM	169,8	66,0	-64,5	168,9
	QM/MM	171,0	76,9	-60,5	160,7
$C_{\gamma'} - C_{\alpha'} - C_{\beta'} - S$	Expérimental	-177,5	175,3	173,8	65,3
	QM	-179,0	173,7	169,6	70,8
	QM/MM	-167,2	177,7	179,7	58,6

Tableau 3.1 – Conformations expérimentales et théoriques (optimisées en DFT) de quatre ponts S-S de la lysozyme. Comparaison entre les angles dièdres impliqués dans les structures des ponts S-S.

degré de dépolarisation bien qu'elles n'appartiennent pas toutes à des modes totalement symétriques. De plus, toutes ces bandes sauf, encore une fois, la bande à 577 cm^{-1} , disparaissent lors de la réduction des ponts S-S. Ce résultat concerne aussi les bandes à 487 et 517 cm^{-1} qui ne sont pas à dominance S-S mais qui contiennent des contributions non négligeables de l'élongation S-S.

Calcul QM/MM

Les bandes Raman simulées présentées sur la Figure 3.23 ont été obtenues en négligeant les effets de l'environnement moléculaire sur la structure et la dynamique des ponts S-S. Des études systématiques sur des peptides en solution aqueuse ont démontré l'importance du solvant dans la prédiction correcte des structures peptidiques et des spectres Raman [106–108]. Dans notre cas, l'environnement des ponts S-S est hydrophobe et les interactions qui devraient être prises en compte en priorité sont celles entre les résidus cystéine et les acides aminés voisins. Ces interactions pourraient entraîner une perturbation du mode de vibration S-S par le couplage des modes à travers la chaîne peptidique.

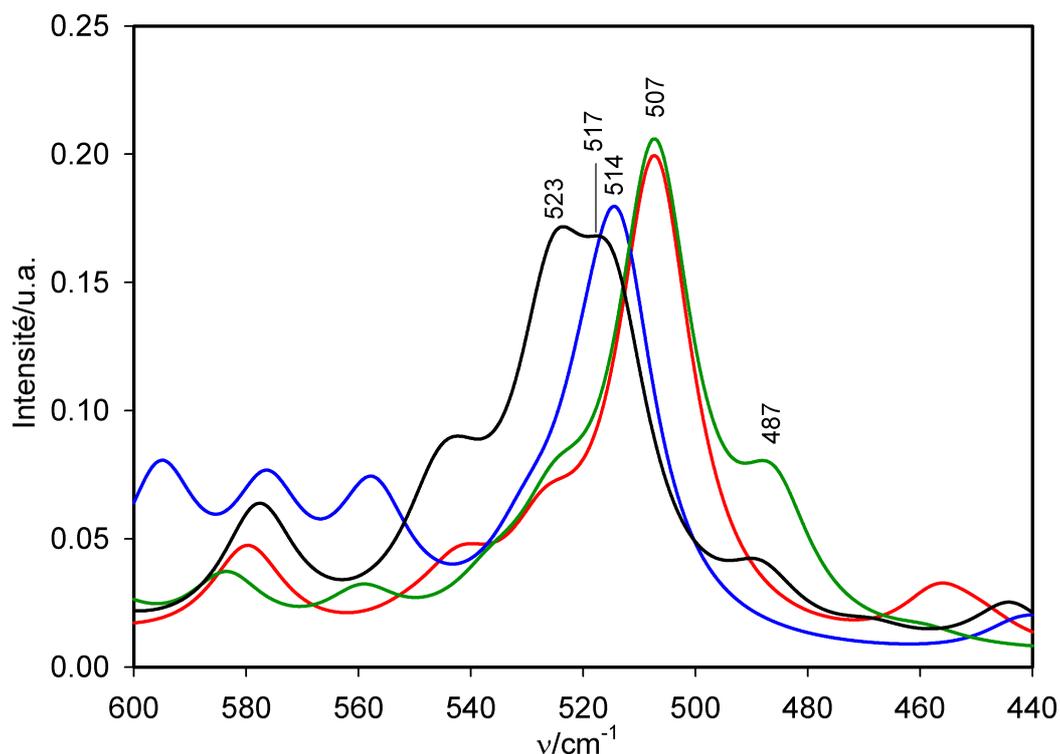


FIGURE 3.23 – Spectres Raman simulés des quatre ponts S-S de la lysozyme : cys6-cys127 (rouge), cys30-cys115 (bleu), cys64-cys80 (vert) et cys76-cys94 (noir).

Afin d'inclure tous ces effets dans nos calculs, nous avons construit des systèmes modèles étendus qui consistent en environ 30 acides aminés entourant un pont S-S (Figure 3.25). Les quatre structures de départ ont été extraites de la structure de diffraction des rayons X de la lysozyme [104] et ont été réoptimisées par des calculs hybrides du type QM/MM. Ces calculs ont été effectués en utilisant la méthode IMOMO [87]. Le système MM a été traité en utilisant le champ de forces AMBER implémenté dans le logiciel Gaussian 03, tandis que le système QM a été traité avec la méthode B3LYP/6-311+G(d,p). Les valeurs des angles dièdres qui caractérisent les conformations des ponts S-S sont données dans le Tableau 3.1. Ces valeurs sont en accord avec les valeurs expérimentales mais elles n'apportent pas d'amélioration par rapport aux valeurs issues du calcul QM.

Le spectre Raman de la lysozyme, obtenu par addition des contributions des quatre modèles QM/MM, est présenté sur la Figure 3.26. Il n'apporte pas d'amélioration par rapport au spectre simulé au niveau QM. Au contraire, bien que les deux bandes, centrées respectivement sur 510 et 523 cm^{-1} sont correctement prédites, le rapport de leurs intensités ne ressemble pas à celui observé dans le spectre expérimental.

Cette première tentative d'application de la méthode QM/MM à la simulation des bandes S-S montre bien les difficultés spécifiques d'une telle approche. Des progrès seront

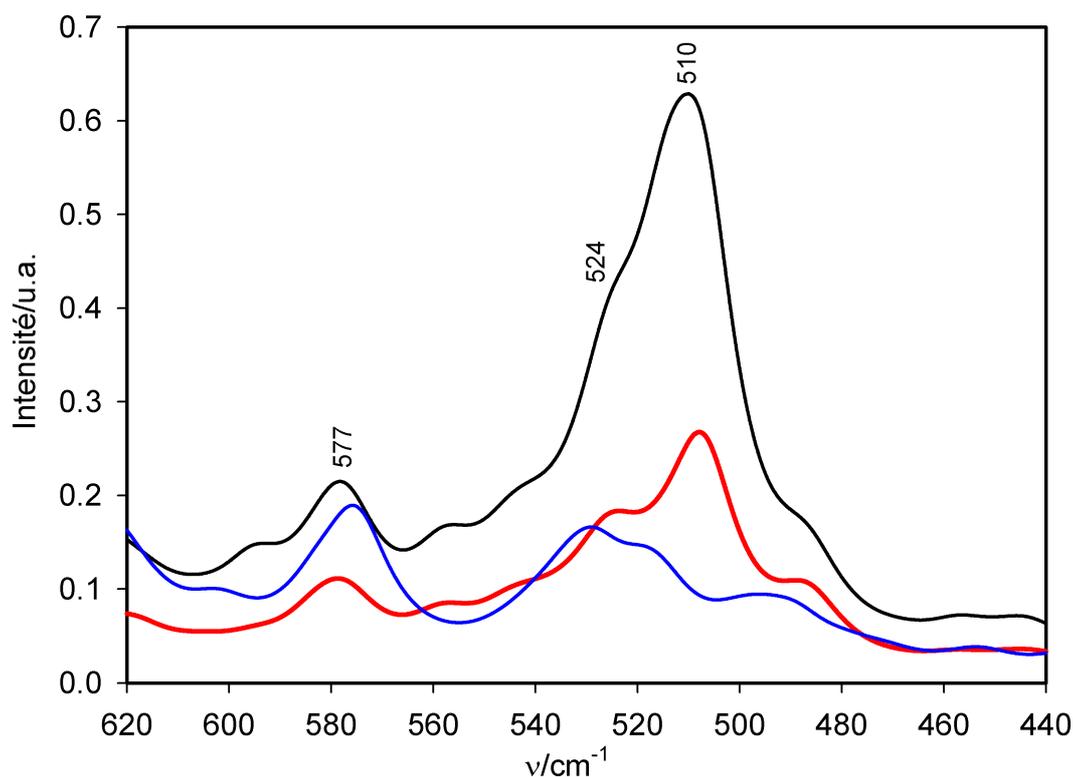


FIGURE 3.24 – Spectres Raman simulés de la lysozyme : spectre Raman total de la lysozyme native (noir), de la lysozyme totalement réduite (bleu) et de la lysozyme en polarisation perpendiculaire (rouge).

certainement possibles à condition que le traitement du couplage QM/MM soit amélioré et optimisé pour ce type de calculs.

Conclusions

Les simulations des bandes Raman S-S de la lysozyme confirment d'une manière directe quelques hypothèses qui ont été avancées dans la première partie de ce chapitre à l'occasion de la discussion de l'attribution des bandes S-S des protéines et des systèmes modèles :

- La conformation *trans* autour de la liaison $C - S$ entraîne un déplacement des bandes S-S vers les hautes fréquences, d'environ 15 cm^{-1} .
- Un pont S-S peut contribuer par plus d'une bande au spectre Raman dans l'intervalle spectral $500\text{-}560 \text{ cm}^{-1}$.
- Dans cet intervalle spectral, il existe des bandes qui ne sont pas des bandes à domi-

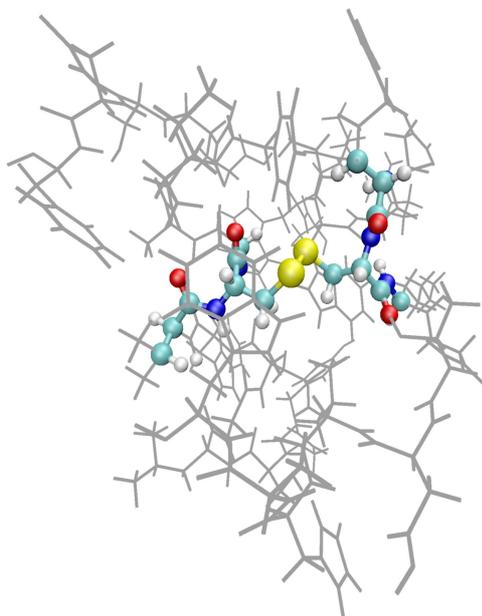


FIGURE 3.25 – Système modèle du pont cys30-cys115 utilisé dans les calculs QM/MM. La partie du système traitée en QM est représentée par des billes.

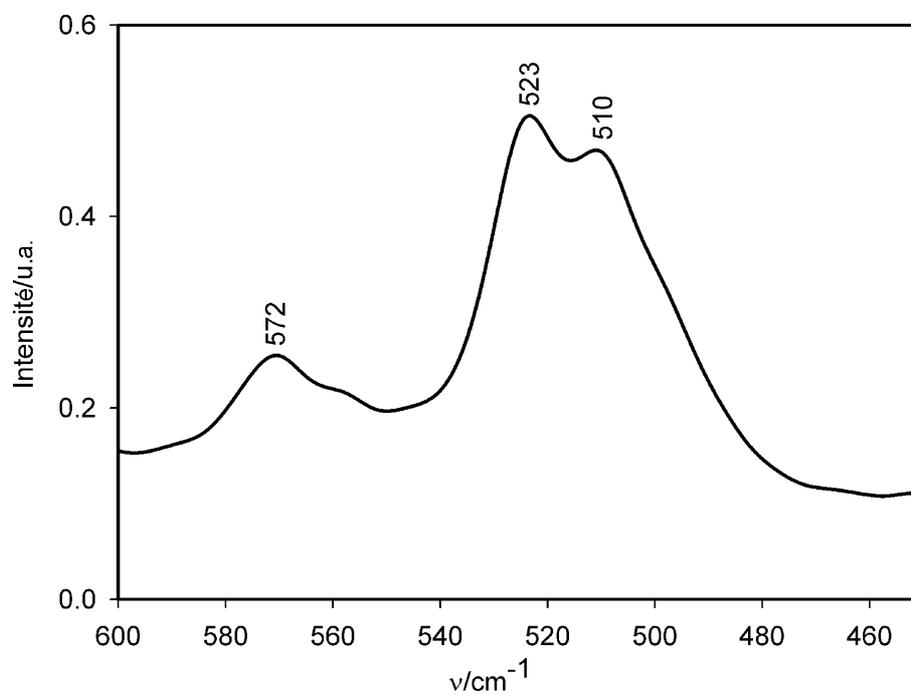


FIGURE 3.26 – Spectre Raman de la lysozyme simulé par des calculs QM/MM.

nance S-S mais qui présentent une contribution non négligeable de l'élongation S-S. Ces bandes disparaissent lors de la réduction des ponts S-S.

- Presque toutes les bandes dans cet intervalle spectral ont un faible degré de dé-polarisation sans pour autant qu'elles soient associées à des modes vibrationnels totalement symétriques.

Chapitre 4

Réduction des ponts S-S dans les protéines : application de la spectroscopie Raman

Sommaire

4.1	Introduction	80
4.2	Réduction des ponts S-S dans les albumines sériques	82
4.2.1	Réaction avec le DTT	82
4.2.1.1	Aspect thermodynamique de la réaction	82
4.2.1.2	Le dépliement de la protéine	88
4.2.1.3	Cinétiques de réaction	91
4.2.1.4	Évaluation de la variation d'enthalpie libre conformationnelle	97
4.2.2	Réaction avec la TCEP	100
4.3	Réduction des ponts S-S dans la lysozyme	105
4.3.1	Aspect thermodynamique de la réaction	106
4.3.2	Cinétiques de réaction	107
4.3.3	Les effets conformationnels	112

4.1 Introduction

Afin d'établir quelques repères expérimentaux concernant les mécanismes de réduction des ponts disulfure dans les protéines, nous avons choisi comme systèmes modèles les albumines sériques et la lysozyme. Les albumines sériques présentent un contenu élevé en ponts disulfure, 17 par molécule. Quant au lysozyme, elle n'en contient que quatre, mais qui sont bien caractérisés du point de vue structural, ce qui rend possibles des études visant la spécificité de ces ponts.

Dans ce travail, la réactivité des ponts disulfure a été étudiée sous deux aspects : thermodynamique, qui est quantifié par l'enthalpie libre de réaction, et cinétique, qui est quantifié par la constante de vitesse de réaction. Dans tous les cas nous nous sommes interrogés principalement sur le rôle joué par la structure protéique dans cette réactivité. La spectroscopie Raman nous a permis de suivre simultanément l'avancement de la réaction de coupure de ponts et le dépliement de la protéine. L'analyse a été basée sur la détection des modifications des bandes de vibration S-S et de la bande amide I. L'attribution des bandes Raman des protéines dans le domaine spectral d'intérêt a joué un rôle primordial dans l'interprétation correcte des cinétiques de réaction.

Les mécanismes de réduction des ponts S-S dépendent d'une manière importante du type d'agent réducteur intervenant dans la réaction. Il existe plusieurs agents réducteurs efficaces pour la réduction des ponts S-S, ayant des structures et des tailles variées. Leur spécificité vient de leur facilité d'accès à la cible (le pont S-S), mais aussi de leur interaction avec la cible. Les antioxydants utilisés dans l'étude présente sont le DTT et la TCEP.

Le DTT

Le DTT est une petite molécule réductrice, également connue sous le nom de réactif de Cleland. La structure moléculaire de la forme réduite de ce composé apparaît sur la Figure 4.1. Elle contient deux groupes thiol, dont les constantes d'acidité sont : $pK_{a1} = 9,2$ et $pK_{a2} = 10,1$. À l'état oxydé, il se cyclise en formant un pont disulfure intramoléculaire. Le DTT est un agent réducteur fort dont le potentiel redox est de $-0,33$ V à pH 7,0. La réduction d'un pont disulfure procède par un mécanisme séquentiel en deux étapes, comme illustré ci-dessous (Figure 4.1). L'intermédiaire contient un pont disulfure mixte et il est peu stable car le second groupe thiol du DTT a une forte tendance à refermer le cycle, donnant ainsi le DTT oxydé et deux groupes thiol libres à la place du pont disulfure initial. Le DTT en solution aqueuse est relativement instable, étant sensible à

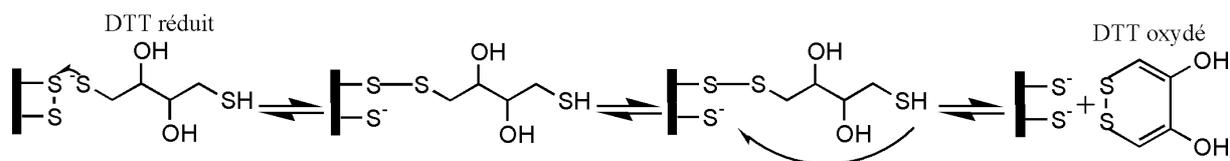


FIGURE 4.1 – Réaction d'échange des ponts disulfure entre la protéine et le DTT.

l'oxydation par l'oxygène de l'air. Afin de tester la vitesse d'autooxydation du DTT en solution aqueuse, nous avons mesuré l'amplitude de la bande Raman S-S du DTT oxydé ainsi formée en fonction du temps. Nous avons trouvé que, même pour des concentrations importantes de DTT (80 mM), au bout de trois heures l'autooxydation était négligeable.

La TCEP

La TCEP réduit les ponts S-S des protéines dans une réaction se déroulant en deux étapes (Figure 4.2) [109].

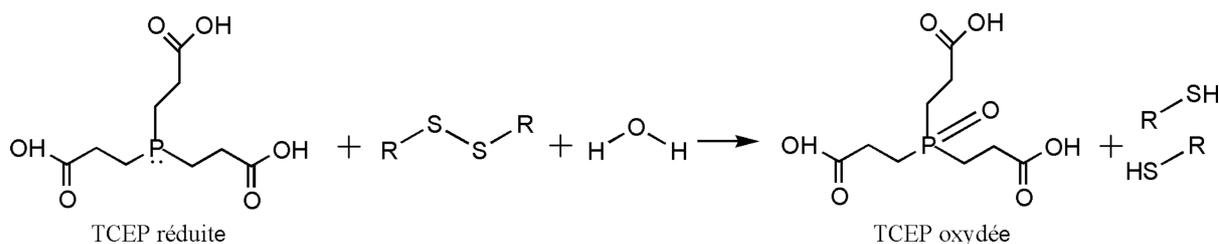


FIGURE 4.2 – Réduction des ponts disulfure par la TCEP.

Dans une première étape l'atome de phosphore (P) se substitue à l'un des deux atomes de soufre (S) dans la liaison S-S : la liaison S-S est clivée et une liaison S-P est formée. La deuxième étape est une réaction d'hydrolyse qui a pour résultat la formation d'une double liaison P=O et la libération du thiol.

En solution aqueuse, les phosphines sont stables. De plus la TCEP est non-réactive par rapport aux autres groupements présents dans la protéine. En comparaison avec le DTT, la TCEP est plus efficace en permettant la réduction des ponts S-S à des valeurs de pH plus basses.

La TCEP est réactive seulement dans sa forme déprotonée en P. À pH acide, il coexiste quatre espèces actives qui diffèrent par le nombre de groupements carboxyle ionisés : TCEPH3 (protonée sur les trois groupements carboxyle), TCEPH2 (protonée sur deux groupements carboxyle), TCEPH1 (protonée sur un groupement carboxyle) et TCEPH0 (déprotonée sur les trois groupements carboxyle).

4.2 Réduction des ponts S-S dans les albumines sériques

4.2.1 Réaction avec le DTT

4.2.1.1 Aspect thermodynamique de la réaction

Les réactions d'échange des ponts S-S sont des réactions d'équilibre. Nous avons étudié cet équilibre dans le cas de la réaction entre le DTT et les albumines sériques en suivant l'avancement de la réaction en fonction de la concentration du DTT présent en solution. Ce suivi a été réalisé à l'aide de la spectroscopie Raman des bandes S-S. Une première difficulté expérimentale vient du fait que le DTT oxydé possède lui aussi une bande S-S. Une analyse détaillée des bandes dans le domaine spectral $440\text{-}620\text{ cm}^{-1}$ a été donc nécessaire afin de séparer la contribution du DTT de celle de la protéine. La fonction modèle utilisée pour cette analyse est la suivante :

$$S(\nu) = a_1L_1(\nu) + a_2L_2(\nu) + a_3L_3(\nu) + a_4L_4(\nu) + a_5f_{DTT}(\nu) + a_6Fl(\nu) + a_7w(\nu) \quad (4.1)$$

où L_1 , L_2 , L_3 , L_4 sont des lorentziennes qui décrivent le spectre Raman de la protéine, f_{DTT} est une fonction qui représente la bande S-S du DTT oxydé, $Fl(\nu)$ est une contribution de fluorescence, $w(\nu)$ est la contribution de l'eau et les paramètres $a_1 - a_7$ sont les amplitudes des ces contributions. Les paramètres $a_1 - a_7$ sont les seuls paramètres ajustables dans la modélisation. La forme des fonctions L_1 , L_2 , L_3 , L_4 et Fl a été préalablement déterminée en analysant le spectre Raman de la protéine native. De même, la forme de la bande S-S du DTT oxydé a été déterminée en mesurant le spectre Raman d'une solution aqueuse de ce produit. Conformément à la discussion présentée dans le troisième chapitre, parmi les quatre lorentziennes appartenant à la protéine, seulement deux ont été attribuées à l'élongation S-S.

Pour toutes les mesures, la concentration des albumines sériques a été de 1 mM. Nous avons fait varier la concentration du DTT réduit ajouté à la solution de protéine native dans l'intervalle 0-30 mM. Le pH des solutions mesurées a été ajusté à 7,8 par l'addition des gouttes de NaOH concentré (2 M). Les mesures ont été effectuées à 4 heures de la mise en contact des réactifs afin que l'équilibre soit effectivement atteint.

Suite à l'addition du DTT à une solution aqueuse d'albumine sérique, les bandes S-S caractéristiques de la protéine, à 507 et 526 cm^{-1} , respectivement, diminuent tandis qu'une nouvelle bande S-S, celle du DTT oxydé, apparaît (Figures 4.3 et 4.4).

Les données expérimentales obtenues à des différentes concentrations de DTT ajoutées

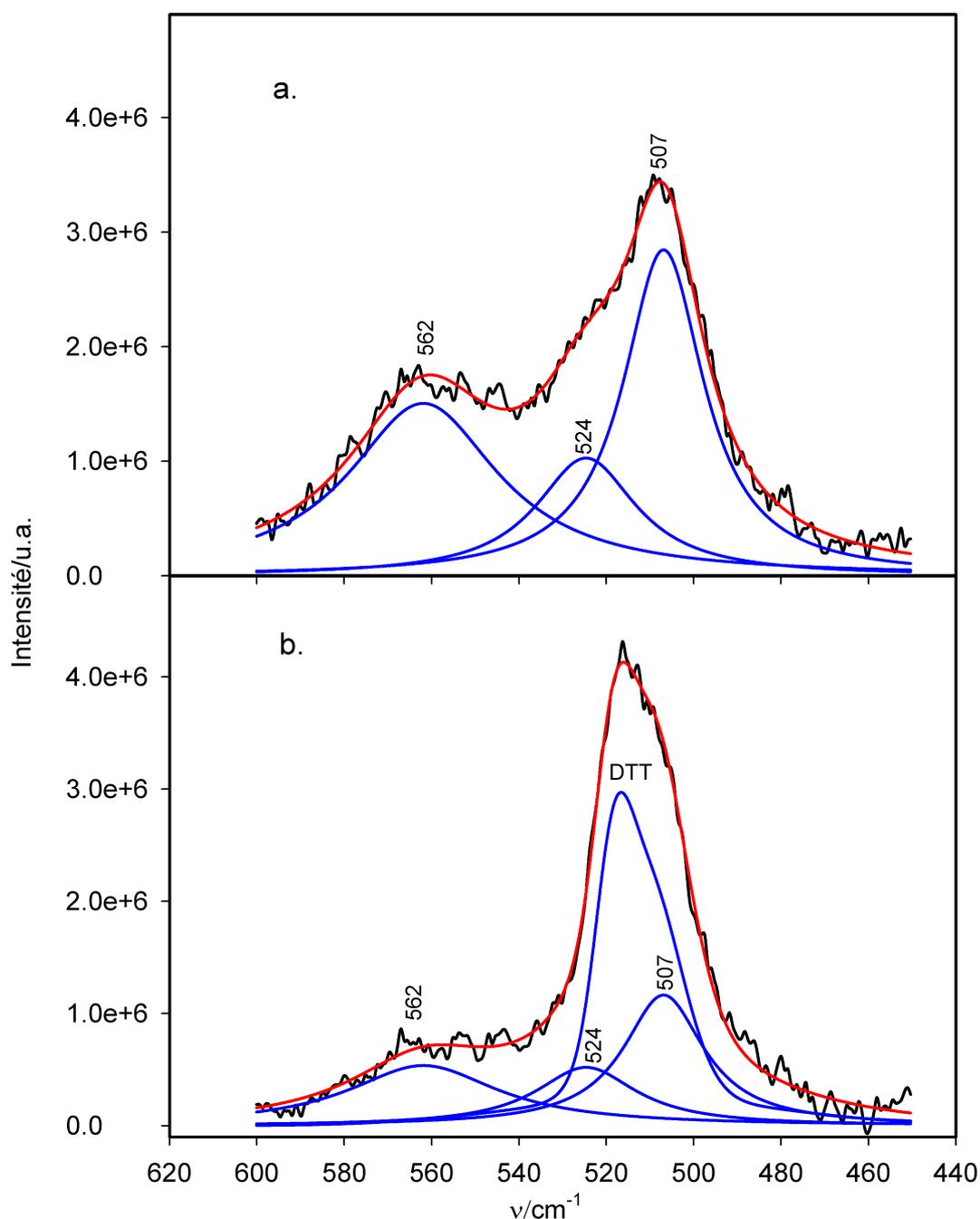


FIGURE 4.3 – Spectre Raman d’une solution aqueuse de HSA à 1 mM et à pH 7,8 : (a) spectre de la protéine native (b) spectre de la protéine (même concentration) en présence du DTT à 30 mM.

ont été quantifiées en convertissant les surfaces des bandes en concentration de ponts S-S. Cette conversion a été réalisée en supposant que chaque pont apporte la même contribution (en surface) au spectre total. L’évolution expérimentale des concentrations des ponts contribuant aux deux bandes S-S de la protéine et, respectivement, à la bande S-S du

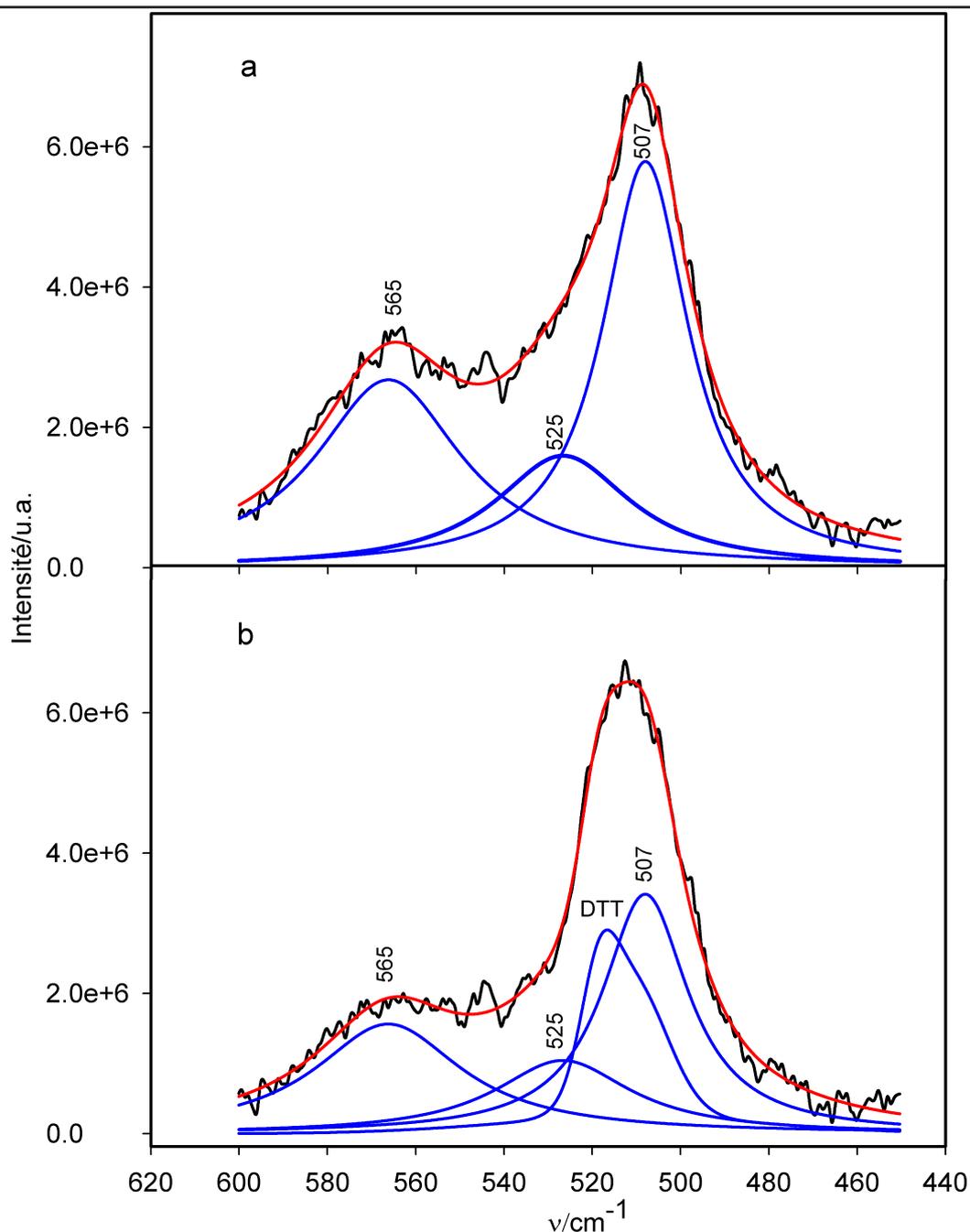


FIGURE 4.4 – Spectre Raman d’une solution aqueuse de BSA à 1 mM et à pH 7,8 : (a) spectre de la protéine native (b) spectre de la protéine (même concentration) en présence du DTT à 30 mM.

DTT sont présentées dans les Figures 4.5 et 4.6 (pour la HSA et la BSA, respectivement). Cette évolution peut ensuite être analysée afin de déterminer la constante d’équilibre de la réaction d’échange des ponts entre la protéine et le DTT.

Le modèle utilisé dans ce traitement est un modèle simplifié basé sur les hypothèses

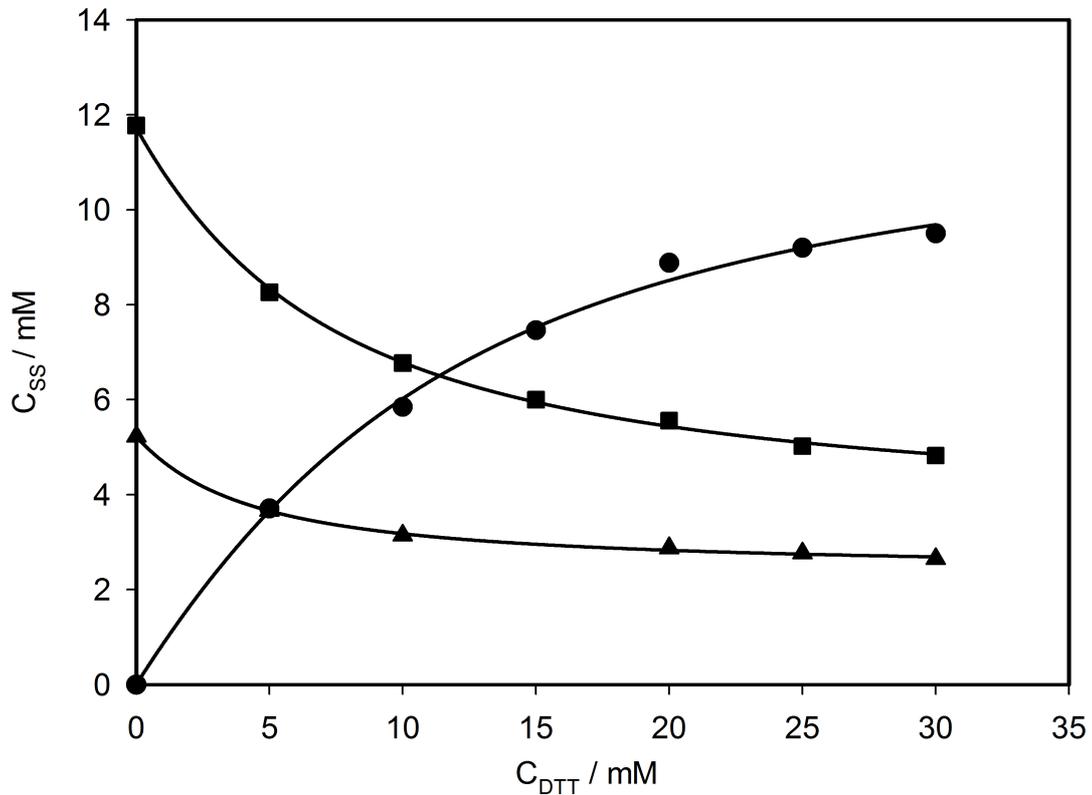


FIGURE 4.5 – Les concentrations de ponts S-S de la HSA et du DTT oxydé sont représentées en fonction de la concentration de DTT ajouté. Les triangles représentent la contribution à la bande Raman S-S de la protéine, centrée sur 507 cm^{-1} , les carrés la contribution à la bande centrée sur 524 cm^{-1} et les cercles la contribution à la bande Raman S-S du DTT oxydé.

suivantes :

- il existe deux types de ponts S-S : réactifs et non-réactifs
- tous les ponts réactifs sont caractérisés par la même constante d'équilibre
- la réduction de certains ponts n'affecte pas la constante d'équilibre des ponts non-réduits

Les différentes étapes de la réaction sont détaillées dans l'annexe.

Dans ce modèle, la constante d'équilibre de l'équation 4.1, a pour expression :

$$K_{eq}^{obs} = \frac{[(Prot_{SH,SH})_U]_{eq}[DTT_{SS}]_{eq}}{[DTT_{SH,SH}]_{eq}[(Prot_{SS})_N]_{eq}} \quad (4.2)$$

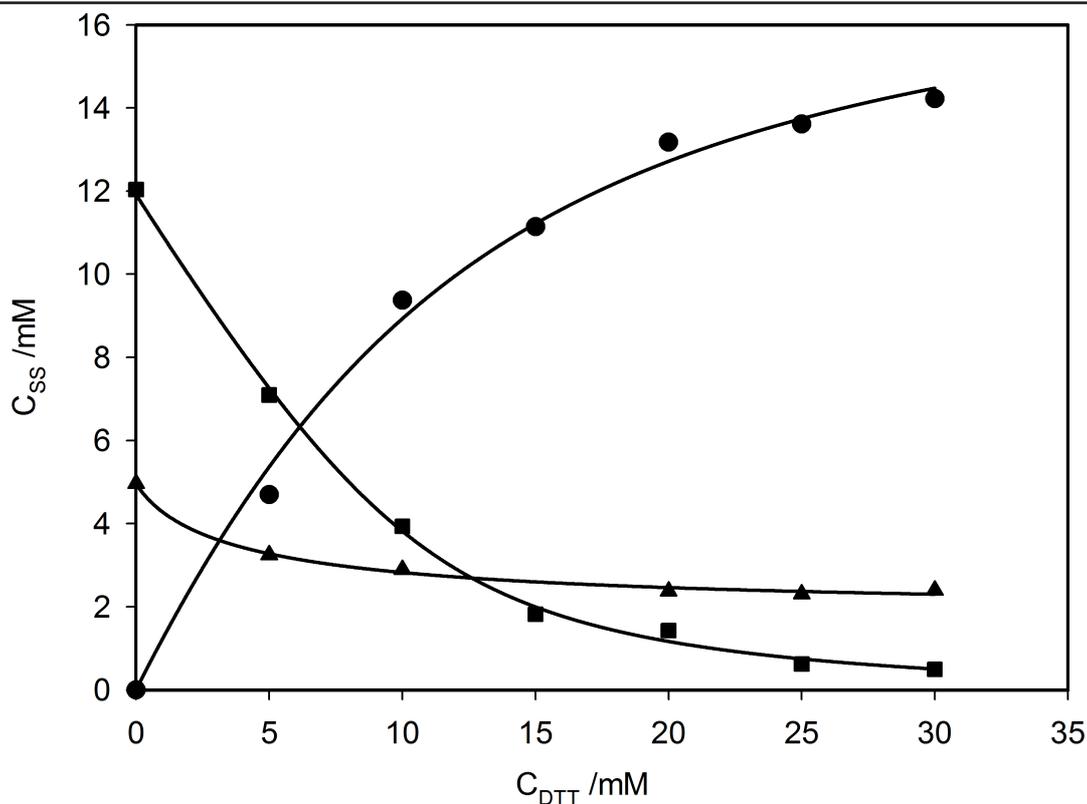


FIGURE 4.6 – Les concentrations de ponts S-S de la BSA et du DTT oxydé sont représentées en fonction de la concentration de DTT ajouté. Les triangles représentent la contribution à la bande Raman S-S de la protéine, centrée sur 507 cm^{-1} , les carrés la contribution à la bande centrée sur 525 cm^{-1} et les cercles la contribution à la bande Raman S-S du DTT oxydé.

où $[(Prot_{SS})_N]$ et $[(Prot_{SH,SH})_U]$ sont les concentrations molaires équivalentes en ponts réactifs et en ponts réduits, respectivement. Les notations N et U indiquent la conformation native de la protéine et la conformation dépliée après la coupure du pont S-S, respectivement. La concentration de DTT oxydé ($[DTT_{SS}]$), la concentration initiale de DTT réduit ($[DTT_{SH,SH}]_I$), la concentration initiale de ponts S-S réactifs ($[(Prot_{SH,SH})_I]$) et la constante d'équilibre sont reliées par l'équation (voir l'annexe) :

$$[DTT_{SS}]_{eq} = \left\{ K_{eq}^{obs} \left([(Prot_{SS})_N]_I + [DTT_{SH,SH}]_I \right) - \left((K_{eq}^{obs})^2 \left([(Prot_{SS})_N]_I + [DTT_{SH,SH}]_I \right)^2 + 4K_{eq}^{obs} \alpha \left([(Prot_{SS})_N]_I [DTT_{SH,SH}]_I \right)^{1/2} \right) \right\} / 2(K_{eq}^{obs} - \alpha) \quad (4.3)$$

Ici α est un paramètre qui dépend du pH et des valeurs des pKa des groupes thiol impliqués dans la réaction d'échange de ponts S-S :

$$\alpha = \frac{1 + 10^{(pH-pK_{a1})} + 10^{(pH-pK_{a2})} + 10^{(2pH-pK_{a1}-pK_{a2})}}{1 + 2 \times 10^{(pH-pK_{ap})} + 10^{2(pH-pK_{ap})}} \quad (4.4)$$

où pK_{a1} et pK_{a2} sont les deux constantes d'acidité des thiols du DTT réduit. pK_{ap} est la constante d'acidité moyenne des thiols de la protéine qui se forment lors de la réduction des ponts. Nous l'avons déterminée en mesurant l'intensité de la bande Raman SH d'une solution aqueuse de BSA (1 mM) réduite pour des différentes valeurs de pH (Figure 4.7). Cette courbe expérimentale a été modélisée en utilisant la fonction modèle suivante :

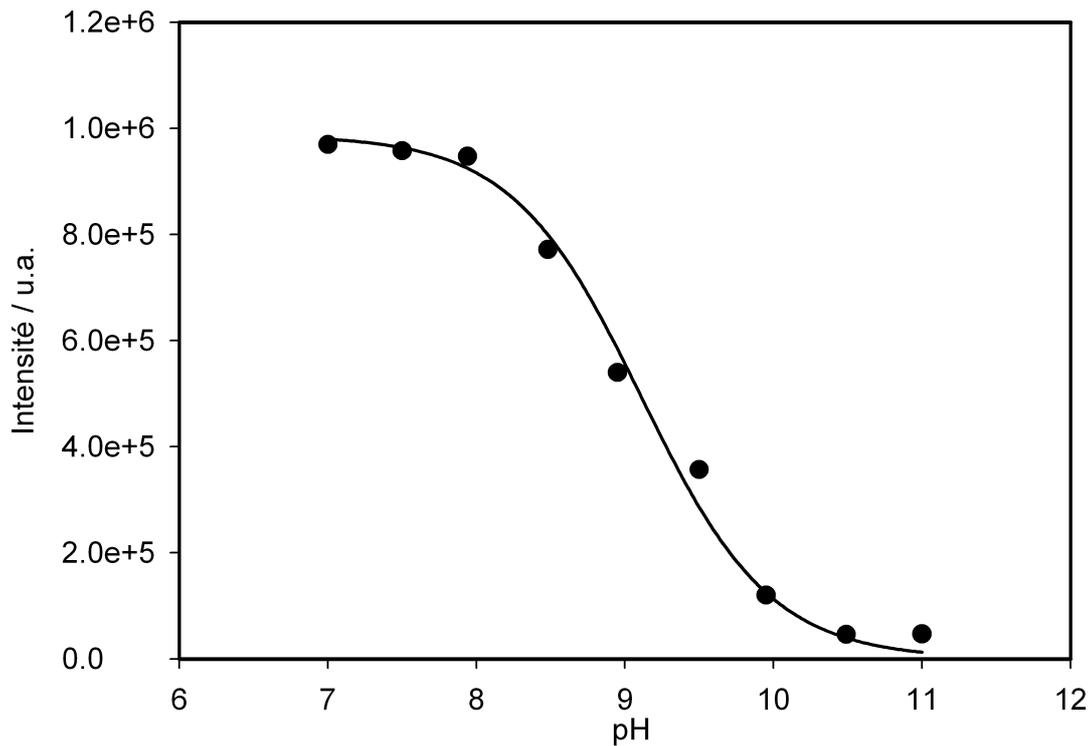


FIGURE 4.7 – Courbe de titrage de l'acidité des groupes thiol dans la BSA obtenue par spectroscopie Raman.

$$[SH] = \frac{[SH_0] \times 10^{-pH}}{10^{-pK_a} + 10^{-pH}} \quad (4.5)$$

La valeur obtenue est de 9,1.

Ainsi, en modélisant les données dans les Figures 4.5 et 4.6 dans l'équation 4.3 nous avons trouvé $K_{eq,HSA}^{obs} = 1,9$ et $K_{eq,BSA}^{obs} = 2,1$. Le nombre des ponts réactifs trouvé pour les deux albumines sériques a été le même, 14.

Le fait que les données expérimentales aient pu être modélisées en utilisant une seule constante d'équilibre pour tous les ponts ne veut pas dire qu'il n'y a pas de différence de réactivité entre les 17 ponts. Tout ce que l'on peut affirmer est que les différentes constantes d'équilibre doivent avoir le même ordre de grandeur.

Une fois la constante d'équilibre connue, il est possible de calculer la variation d'en-

thalpie libre (ΔG_{SS}) lors de la réaction de coupure d'un pont S-S :

$$K_{eq}^{obs} = e^{-\Delta G_{SS}/RT} \quad (4.6)$$

où R est la constante universelle des gaz et T est la température. Pour $K_{eq}^{obs} \approx 2$ on obtient (ΔG_{SS}) = $-1,7 \text{ kJ.mol}^{-1}$. C'est une valeur très faible qui indique que la réduction des ponts sera complète seulement pour une concentration de DTT largement en excès par rapport à la concentration des ponts. Néanmoins, les valeurs des constantes d'équilibre calculées pour la réaction de réduction des ponts des albumines sériques ont le même ordre de grandeur que celles calculées pour les ponts les plus réactifs de BPTI [7].

4.2.1.2 Le dépliement de la protéine

La bande Raman amide I des protéines est très sensible aux changements de structure secondaire, sa forme présentant des différences notables d'une protéine à l'autre. Elle est souvent analysée comme une superposition de plusieurs gaussiennes. Chaque gaussienne est attribuée à une structure secondaire spécifique. Dans ce travail, nous avons modélisé les spectres Raman des albumines sériques dans l'intervalle spectral $1500\text{-}1800 \text{ cm}^{-1}$ par une somme de huit gaussiennes : quatre pour représenter la bande amide I et quatre autres pour représenter les vibrations des résidus aromatiques :

$$S(\nu) = \sum_{i=1}^8 a_i g_i(\nu) + a_9 w(\nu) \quad (4.7)$$

Dans l'équation 4.7, les amplitudes a_1 à a_9 sont des paramètres ajustables et les fonctions $g_i(\nu)$ des gaussiennes dont les positions et les largeurs à mi-hauteur sont maintenues fixes telles qu'elles ont été déterminées en analysant le spectre Raman de la protéine native. La fonction $w(\nu)$ représente la bande Raman de l'eau.

La décomposition de la bande amide I des albumines sériques, natives et réduites, est donnée dans les Figures 4.8 et 4.9. Les bandes (1-4) correspondent aux vibrations des anneaux des résidus aromatiques tandis que les bandes (5-8) sont attribuées effectivement à la bande amide I. La bande centrée sur 1655 cm^{-1} (la bande 6) est attribuée dans la littérature aux hélices α orientées (ho), celle centrée sur 1671 cm^{-1} (la bande 7) aux hélices α désorientées (hu) et celle centrée sur 1684 cm^{-1} (la bande 8) aux structures secondaires du type *turn* (t). La contribution centrée autour de 1630 cm^{-1} (la bande 5) est d'habitude attribuée aux structures du type feuillet- β [96, 97, 110–112]. Dans le cas des albumines sériques (où les structures de type feuillet- β sont absentes [102]) la bande

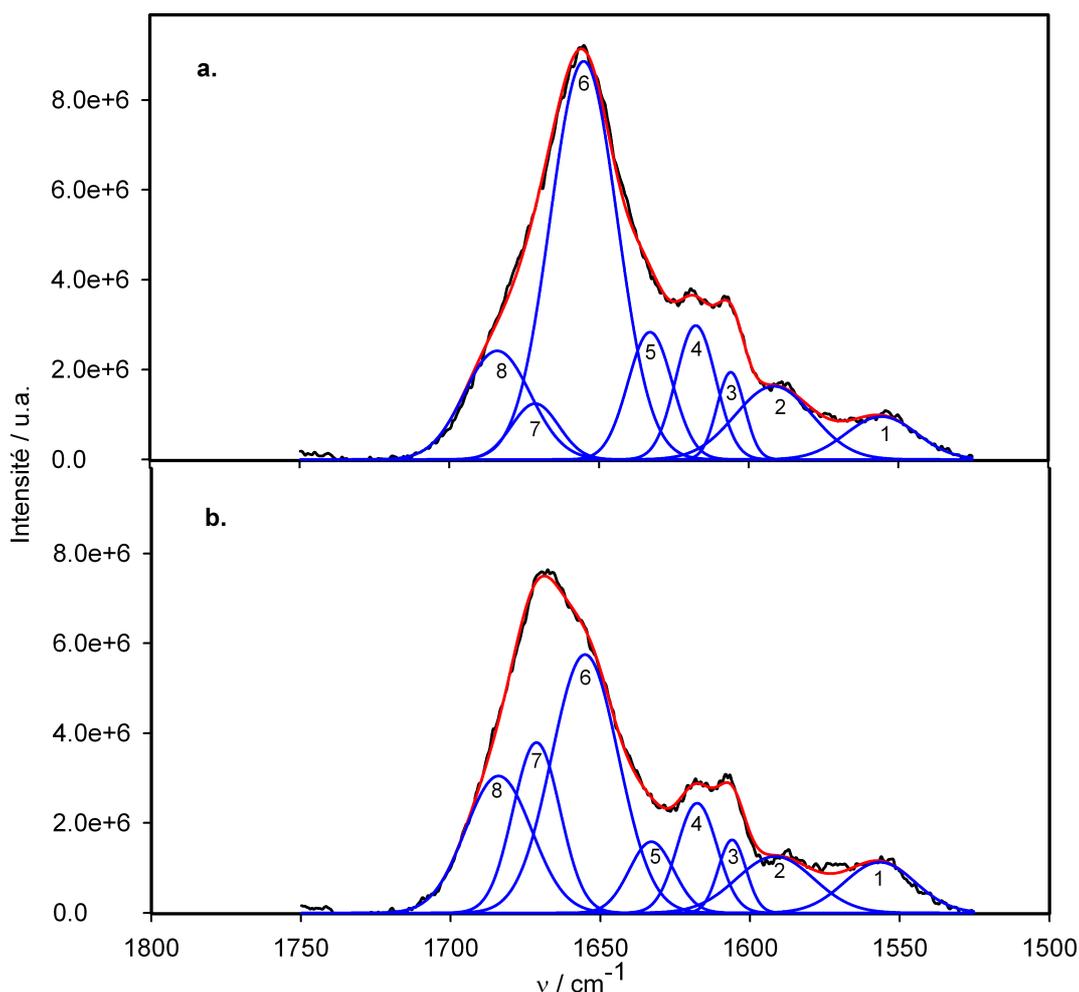


FIGURE 4.8 – Analyse de la bande Raman amide I d’une solution aqueuse de BSA à 1 mM (a) et d’une solution mixte : BSA à 1 mM et DTT à 30 mM (b). Le pH des solutions est de 7,8.

à 1630 cm^{-1} a été attribuée aux chaînes courtes liant les segments hélicoïdaux [6].

L’addition du DTT dans la solution aqueuse de BSA ou de HSA, affecte la bande amide I d’une manière évidente (Figures 4.8b et 4.9b). Cette évolution consiste en une diminution des bandes 5 et 6 et une augmentation des bandes 7 et 8. De ce fait, les changements de la structure secondaire de la protéine peuvent être représentés d’une manière approximative par la diminution de la fraction des hélices α orientées (la bande 6) et l’augmentation de la fraction des hélices α désorientées (la bande 7). Par conséquent, les changements structuraux qui accompagnent la réaction de réduction de ponts S-S affectent principalement les structures du type hélice α , probablement en modifiant leur pas. Nous notons aussi la grande similitude qui existe entre les changements observés dans les deux cas : celui de la BSA et celui de la HSA.

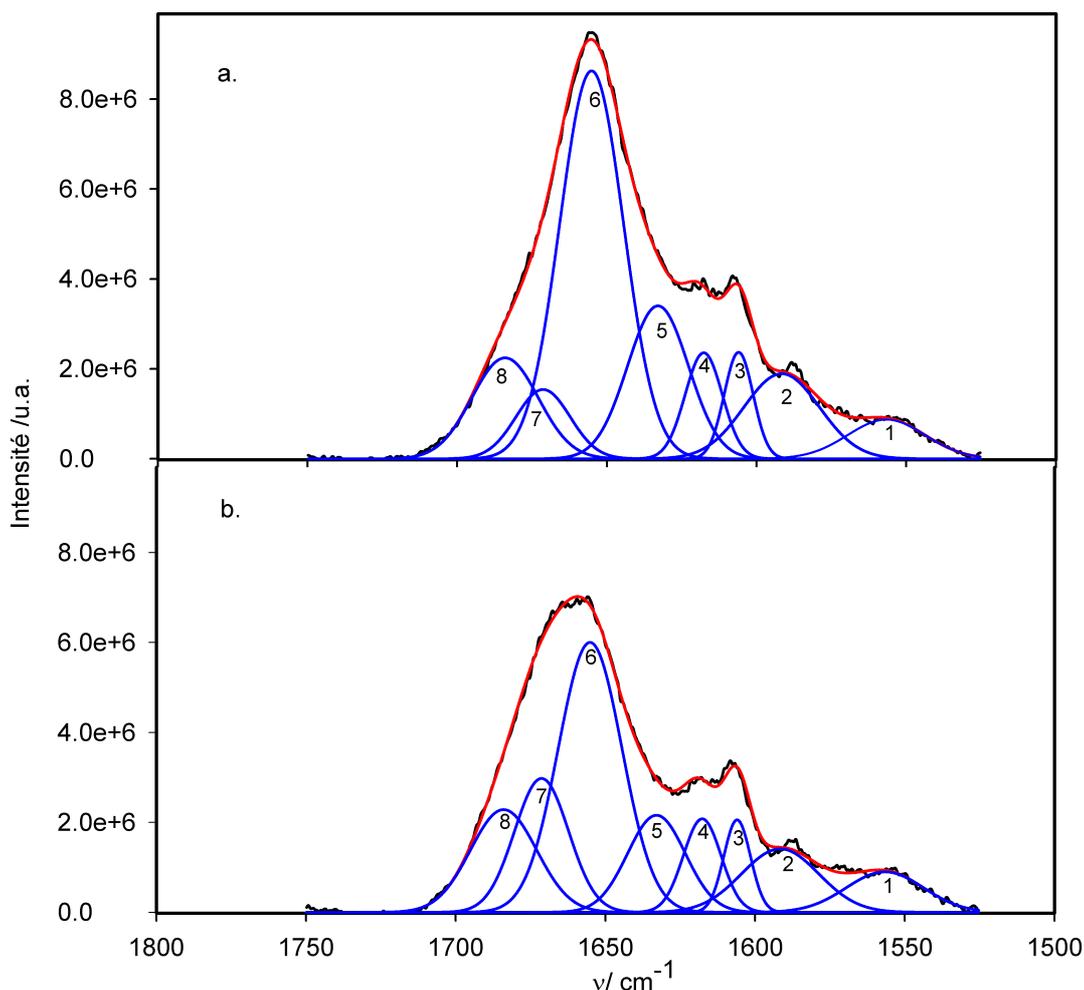


FIGURE 4.9 – Analyse de la bande Raman amide I d’une solution aqueuse de HSA à 1 mM (a) et d’une solution mixte : HSA à 1 mM et DTT à 30 mM (b). Le pH des solutions est de 7,8.

Le fait que la réduction des ponts S-S soit une réaction d’équilibre signifie qu’à des concentrations suffisamment importantes de DTT réduit, les ponts S-S sont coupés et reformés d’une manière continue. Dans ces conditions, nous pouvons nous demander si parmi les ponts S-S reformés, il se trouve aussi des ponts qui ne sont pas natifs. Si cela avait été le cas, le processus de dépliement de la protéine aurait surpassé celui de la réduction des ponts, puisque la formation des ponts non-natifs ne pourrait pas induire de processus de repliement inverse. Conformément aux remarques précédentes, le repliement de la protéine pourrait être quantifié par le rapport entre la surface de la bande 6 (ho) et celle de la bande 7 (hu). La dépendance de ce rapport en fonction de la concentration de DTT ajouté est représentée sur la Figure 4.10.

Ces derniers résultats mettent en évidence le fait que le processus de dépliement suit de

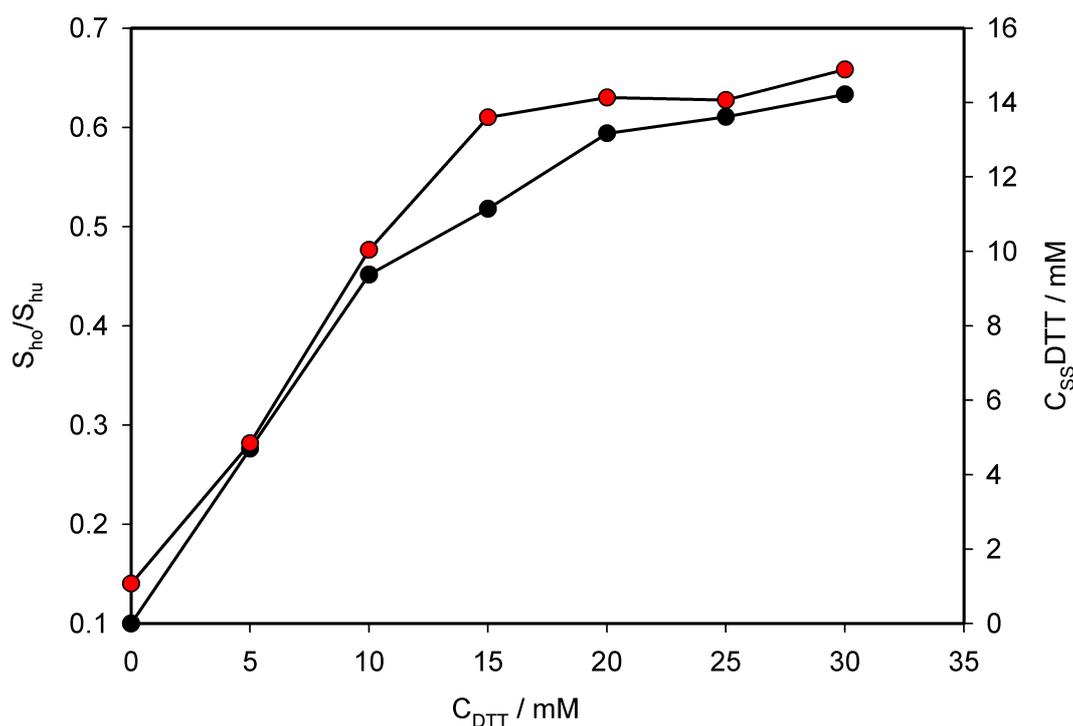


FIGURE 4.10 – Dépendance du rapport des surfaces des bandes 6 et 7 en fonction de la concentration de DTT ajouté (en rouge) comparée avec la dépendance de la concentration des ponts S-S formés dans le DTT en fonction du même paramètre (en noir).

près la réduction des ponts S-S. Nous avons aussi conclu que pour les concentrations de DTT réduit utilisées dans cette étude, la fraction de ponts S-S non-natifs éventuellement formés n'est pas détectable dans les spectres Raman.

4.2.1.3 Cinétiques de réaction

Étant donné que la réaction d'échange de ponts S-S implique la participation du thiol déprotoné (thiolate), la constante de vitesse de réaction observée k_f^{obs} est évidemment dépendante de pH. De ce fait, la vitesse effective de la réaction peut être convenablement ajustée en choisissant le pH de l'échantillon. La valeur de pH de 7,8 a été jugée satisfaisante pour suivre la réaction de réduction des ponts par la technique de spectroscopie Raman. Une solution aqueuse mère de 1 mM de BSA et 17 mM de DTT a été préparée à pH=7,8 et maintenue à 22°C. Des échantillons ont été prélevés et mesurés à des instants différents dans un intervalle de temps de 3 heures après la préparation de la solution mère. La dépendance de la concentration du DTT oxydé en fonction du temps obtenue en analysant

les bandes Raman S-S est présentée sur la Figure 4.11.

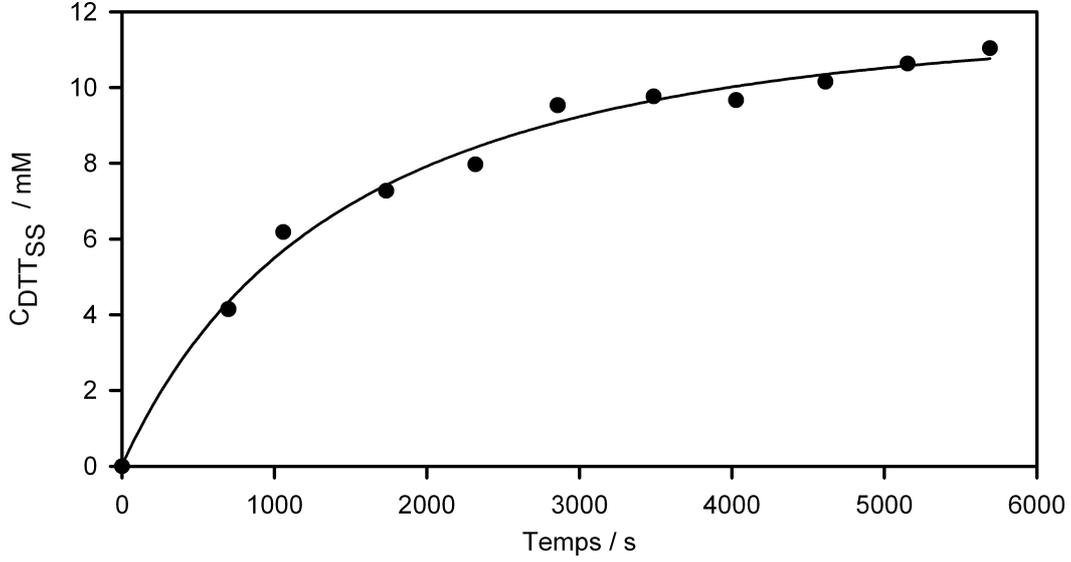


FIGURE 4.11 – Cinétique de réduction des ponts S-S mesurée pour une solution de BSA à 1 mM et de DTT à 17 mM à pH 7,8 et à la température de 22°C.

La fonction théorique correspondante est la solution de l'équation différentielle suivante :

$$\frac{d[DTT_{SS}]}{dt} = k_f^{obs}[DTT_{SH,SH}][(Prot_{SS})_N] - k_b^{obs}[DTT_{SS}][(Prot_{SH,SH})_N] \quad (4.8)$$

Les concentrations de ponts S-S et de thiols de la protéine qui apparaissent dans l'équation 4.8 concernent seulement les ponts S-S réactifs. Entre les constantes de vitesse de la réaction directe (k_f^{obs}) et de la réaction inverse (k_b^{obs}) il existe la relation suivante :

$$\frac{k_f^{obs}}{k_b^{obs}} = K_{eq}^{obs} \quad (4.9)$$

L'équation 4.9 a été obtenue en imposant la condition d'équilibre $\frac{d[DTT_{SS}]}{dt} = 0$ dans l'équation 4.8. L'intégration de l'équation 4.8 conduit à (voir l'Annexe C) :

$$[DTT_{SS}](t) = \frac{X_1 X_2 [e^{(k'(X_1 - X_2)t)} - 1]}{X_1 e^{(k'(X_1 - X_2)t)} - X_2} \quad (4.10)$$

où

$$k' = \frac{k_f^{obs}}{1 + 2 \times 10^{(pH - pK_{a1})} + 10^{(2pH - pK_{a1} - pK_{a2})}} \times \left(1 - \frac{\alpha}{K_{eq}^{obs}}\right) \quad (4.11)$$

et où X_1 et X_2 sont les solutions de l'équation :

$$([DTT_{SH,SH}]_I - X)[(Prot_{SS})_N]_I - X) - \frac{\alpha}{K_{eq}^{obs}} X^2 = 0 \quad (4.12)$$

L'équation 4.10 a ensuite été utilisée pour modéliser les données expérimentales sur la Figure 4.11, en prenant $K_{eq}^{obs} = 2$, $[(Prot_{SS})_N]_I = 14 \text{ mM}$, $[DTT_{SH,SH}]_I = 17 \text{ mM}$. La modélisation a mis en évidence une constante de vitesse de la réaction directe, $(k_f^{obs}) = (2,8 \pm 0,3) \times 10^{-2} M^{-1} s^{-1}$.

La relation entre la constante de vitesse de réaction observée et la constante de vitesse intrinsèque de réaction a été obtenue en exprimant l'équation 4.8 en fonction des réactifs déprotonés :

$$\frac{d[DTT_{SS}]}{dt} = k_f([DTT_{SH,S^-}])[(Prot_{SS})_N] - k_b[DTT_{SS}][(prot_{SH,S^-})] \quad (4.13)$$

Puisque $[DTT_{SH,S^-}] = 2[DTT_{SH,SH}] \times 10^{(pH-pK_{a1})}$ on obtient :

$$k_f^{obs} = 2k_f \times (10^{(pH-pK_{a1})}) \quad (4.14)$$

Cette équation nous a permis de trouver, pour la constante de vitesse intrinsèque k_f , une valeur de $0,68 \pm 0,11 M^{-1} s^{-1}$.

Il est intéressant de constater que les données expérimentales ont pu être correctement modélisées en utilisant pour $[(Prot_{SS})_N]_I$ la valeur déterminée par l'étude de la réaction d'équilibre. Cela nous indique que les cinétiques de réaction ne présentent pas une composante plus rapide qui ne serait pas détectable par notre système de mesure. En outre, nous avons trouvé qu'il n'existe aucune composante lente, puisque pour une concentration de DTT ajouté de 17 mM, la limite asymptotique de la courbe théorique (Figure 4.11) est égale à la concentration d'équilibre de DTT oxydé obtenue à 17 mM de DTT ajouté (Figure 4.6).

Les résultats obtenus dans l'étude de l'équilibre de réaction ont montré que la réduction des ponts S-S et le dépliement qui accompagne ce processus sont corrélés. Si le dépliement de la protéine suit la réduction dans un délai détectable, alors les cinétiques de réaction pourront aussi être étudiées en analysant l'évolution de la bande amide I en fonction du temps (Figure 4.12). Étant donné que la différence entre les surfaces des bandes 6 (ho) et 7 (hu) diminuent proportionnellement avec la concentration de ponts S-S réduits, l'équation qui va modéliser les données expérimentales s'écrit :

$$(S_{ho} - S_{hu})(t) = a - b[DTT_{SS}](t) \quad (4.15)$$

où a et b sont des constantes et $[DTT_{SS}]$ est donnée par l'équation 4.10. La modélisation des données expérimentales par cette équation conduit à une constante de vitesse intrinsèque, k_f , de $0,96 \pm 0,07 M^{-1} s^{-1}$. Ce résultat est en bon accord avec celui obtenu par l'étude des cinétiques basée sur la détection des bandes Raman S-S. Il indique que le

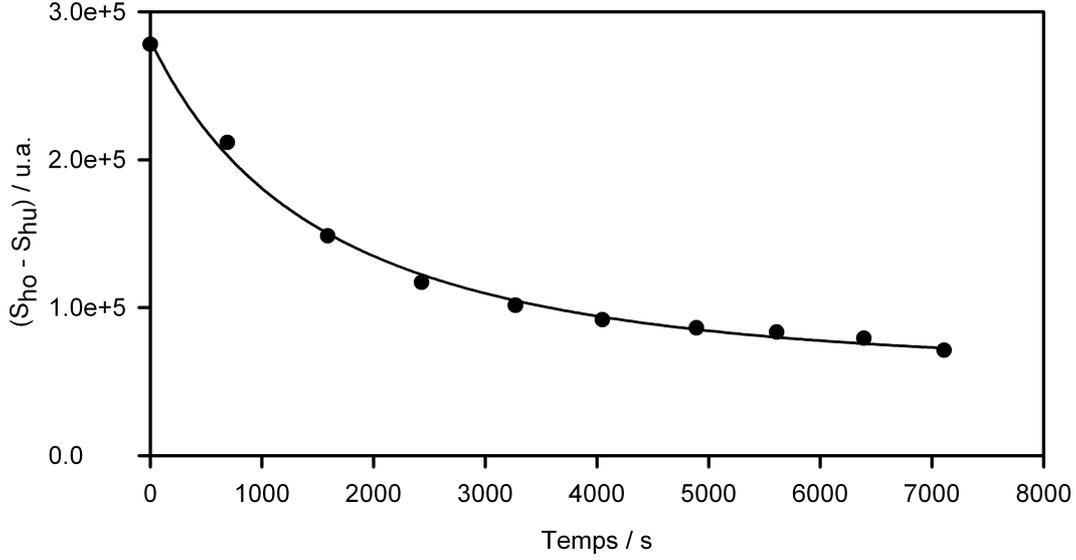
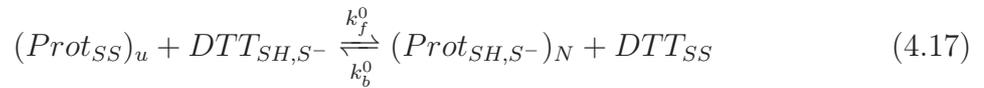


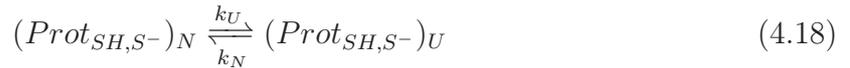
FIGURE 4.12 – Cinétique de réduction des ponts S-S mesurée pour une solution de BSA à 1 mM et de DTT à 17 mM à pH 7,8 et à la température de 22°C : dépendance de la différence des surfaces des bandes 6(ho) et 7(hu) en fonction du temps.

déplieement de la protéine suit la réaction de réduction des ponts S-S dans un délai qui n'est pas détectable. En effet, on s'attend à ce que le processus de déplieement lors de la réduction des ponts S-S soit un processus rapide. Les simulations de dynamique moléculaire effectuées sur la lysozyme réduite ont prédit une durée caractéristique d'environ 1 ns pour ce type de processus [113].

La réaction globale de réduction de ponts S-S des protéines peut être décrite par un modèle simple. La première étape est un déplieement local qui est suivi par la réduction du pont S-S ainsi exposé au solvant :



Les indices u et n indiquent respectivement les structures dépliées ou repliées localement. La dernière étape de la réaction est le déplieement global de la structure protéique réduite :



Conformément à la discussion antérieure, ce dernier pas est très rapide et de ce fait la constante de vitesse intrinsèque globale, k_f , ne dépend que des deux premiers pas de la réaction (équations 4.16 et 4.17) :

$$k_f = \frac{k_u}{k_n + 2k_f^0[DTT_{SH,S^-}]} k_f^0 \quad (4.19)$$

où k_f^0 est la constante intrinsèque de vitesse de réaction pour la réduction d'un pont S-S complètement exposé au solvant, k_u est la constante de vitesse de dépliement local et k_n est la constante de vitesse du processus de repliement. L'équation 4.19 a été obtenue à partir de l'équation de vitesse écrite pour $[(Prot_{SS})_u]$ dans les conditions stationnaires. En supposant que $k_n \gg k_f^0[DTT_{SH,S-}]$ (le processus de repliement se produit en quelques picosecondes), on obtient :

$$k_f = \frac{k_u}{k_n} k_f^0 \quad (4.20)$$

Le rapport $\frac{k_u}{k_n}$ est directement lié à la différence d'enthalpie libre, ΔG^0 , entre l'état "déplié" et l'état "plié" :

$$\Delta G^0 = -RT \ln \frac{k_u}{k_n} = RT \ln \frac{k_f}{k_f^0} \quad (4.21)$$

Afin d'évaluer ΔG^0 nous avons besoin de connaître la constante k_f^0 . Elle peut être définie comme étant la constante de vitesse pour la réduction par le DTT d'une molécule modèle ayant un pont S-S exposé. Un bon modèle dans ce sens est la molécule de glutathion oxydé (GSSG). La constante de vitesse de la réaction d'échange de ponts S-S entre le GSSG et le DTT a été déterminée par la même méthode expérimentale. Les mesures ont été effectuées sur une solution de GSSG à 15 mM et de DTT à 15 mM à pH 5,9 et à la température de 22°C. Dans le cas où la concentration de DTT ajouté est égale à la concentration initiale de GSSG, la dépendance théorique de la concentration de DTT oxydé en fonction du temps s'exprime de la manière suivante :

$$[DTT_{SS}](t) = c_0 - \frac{c_0}{c_0 k_f^{0,obs} t + 1} \quad (4.22)$$

où c_0 est la concentration de DTT ajouté et les constantes k_f^0 et $k_f^{0,obs}$ sont reliées par l'équation 4.14. La modélisation des données expérimentales (Figure 4.13) par l'équation 4.22 conduit à une constante de vitesse intrinsèque de réaction, k_f^0 , de $71 \pm 10 M^{-1}s^{-1}$. Cette valeur est légèrement supérieure à celle rapportée dans la littérature ($51 \pm 8 M^{-1}s^{-1}$) [3].

À partir de la valeur de k_f^0 trouvée, les équations 4.20 et 4.21 donnent une enthalpie libre de dépliement local de la BSA, $\Delta G^0 = 11,6 kJ.mol^{-1}$. Visiblement, cette enthalpie libre est une moyenne faite sur tous les ponts S-S réactifs de la BSA. Comme on pourrait s'attendre, elle est plus faible que les valeurs données dans la littérature pour l'enthalpie libre de dépliement global ($18-40 kJ.mol^{-1}$) [10, 114–116]. D'autre part, la valeur ΔG^0 calculée pour la BSA est 2,6 fois plus faible que celle rapportée dans la littérature pour les deux ponts S-S les moins réactifs du BPTI [117]. Cela suggère que les ponts S-S de la BSA sont en moyenne plus exposés au solvant. Enfin, le fait que les cinétiques étudiées aient

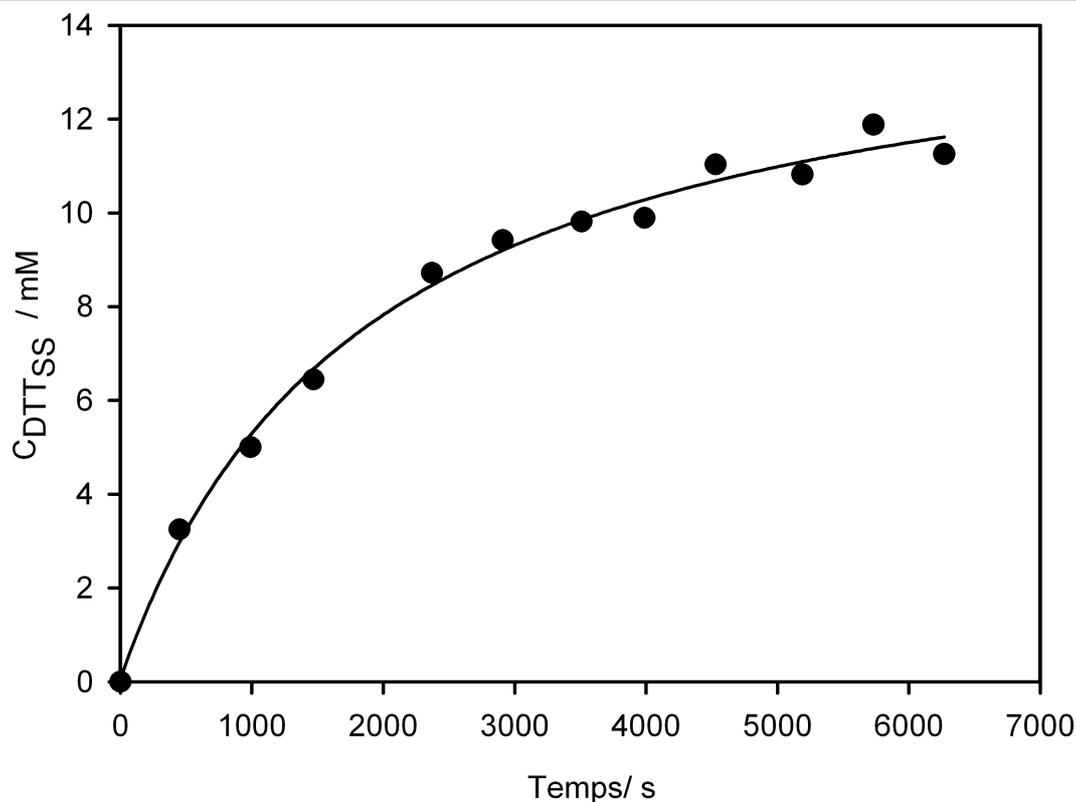


FIGURE 4.13 – Cinétique de réduction du GSSG par le DTT. Les mesures des spectres Raman ont été effectuées sur des solutions aqueuses contenant 15 mM de GSSG et 15 mM de DTT à pH 5,9 et à la température de 22°C.

pu être modélisées d'une manière satisfaisante en utilisant une seule constante de vitesse, nous indique que que l'environnement des 14 ponts S-S réactifs de la BSA est relativement uniforme.

Il est aussi intéressant de noter que la constante de vitesse de réaction moyenne k_f , que nous avons trouvée pour la réduction de la BSA par le DTT, a le même ordre de grandeur que celle rapportée pour la réduction de trois (parmi quatre) ponts S-S de l' α -lactalbumine ($3,7 \pm 10 M^{-1}s^{-1}$) [19]. La constante de vitesse pour la réduction du quatrième pont S-S de l' α -lactalbumine est beaucoup plus grande : $(4,56 \pm 0,05) \times 10^3 M^{-1}s^{-1}$. D'autre part, la valeur de la constante de vitesse k_f rapportée pour la réduction des ponts les plus réactifs de la RNase par le DTT est de $6,3 \times 10^{-4} M^{-1}s^{-1}$ [50]. Il est peu probable que les différences d'enthalpie libre de dépliement soient responsables de cette dispersion des valeurs entre les différents ponts et les différentes protéines. Une explication alternative doit être cherchée du côté des interactions électrostatiques. En effet, il a été montré que les groupements chargés positivement qui se trouvent proche du site réactionnel, peuvent influencer d'une manière critique les constantes de vitesse par l'augmentation locale de la

concentration effective du thiolate [118–120].

4.2.1.4 Évaluation de la variation d’enthalpie libre conformationnelle

La réactivité des ponts S-S dans les protéines avec le DTT est directement liée à l’enthalpie libre de la réaction d’échange des ponts. Afin d’identifier les facteurs qui influencent cette réactivité, nous avons décomposé l’enthalpie libre de réaction en plusieurs contributions : la contribution locale ΔG_{locale} , la contribution conformationnelle ΔG_{conf} et la contribution responsable de la variation d’enthalpie libre lors de l’oxydation du DTT en solution aqueuse :

$$\Delta G \cong \Delta G_{DTT_{SH,SH}/DTT_{SS}} + \Delta G_{locale} + \Delta G_{conf} \quad (4.23)$$

où

$$\Delta G_{DTT_{SH,SH}/DTT_{SS}} = G_{aq}(DTT_{SH,SH}) - G_{aq}(DTT_{SS}) \quad (4.24)$$

La contribution locale regroupe les variations de l’enthalpie libre électronique et vibrationnelle lors de la coupure d’un pont S-S. La contribution conformationnelle représente la variation d’enthalpie libre due au dépliement de la protéine qui accompagne le processus de réduction. Elle inclut aussi la contribution de l’interaction entre les atomes du groupement local et leur environnement, la chaîne protéique et le solvant. Évidemment, la variation d’énergie de solvatation de la protéine lors du changement conformationnel fait partie de la contribution conformationnelle, à l’enthalpie libre.

Les termes $\Delta G_{DTT_{SH,SH}/DTT_{SS}}$ et ΔG_{locale} ont été évalués par calcul *ab initio* en utilisant la méthode DFT. Les réactifs et les produits ont été optimisés avec le potentiel hybride B3LYP et le jeu de fonctions de base 6-311+G(d,p). Les structures ainsi optimisées ont fait l’objet d’un calcul d’énergie potentielle avec un jeu de fonctions de base plus large 6-311+G(2df,2pd). Les corrections thermiques à l’enthalpie libre, ΔG^{corr} (équation 2.26), ont été obtenues par des calculs de fréquences au niveau de théorie B3LYP/6-311+G(d,p).

La variation d’enthalpie libre locale ΔG_{locale} a été évaluée en utilisant un pont S-S formé par l’assemblage de deux résidus de cystéine modèle. Ces résidus modèles ont été construits à partir de la structure de la L-cystine, en remplaçant les groupes terminaux $-NH_3^+$ et $-COO^-$ par les groupes $-NH - CO - CH_3$ et $-CO - NH - CH_3$ respectivement (figure 4.14). Étant donné que les enthalpies libres de rotation, de translation et de solvatation ne contribuent pas à $\Delta G_{(locale)}$ mais sont incluses dans l’enthalpie libre

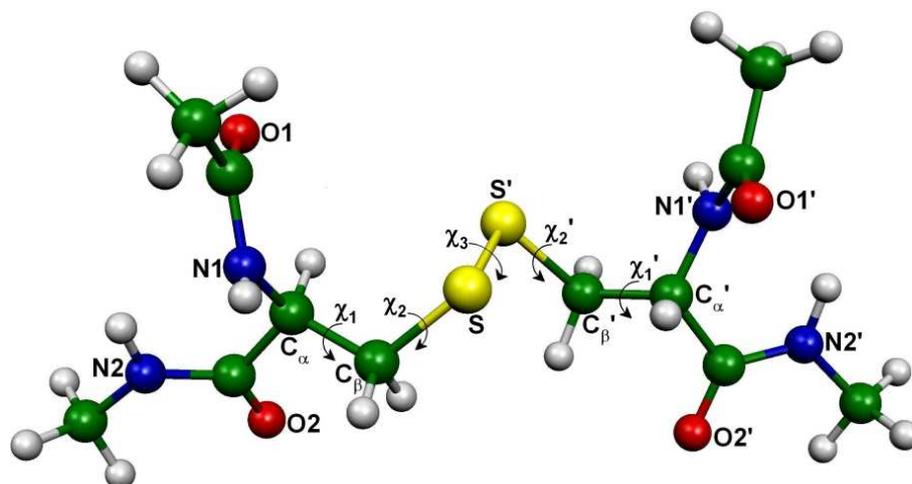


FIGURE 4.14 – Modèle de pont S-S dans la protéine optimisé au niveau de théorie B3LYP/6-311+G(d,p). Les angles dièdres qui caractérisent la conformation du pont sont : $\chi_1 = -61,1^\circ$, $\chi_2 = -71,0^\circ$, $\chi_3 = -89,3^\circ$, $\chi_1' = -61,1^\circ$, $\chi_2' = -70,9^\circ$.

conformationnelle, on peut écrire :

$$\Delta G_{locale} = 2[G_g(cys_{SH,mod}) - G_{tr,rot}(cys_{SH,mod})] - [G_g(cys_{SS,mod}) - G_{tr,rot}(cys_{SS,mod})] \quad (4.25)$$

où $cys_{SS,mod}$ représente le fragment présenté dans la Figure 4.14 et $cys_{SH,mod}$ l'un des fragments qui résulte de la coupure du pont S-S. D'autre part, G_g est l'enthalpie libre en phase gazeuse et $G_{tr,rot}$ est la contribution translationnelle et rotationnelle à cette enthalpie. Après substitution des valeurs calculées dans les équations 4.24 et 4.25, on obtient $\Delta G_{DTT_{SH,SH}/DTT_{SS}} + \Delta G_{locale} = 25 \text{ kJ.mol}^{-1}$. Il est prévisible que la coupure d'un pont S-S dans la protéine conduira à une variation d'enthalpie libre plus faible que celle-ci. En effet, les possibles contraintes imposées par la structure globale de la protéine peuvent induire des distorsions qui déstabiliseront le pont. Cette énergie de distorsion pourrait être exprimée sous une forme empirique basée sur les valeurs des angles dièdres autour des liaisons $C - C$ et $S - S$ [19, 34] bien que du point de vue de la mécanique quantique elle a ses origines dans la répulsion électronique entre les liaisons voisines. Ainsi, nous avons effectué des analyses *ab initio* sur la coupure du pont Cys75-Cys91 de la HSA (Figure 4.15). Le pont analysé a été extrait de la structure de diffraction de rayons X de la HSA et ensuite réoptimisé. Afin de bien préserver la conformation de l'angle dièdre $-C - C - S - S - C - C-$ du pont, les quatre groupements méthyle ont été fixés pendant la réoptimisation. Les mêmes contraintes ont été imposées pour l'optimisation des fragments $cys_{SH,mod}$ résultant de la coupure du pont. Dans l'évaluation de l'enthalpie libre (équation 2.25) des structures contraintes nous avons utilisé le même ΔG^{corr} que pour les structures

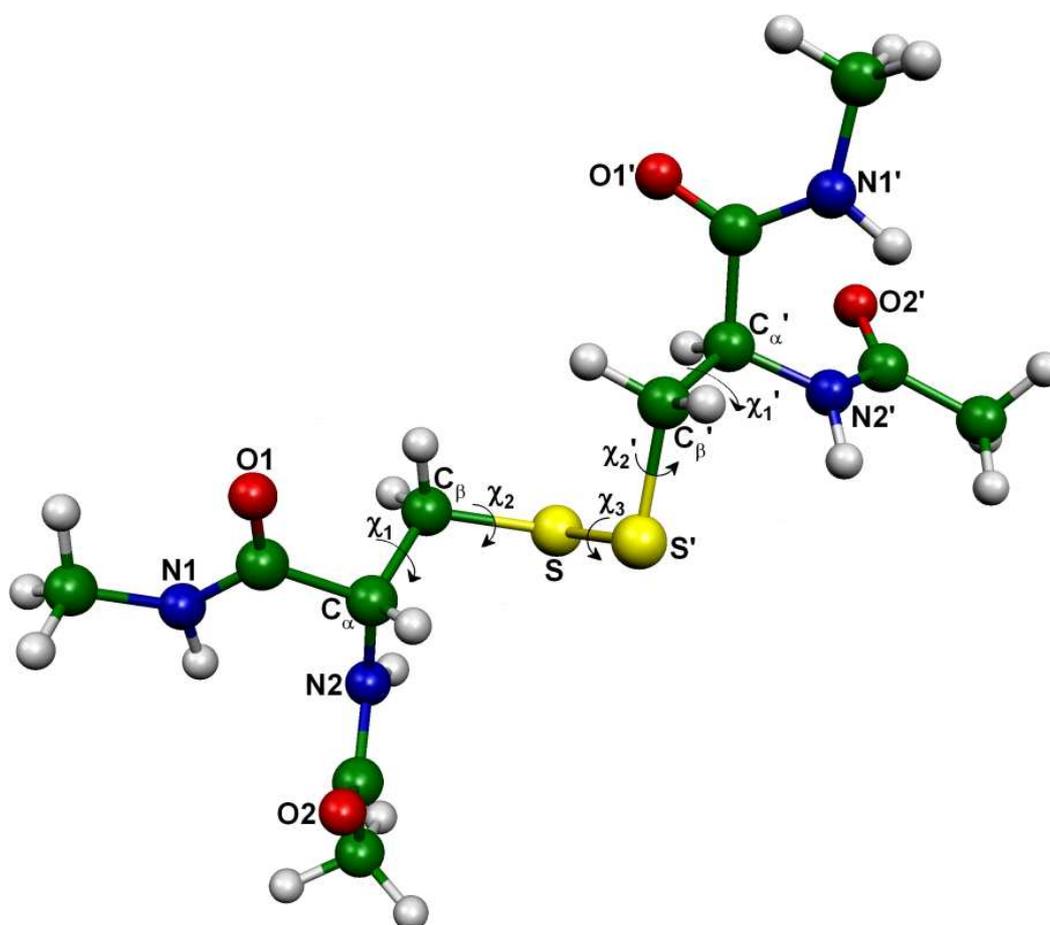


FIGURE 4.15 – Pont Cys75-Cys91 de la HSA réoptimisé au niveau de théorie B3LYP/6-311+G(d,p) en fixant les groupements méthyle. Les angles dièdres qui caractérisent la conformation du pont sont : $\chi_1 = -63,4^\circ$, $\chi_2 = -68,1^\circ$, $\chi_3 = -83,3^\circ$, $\chi'_1 = -59,0^\circ$, $\chi'_2 = -64,1^\circ$.

non-contraintes. Cette approximation n'affecte pas d'une manière significative l'enthalpie libre de réaction car elle ne concerne que les extrémités du système. La variation d'enthalpie libre calculée, $\Delta G_{DTT_{SH,SH}/DTT_{SS}} + \Delta G_{locale}$, pour la réduction du pont Cys75-Cys91 de la HSA par DTT est de $4,0 \text{ kJ.mol}^{-1}$. On observe que cette valeur est bien plus faible que celle trouvée pour le pont modèle. Etant donné que les deux structures (Figures 4.14 et 4.15) n'ont pas la même conformation des angles dièdres autour des liaisons C-S, nous pouvons attribuer cette différence à l'énergie de distorsion du pont.

Comme il a été montré ci-dessus, l'analyse des équilibres de réaction conduit à une valeur expérimentale de l'enthalpie libre de réduction d'un pont S-S d'environ -2 kJ.mol^{-1} , autant pour la BSA que pour la HSA. L'utilisation de l'équation 4.23 et de la valeur théorique des termes $\Delta G_{DTT_{SH,SH}/DTT_{SS}} + \Delta G_{locale}$ (25 kJ.mol^{-1}), nous a permis d'es-

timer ΔG_{conf}^{HSA} à -27 kJ.mol^{-1} . Évidemment, la contribution conformationnelle doit être négative. En effet, l'enthalpie libre conformationnelle pour une protéine ayant les ponts S-S coupés représente un minimum par rapport à l'espace configurationnel entier de la protéine, y compris la protéine avec les ponts intacts.

Ces résultats mettent en évidence le rôle essentiel de la contribution conformationnelle dans la réactivité des ponts S-S : une augmentation de cette contribution de quelques kJ.mol^{-1} seulement peut affecter de manière spectaculaire la réactivité d'un pont S-S dans les albumines sériques. Il est donc probable que la réduction de quelques ponts S-S de la BSA (ou la HSA) soit associée à des faibles variations d'enthalpie libre conformationnelle, ce qui les rend non-réactifs. Dans cette approche, les ponts S-S qui sont le plus exposés au solvant ne seraient pas nécessairement les plus réactifs.

Cependant, les valeurs obtenues pour ΔG_{conf} dans la présente étude dépendent d'une manière critique de la valeur choisie pour la somme $\Delta G_{DTT_{SH,SH}/DTT_{SS}} + \Delta G_{locale}$. Une meilleure estimation de cette valeur pourrait être obtenue après un évaluation plus spécifique de la ΔG_{locale} . Cela serait envisageable en effectuant des calculs *ab initio* sur les 17 ponts S-S des albumines sériques.

4.2.2 Réaction avec la TCEP

La réduction des ponts S-S des protéines par la TCEP présente l'avantage, par rapport à la réduction par le DTT, d'être une réaction totale. D'autre part, la taille de la molécule de TCEP est supérieure à celle du DTT, ce qui pourrait rendre les ponts S-S moins accessibles pour cette réaction. Enfin, le fait que la TCEP se présente sous plusieurs états de protonation actifs dans la réaction, rend l'analyse de la cinétique de réaction plus compliquée. Cependant, du point de vue technique, le suivi de la réaction par la détection des bandes Raman S-S est plus facile, étant donné que la TCEP et la TCEP oxydé ne possèdent pas de bandes Raman intenses dans le domaine spectral respectif.

Cinétiques de réaction

Comme dans le cas de la réduction des ponts S-S par le DTT, l'avancement de la réaction a été suivi en analysant les modifications des bandes S-S de la protéine. La solution aqueuse mère de BSA (ou HSA) à 0,5 mM et TCEP à 20 mM a été préparée à pH 2,0 et à la température de 22°C. Le pH des solutions a été ajusté par l'addition de quelques gouttes

de solution concentrée de HCl ou de NaOH. Cette valeur de pH a été choisie pour obtenir à une vitesse de réaction apparente convenable, qui permette une bonne acquisition des spectres. Les échantillons à mesurer ont été prélevés à des instants couvrant un intervalle temporel de 7 heures qui débute au moment du mélange des réactifs. La stabilité de la TCEP réduit libre en solution aqueuse a été testée en mesurant sa bande Raman à 2468 cm^{-1} . Au bout de 8 heures nous n'avons pas détecté de changement d'amplitude de cette bande.

Les espèces actives de TCEP sont celles déprotonées sur l'atome de phosphore. La vitesse de réaction peut être donc ajustée en modifiant la concentration de ces espèces par un changement de pH. D'autre part, la détermination des constantes intrinsèques de réaction à partir de constantes de vitesse apparentes exige une connaissance précise de toutes les valeurs de pKa de la TCEP. Des études rapportées dans la littérature montrent que le pKa du phosphore (pKa_P) est fortement affecté par la déprotonation des groupements carboxyle de la TCEP [109]. Puisque les trois valeurs de pKa des groupements carboxyle (pKa_1, pKa_2, pKa_3) sont inférieures à la valeur du pKa_P , il n'est pas possible de mesurer directement le pKa_P des espèces protonées sur le groupement carboxyle de la TCEP. L'alternative est de faire ces mesures sur des composés modèles dans lesquels les groupements carboxyle ont été substitués par des groupements méthyle. Les valeurs du pKa_P mesurées pour le triméthyle-TCEP, le diméthyle-TCEP et le méthyle-TCEP [109] permettent de déduire les valeurs du pKa_P pour les espèces de TCEP ayant trois, deux et un groupement carboxyle protoné : 4,7 ; 5,9 et 6,6, respectivement. Dans le cas de la TCEP ayant les trois groupements carboxyle déprotonés, Cline et al. [109] ont rapporté une valeur du pKa_P de 7,5. La coexistence des plusieurs espèces actives de TCEP dans un domaine de pH acide est une conséquence directe de l'augmentation du pKa_P lors de l'ionisation des groupements carboxyle. Dans notre étude ces espèces actives ont été désignées de la manière suivante : TCEPH3, TCEPH2, TCEPH1 et TCEPH0, avec TCEPH0 pour désigner la forme complètement déprotonée. Parmi ces espèces, la TCEPH3 est neutre et les autres sont chargées négativement. On s'attend à ce que ces espèces soient caractérisées par des constantes de vitesse intrinsèques différentes. En effet, il a été déjà rapporté que les interactions électrostatiques peuvent fortement affecter les constantes de vitesse de réaction des protéines avec des réactifs chargés [120].

Pour une valeur de pH donnée, la constante apparente de vitesse de réaction s'exprime par l'équation suivante :

$$k_{ap} = w_1k_{f1} + w_2k_{f2} + w_3k_{f3} + w_4k_{f4} \quad (4.26)$$

où k_{fi} et w_i sont les constantes de vitesse intrinsèques et les fractions moléculaires des es-

pèces TCEPH3, TCEPH2, TCEPH1 et TCEPH0, respectivement. Les fractions molaires ont été calculées en tenant compte des trois constantes d'acidité des groupements carboxyle dans la TCEP. Ces valeurs ont été déterminées par titrage pH-métrique, ce qui a donné : $pK_{a1} = 2,9$, $pK_{a2} = 3,3$ et $pK_{a3} = 3,9$. Les calculs montrent qu'à pH=2 (utilisé dans nos mesures) la concentration de l'espèce neutre, TCEPH3, est supérieure de plusieurs ordres de grandeur aux concentrations des autres espèces actives. Nous pouvons donc simplifier l'analyse en considérant que :

$$k_{ap} \approx w_1 k_{f1} \quad (4.27)$$

Les cinétiques expérimentales (Figures 4.16 et 4.17) sont hétérogènes et ont été modélisées en utilisant deux constantes de vitesse pour chacune. Les valeurs de ces constantes

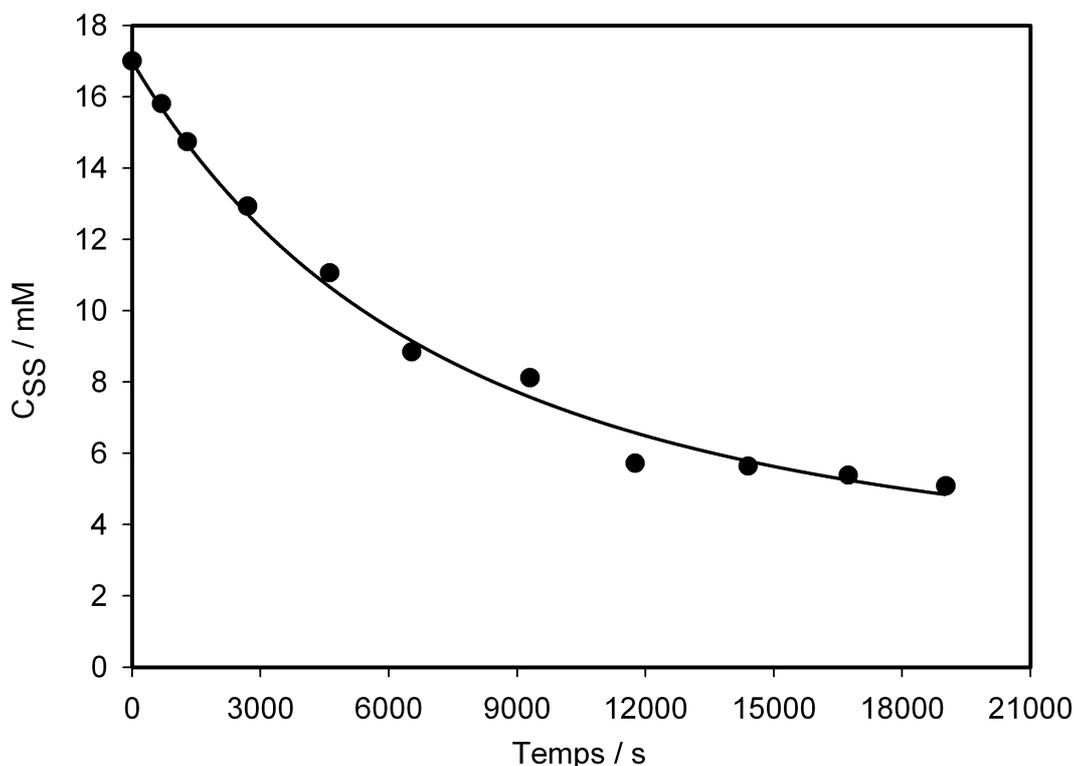


FIGURE 4.16 – Cinétique de la réaction de réduction des ponts S-S dans la BSA par la TCEP mesurée à pH=2,0 et à la température de 22°C. La concentration des ponts S-S a été évaluée à partir des surfaces des bandes Raman S-S. Les mesures ont été effectuées sur une solution aqueuse de BSA à 0,5 mM et TCEP à 20 mM.

sont données dans le Tableau 4.1. Nous constatons que la cinétique de réduction de la HSA est légèrement plus lente que celle de la BSA (à remarquer que les échelles temporelles sur les Figures 4.16 et 4.17 sont légèrement différentes).

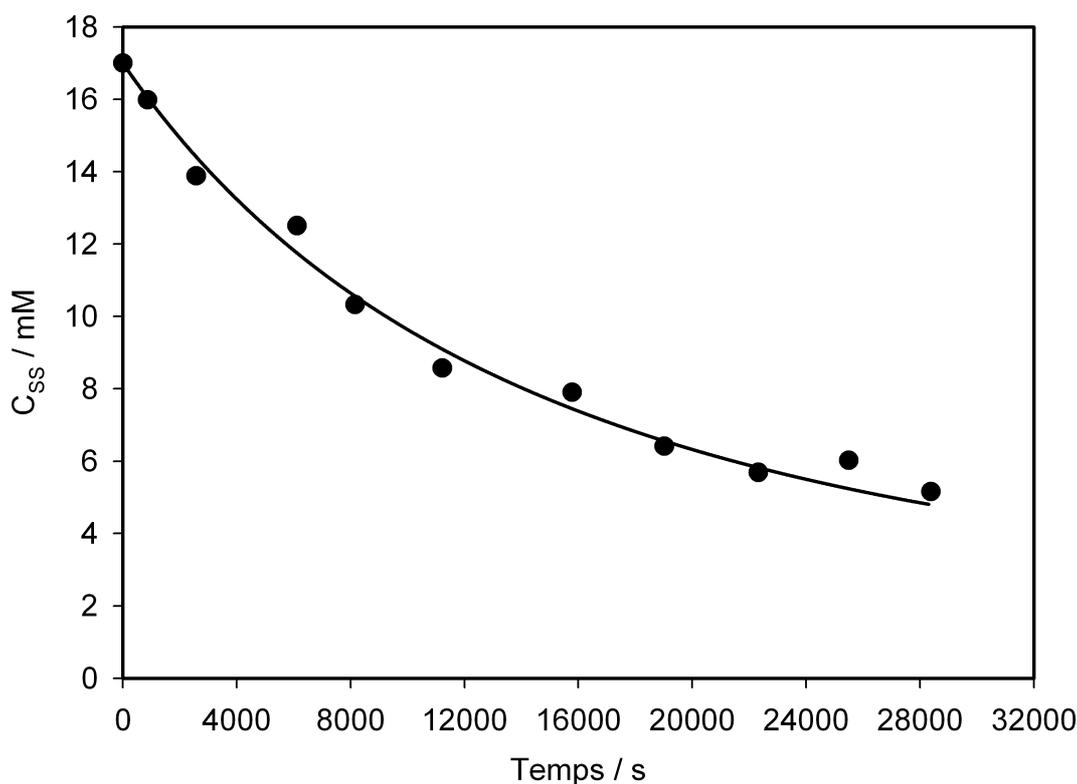


FIGURE 4.17 – Cinétique de la réaction de réduction des ponts S-S dans la HSA par la TCEP mesurée à pH=2,0 et à la température de 22°C. La concentration des ponts S-S a été évaluée à partir des surfaces des bandes Raman S-S. Les mesures ont été effectuées sur une solution aqueuse de BSA à 0,5 mM et TCEP à 20 mM.

	pH	w_1	$k_{ap}^{(1)}/M^{-1}s^{-1}$	$k_{ap}^{(2)}/M^{-1}s^{-1}$
BSA	2,0	$3,5 \times 10^{-5}$	$7,7 \times 10^{-3}$, 76%	$0,4 \times 10^{-3}$, 24%
HSA	2,0	$3,5 \times 10^{-5}$	$4,0 \times 10^{-3}$, 76%	$\leq 0,4 \times 10^{-3}$, 24%

Tableau 4.1 – Constantes apparentes de vitesse de la réaction de réduction des albumines sériques par la TCEP à pH = 2,0. Sont données aussi les fractions molaires w_1 de l'espèce active TCEPH3.

L'existence de différents types d'espèces de TCEP actives n'est pas responsable de l'hétérogénéité de ces cinétiques. En effet, à pH 2,0, les espèces ionisées actives de TCEP sont quasiment inexistantes. De plus, dans le milieu réactionnel, les espèces actives se trouvant en équilibre thermodynamique, elles se comportent comme une seule espèce caractérisée par une constante de vitesse moyenne. L'hétérogénéité de la cinétique indique alors une hétérogénéité de réactivité des ponts S-S.

Les fractions molaires des espèces actives de TCEP et les constantes de vitesse appa-

	TCEP	$k_{f1}^{(1)}/M^{-1}s^{-1}$	$k_{f1}^{(2)}/M^{-1}s^{-1}$
BSA	espèce neutre	4,4	0,22
HSA	espèce neutre	2,28	$\leq 0,22$

Tableau 4.2 – Constantes intrinsèques de vitesse de réaction, rapide ($k_{f1}^{(1)}$) et lente ($k_{f2}^{(2)}$), calculées conformément aux équations 4.28 et 4.29.

	TCEP		DTT
	$\Delta G_u^1/kJ.mol^{-1}$	$\Delta G_u^2/kJ.mol^{-1}$	$\Delta G_u^0/kJ.mol^{-1}$
BSA	3,10	10,58	11,60
HSA	4,74	10,80	-

Tableau 4.3 – Enthalpie libre de dépliement local pour la réaction de réduction des ponts S-S des albumines sériques par la TCEP et par le DTT.

rentes ont été utilisées pour déterminer les constantes de vitesse intrinsèques respectives conformément aux équations :

$$k_{ap}^{(1)} = w_1 k_{f1}^{(1)} \quad (4.28)$$

$$k_{ap}^{(2)} = w_1 k_{f1}^{(2)} \quad (4.29)$$

Les valeurs ainsi trouvées sont données dans le Tableau 4.2. En suivant le même raisonnement que dans le cas de la réduction des ponts S-S par le DTT, nous pouvons évaluer une enthalpie libre de dépliement local en utilisant l'expression :

$$\Delta G_u = -RT \ln \frac{k_f}{k_f^0} \quad (4.30)$$

où la constante de vitesse intrinsèque globale, k_f , doit être remplacée par $k_{f1}^{(1)}$ ou $k_{f2}^{(2)}$. La constante k_f^0 a été déterminée en étudiant la cinétique de la réaction de réduction du pont du GSSG par la TCEP. La valeur obtenue par la modélisation des données expérimentales est de $15,2 M^{-1}s^{-1}$. Cette valeur a ensuite été substituée dans l'équation 4.30 afin d'obtenir le ΔG_u à partir des deux constantes intrinsèques de vitesse, rapide et lente, respectivement. Les résultats de ces calculs sont affichés dans le Tableau 4.3. Contrairement à nos attentes, l'enthalpie de dépliement local de la BSA (la composante rapide) est sensiblement plus faible que celle calculée à partir de la réaction avec le DTT. Cela met en lumière les limitations du modèle de réaction en deux étapes utilisé ici.

4.3 Réduction des ponts S-S dans la lysozyme

La structure de la lysozyme (Hen Egg White Lysozyme, HEWL) avec ses quatre ponts S-S dont un exposé au solvant est illustrée sur la Figure 4.18. À pH neutre et à des

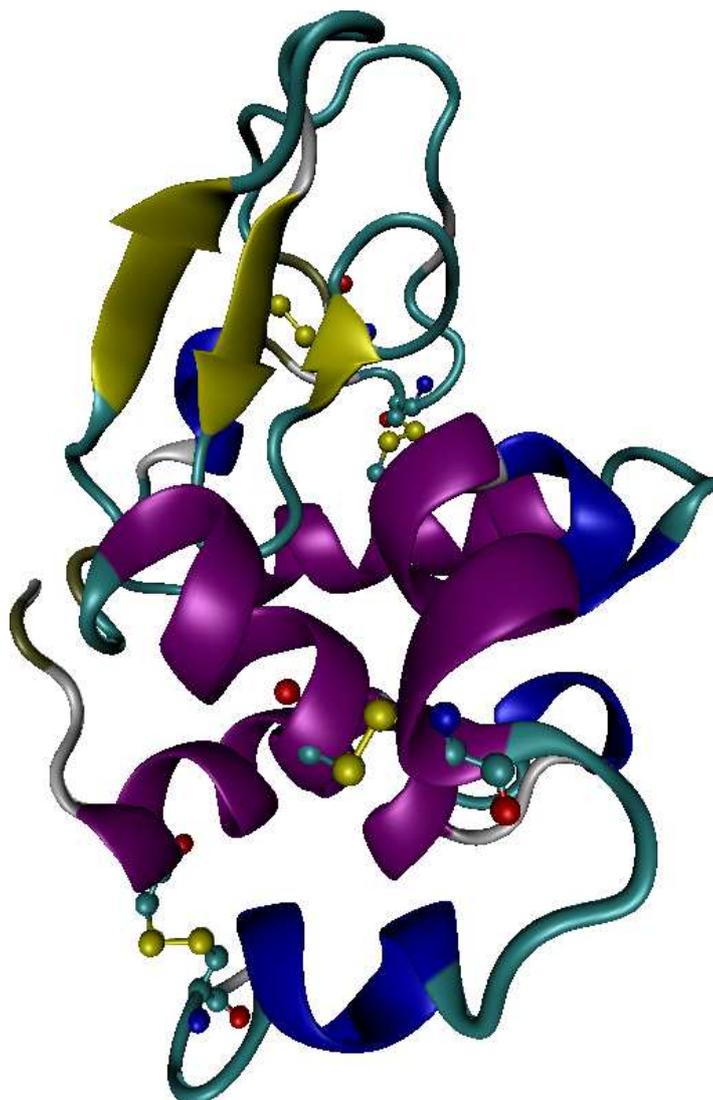


FIGURE 4.18 – Représentation de la structure 3D de la lysozyme (HEWL). Les quatre ponts S-S sont : Cys6-Cys127, Cys30-Cys115, Cys64-Cys80 et Cys76-Cys94.

concentrations de l'ordre de 1 mM, la lysozyme en solution aqueuse précipite. C'est la raison pour laquelle nous n'avons pas pu étudier sa réaction avec le DTT, qui est actif seulement à pH neutre. À pH acide l'agent réducteur idéal est la TCEP.

4.3.1 Aspect thermodynamique de la réaction

Du point de vue thermodynamique, la réduction des ponts S-S est possible dans les cas où l'enthalpie libre de réaction est négative. Il est évident que ce paramètre ne dépend pas seulement des agents réducteurs utilisés mais aussi de l'environnement protéique. On s'attend à ce que la réduction d'un pont S-S dans les molécules simples, comme la cystine, soit plus favorable du point de vue thermodynamique par rapport à la réduction d'un pont S-S dans les protéines. En effet, dans le cas des molécules simples, les produits de réaction sont stabilisés par l'augmentation de l'entropie de translation et de rotation associées à la séparation du système moléculaire initial en deux fragments individuels. Dans le cas de la réduction des ponts S-S des protéines, la stabilisation des produits de réaction peut survenir grâce au dépliement global qui suit la coupure des ponts. Cette contribution est spécifique pour chaque pont et chaque protéine. Un autre facteur qui peut contribuer à l'instabilité des ponts S-S de la protéine est la distorsion de la conformation du pont induite par les contraintes structurales de la protéine.

Afin de caractériser la réactivité thermodynamique des ponts S-S dans la lysozyme, nous avons effectué le même type d'analyse que pour les albumines sériques, c'est à dire que nous avons séparé la variation de l'enthalpie libre en deux contributions, une locale et l'autre conformationnelle. Nous avons ensuite déterminé par calcul de type DFT la variation ΔG_{locale} pour chacun des quatre ponts de la lysozyme. La réaction à considérer dans ce cas est la suivante :



où la TCEP se trouve dans la forme protonée (TCEPH3). Le Cys_{SS} est un modèle de pont S-S très similaire à celui illustré sur la Figure ???. Sa conformation est différente d'un pont à l'autre. On peut donc écrire, comme dans le cas des albumines sériques :

$$\Delta G_{locale} = \Delta G_{mod} - [2G_{trans}(Cys_{SH}) + 2G_{rot}(Cys_{SH}) - G_{trans}(Cys_{SS}) - G_{rot}(Cys_{SS})] \quad (4.32)$$

Ici ΔG_{mod} est la variation d'enthalpie libre totale pour la réaction 4.31. Il est évident que cette réaction a lieu en solution aqueuse. Malgré cela, la variation d'enthalpie de solvatation du pont S-S modèle lors de sa réduction n'est pas incluse dans ΔG_{mod} . En raison de sa forte dépendance de la conformation protéique, elle est incluse dans ΔG_{conf} .

Les valeurs de ΔG_{locale} obtenues pour les quatre ponts de la lysozyme sont consignées dans le Tableau 4.4. On observe que toutes les valeurs sont négatives. Puisque ΔG_{conf} est toujours négative (le dépliement conduit vers des structures plus stables), nous pouvons

Ponts	$\Delta G_{locale}/kJ.mol^{-1}$
cys6-cys127	-27,0
cys30-cys115	-63,0
cys64-cys80	-33,0
cys76-cys94	-30,0

Tableau 4.4 – Variation d’enthalpie libre locale lors de la réduction des ponts S-S de la lysozyme par la TCEP.

affirmer que la réduction de tous les ponts S-S de la lysozyme par la TCEP est une réaction favorable du point de vue thermodynamique et que tous les ponts de la lysozyme sont réactifs avec la TCEP. Cette conclusion est différente par rapport de celle tirée pour la réduction des ponts S-S dans les albumines sériques par le DTT. Dans ce cas nous avons trouvé que la ΔG_{conf} représente une contribution nécessairement critique à la stabilisation des produits de réaction.

Parmi les ponts S-S de la lysozyme, le Cys30-Cys115 se fait remarquer par sa valeur très basse de ΔG_{locale} . De ce fait, il semble être le pont le plus instable. Cette instabilité est certainement due à la distorsion de sa conformation induite par les contraintes extérieures. Cependant, une analyse plus détaillée montre que la distorsion en cause est située au niveau du fragment peptidique porté par la Cys30. Il n’est pas évident que cette distorsion soit récupérée après la coupure du pont et il se peut donc qu’elle ne contribue pas à l’instabilité du pont.

Les valeurs de la ΔG_{locale} dans le Tableau 4.4 indiquent que la réaction de réduction des ponts S-S de la lysozyme par la TCEP est pratiquement irréversible. Il est donc possible de simplifier l’analyse des cinétiques de réaction en négligeant le processus inverse.

4.3.2 Cinétiques de réaction

L’aspect cinétique de la réactivité des ponts avec la TCEP a été abordé au niveau expérimental. Les mesures ont été effectuées sur des solutions aqueuses de lysozyme et $TCEP \bullet HCl$ (TCEP cristallisée avec le HCl). L’évolution des bandes Raman S-S a été suivie sur un intervalle temporel de plusieurs heures après le mélange des réactifs. Les cinétiques de réaction ont été mesurées pour deux valeurs de pH, 3,2 et 4,4, à la température de 22°C. Dans les deux cas, le pH a été ajusté avant le mélange des réactifs

par l'addition des gouttes de solution concentrée de HCl ou de NaOH. La solution obtenue après le mélange des réactifs contenait 4 mM de lysozyme et 60 mM de TCEP, dans le cas des échantillons préparés à pH 3,2, et 2 mM de lysozyme et 20 mM de TCEP, dans le cas de ceux préparés à pH 4,4. Des mesures des bandes S-S effectuées sur une solution de lysozyme à 4 mM montrent qu'en l'absence de TCEP, la solution est stable pendant plus de dix heures.

La lysozyme est une macromolécule très polaire : à pH inférieur à 3,0 la charge nette est de +18e et à pH neutre de +8e. De ce fait, les interactions électrostatiques pourraient influencer d'une manière critique les constantes de vitesse intrinsèques de la réaction de réduction des ponts S-S. Étant donné la coexistence en solution aqueuse de quatre espèces actives de TCEP différemment chargées, il faudrait envisager des mesures des cinétiques à plusieurs valeurs de pH afin de déterminer les quatre constantes de vitesse possibles. En réalité, dans nos conditions expérimentales, l'intervalle de pH utilisable est limité par la gélification de la solution de lysozyme réduite observée au-dessus de pH=5,0 et par la constante de vitesse qui devient trop lente au-dessous de pH=3,0. C'est la raison pour laquelle nous nous sommes limités à mesurer les cinétiques à deux valeurs de pH seulement. Dans ce cas on ne peut espérer que de déterminer deux constantes de vitesse. Par conséquent, nous avons groupé les constantes de vitesse des espèces chargées en une seule constante effective, k_{fa} . La constante apparente de réaction s'écrit alors :

$$k_{ap} = w_1 k_{f1} + w_a k_{fa} \quad (4.33)$$

w_1 et w_a étant la fraction molaire de l'espèce neutre, et respectivement, la somme des fractions molaires des espèces chargées.

L'avancement de la réaction de réduction des ponts S-S de la lysozyme par la TCEP, suivi dans la bande Raman S-S, est présenté sur la Figure 4.19. Les bandes Raman S-S sont celles centrées sur 509 et 528 cm^{-1} . Nous observons qu'elles diminuent d'une manière similaire pendant la réaction. Les cinétiques expérimentales mesurées à pH=3,2 et pH=4,4 sont présentées sur les Figures 4.20 et 4.21. À chaque valeur de pH, la cinétique de réaction a été modélisée en utilisant deux constantes de vitesse apparentes, $k_{ap}^{(1)}$ (constante de vitesse rapide) et $k_{ap}^{(2)}$ (constante de vitesse lente). Les constantes de vitesse apparentes ainsi que les fractions molaires w_1 et w_a sont données dans le Tableau 4.5 pour les deux valeurs de pH. Nous pouvons écrire à chaque pH :

$$k_{ap}^{(1)}(pH) = w_1(pH)k_{f1}^{(1)} + w_a(pH)k_{fa}^{(1)} \quad (4.34)$$

$$k_{ap}^{(2)}(pH) = w_1(pH)k_{f1}^{(2)} + w_a(pH)k_{fa}^{(2)} \quad (4.35)$$

On obtient finalement quatre équations pour les quatre inconnues, $k_{f1}^{(1)}$, $k_{fa}^{(1)}$, $k_{f1}^{(2)}$ et $k_{fa}^{(2)}$. Les solutions de ces quatre équations sont consignées dans le Tableau 4.6.

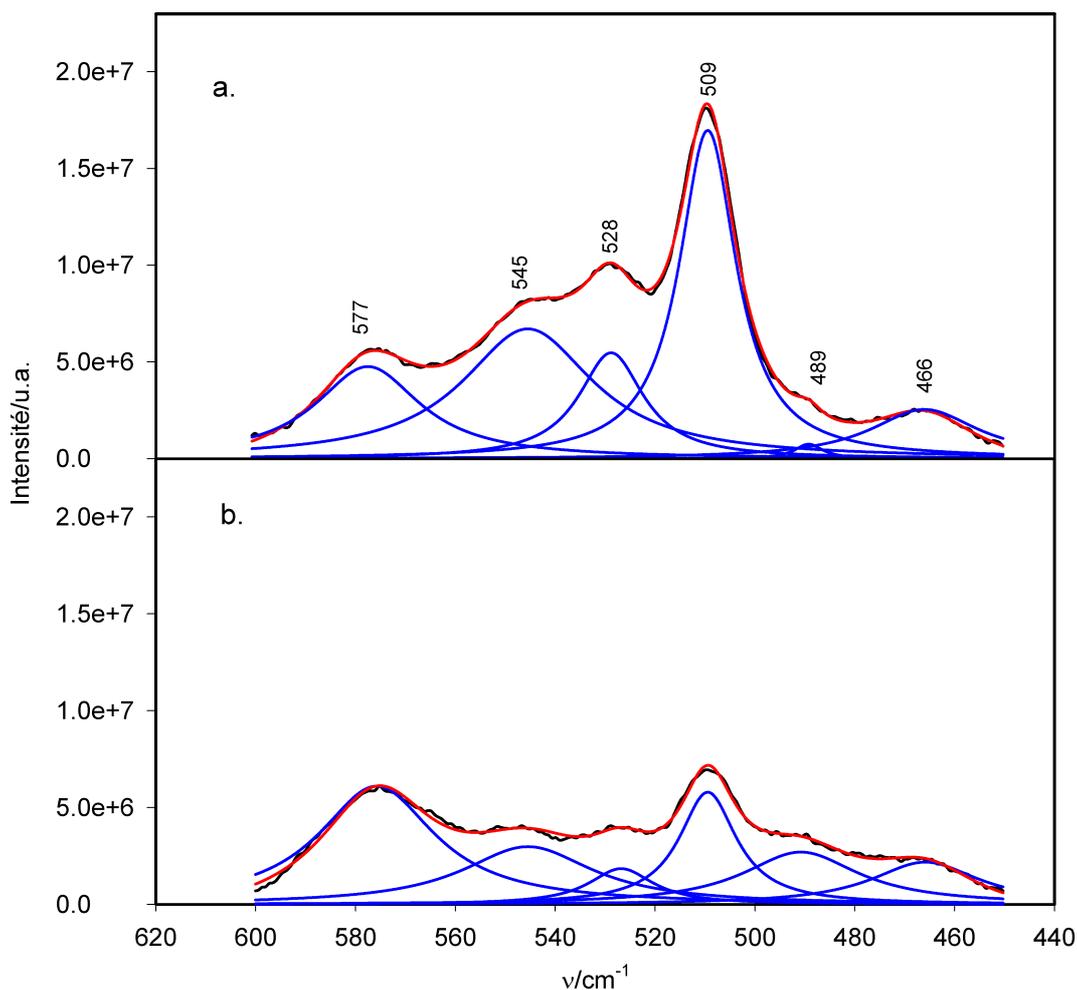


FIGURE 4.19 – Réduction des ponts S-S dans la lysozyme par la TCEP à pH=3,2 et à la température de 22°C : (a) spectre Raman d’une solution aqueuse de lysozyme native à 4 mM et (b) spectre Raman d’une solution aqueuse de lysozyme à 4 mM et TCEP à 60 mM mesuré 8 heures après le mélange.

En suivant le même raisonnement que dans le cas des albumines sériques réduites par la TCEP, on peut conclure que la seule explication possible pour l’hétérogénéité de la cinétique est la différence de réactivité entre les quatre ponts S-S de la lysozyme. Il est d’autant plus étonnant de noter que les deux bandes Raman S-S de la lysozyme ont la même cinétique : le rapport entre la surface de la bande centrée sur 509 cm^{-1} et celle centrée sur 528 cm^{-1} est constant par rapport à l’avancement de la réaction, avec une valeur de 2,8 (Figure 4.22). Ce résultat pourrait être expliqué soit en supposant que les quatre ponts S-S ont des cinétiques de réaction identiques, soit qu’il y a un recouvrement significatif des contribution de ces ponts aux bandes Raman. La deuxième hypothèse est partiellement soutenue par l’analyse théorique des bandes Raman S-S de la lysozyme pré-

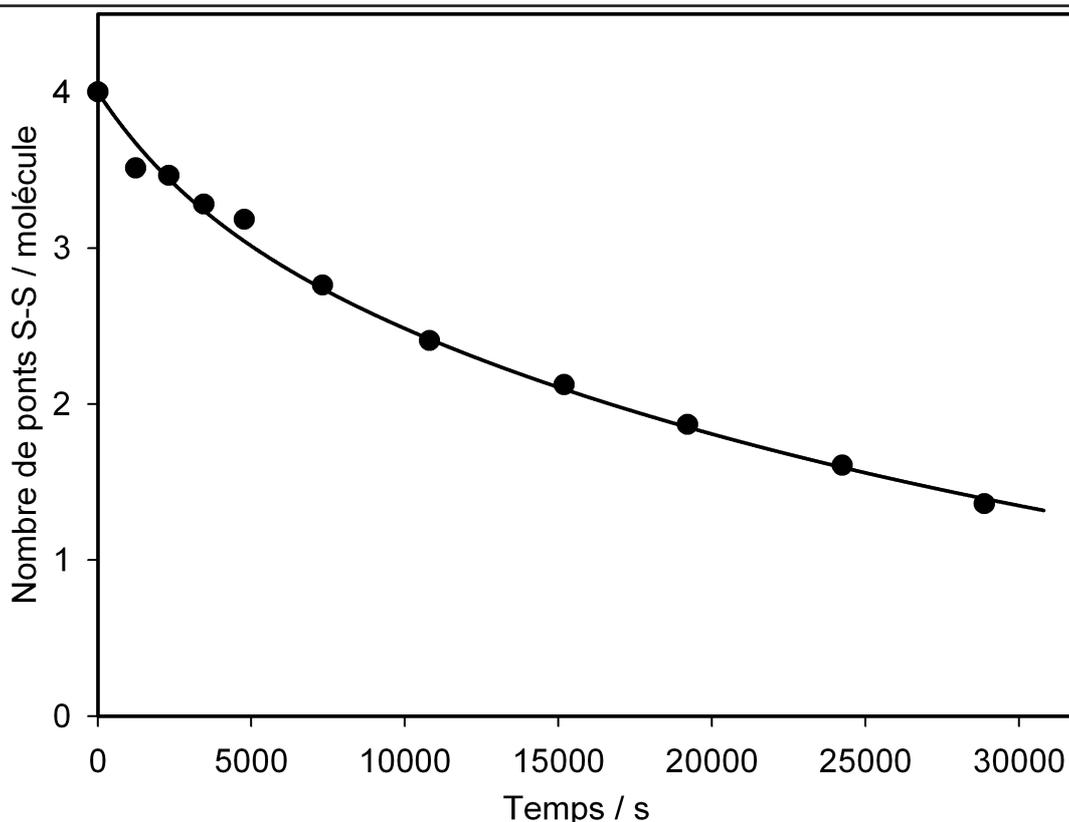


FIGURE 4.20 – Cinétique de la réduction des ponts S-S dans la lysozyme par la TCEP à pH=3,2 et à la température de 22°C. Le nombre de ponts S-S par molécule a été évalué à partir des surfaces des bandes Raman S-S. Les concentrations en lysozyme et TCEP sont de 4 mM et 60 mM, respectivement.

pH	w_i	w_a	$k_{ap}^{(1)}/M^{-1}s^{-1}$	$k_{ap}^{(2)}/M^{-1}s^{-1}$
3,25	0,006	0,0014	$5,0 \times 10^{-3}$, 25%	$0,58 \times 10^{-3}$, 75%
4,40	0,0005	0,0038	$6,5 \times 10^{-3}$, 25%	$0,90 \times 10^{-3}$, 75%

Tableau 4.5 – Constantes de vitesse de réaction rapide ($k_{ap}^{(1)}$) et lente ($k_{ap}^{(2)}$) déterminées pour deux valeurs du pH différentes. Les fractions des espèces actives, neutres et ionisées, de TCEP ont été évaluées pour chaque valeur du pH.

sentée dans le chapitre précédent. Par ailleurs, si on accepte le fait que les quatre ponts S-S ont des cinétiques de réaction identiques, alors il est nécessaire d'envisager une autre explication pour l'hétérogénéité des cinétiques, ce qui paraît difficile.

Les résultats du Tableau 4.6 indiquent des différences non négligeables entre les constantes de vitesse de l'espèce neutre de TCEP et celle des espèces chargées. Pour les deux composantes, rapide et lente, la constante de vitesse pour les espèces chargées négativement est

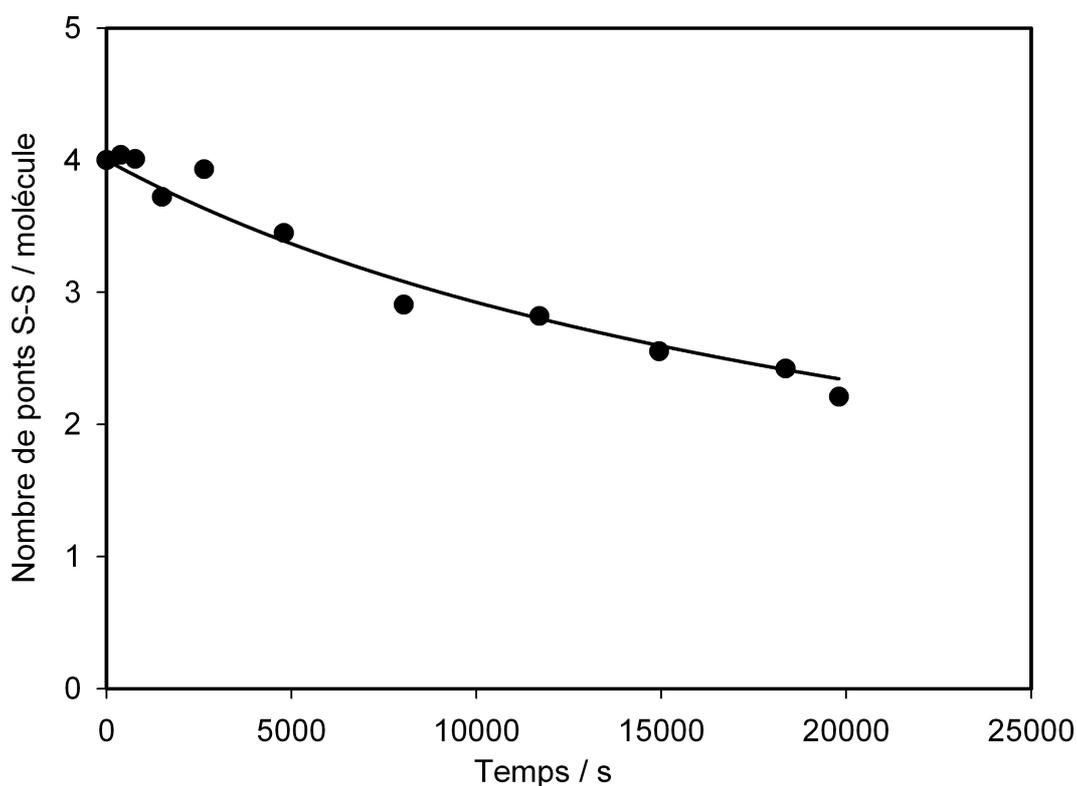


FIGURE 4.21 – Cinétique de la réduction des ponts S-S dans la lysozyme par la TCEP à pH=4,4 et à la température de 22°C. Le nombre de ponts S-S par molécule a été évalué à partir des surfaces des bandes Raman S-S. Les concentrations en lysozyme et TCEP sont de 2 mM et 20 mM, respectivement.

TCEP	$k_f^{(1)}/M^{-1}s^{-1}$	$k_f^{(2)}/M^{-1}s^{-1}$
Neutre	0,45	0,052
Ionisé	1,65	0,23

Tableau 4.6 – Constantes intrinsèques de vitesse de réaction rapide ($k_f^{(1)}$) et lente ($k_f^{(2)}$) déterminées pour les espèces actives, neutres et ionisées, de TCEP impliquées dans la réaction de réduction des ponts S-S de la lysozyme.

quatre fois supérieure à celle déterminée pour les espèces neutres. Cela représente une augmentation due aux interactions électrostatiques qui est plutôt modeste. On peut conclure que l'environnement proche des ponts S-S dans la lysozyme n'est pas très polaire.

Les valeurs des constantes de vitesse ainsi obtenues peuvent être utilisées pour déterminer l'enthalpie libre de dépliement local dans le cadre du modèle réactionnel en deux étapes (équation 4.30). Les valeurs obtenues pour la composante rapide et lente de la cinétique de l'espèce neutre sont $\Delta G_u^{(1)} = 8,5 \text{ kJ.mol}^{-1}$ et $\Delta G_u^{(2)} = 13,8 \text{ kJ.mol}^{-1}$. Ces

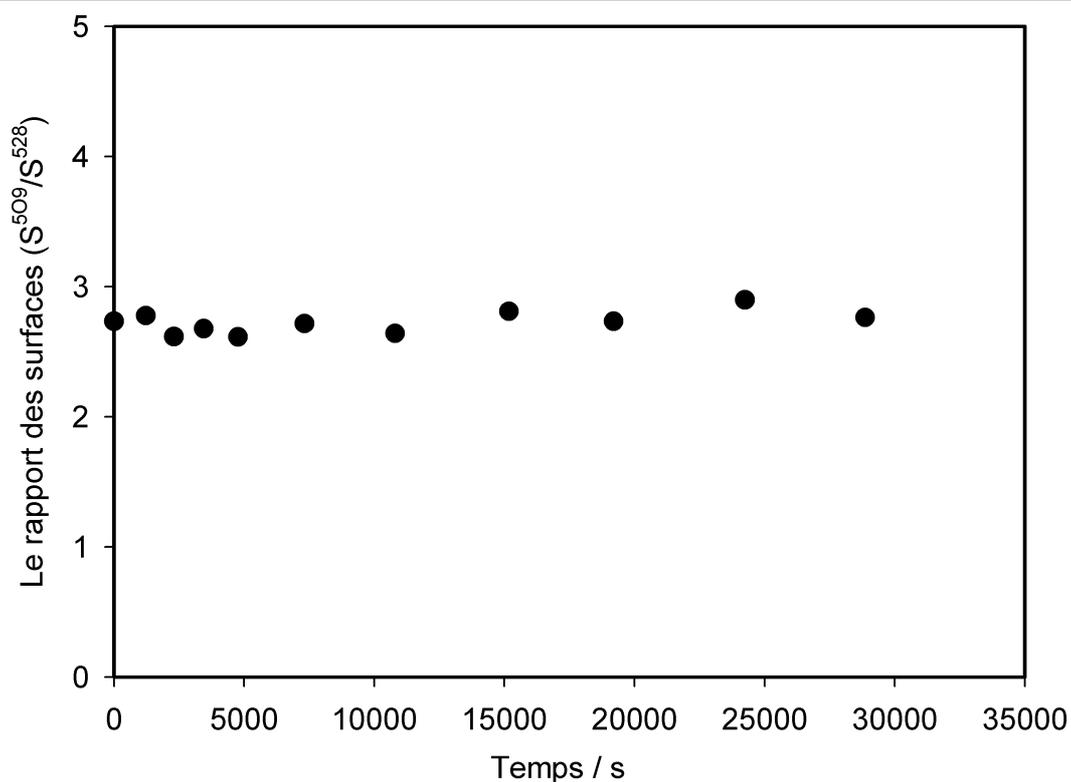


FIGURE 4.22 – Rapport entre la surface de la bande Raman S-S centrée sur 509 cm^{-1} (S^{509}) et celle centrée sur 528 cm^{-1} (S^{528}), en fonction du temps pendant la réaction de réduction des ponts S-S dans la lysozyme.

valeurs sont légèrement supérieures à celles déterminées dans le cas des albumines sériques.

4.3.3 Les effets conformationnels

La stabilité conformationnelle de la lysozyme lors de la réaction de réduction des ponts S-S a été vérifiée en analysant l'évolution de la bande amide I de la lysozyme. Il est connu que cette bande est surtout sensible aux modifications de la structure secondaire des protéines [27]. Dans le cas de la lysozyme nous avons trouvé que la réduction d'environ 30% de ponts S-S n'affecte pas la bande amide I d'une manière détectable (Figure 4.23). Au-dessus de ce taux de réduction on observe un faible changement (Figure 4.24). Ces résultats diffèrent beaucoup de ceux obtenus pour les albumines sériques (Figure 4.25) et indiquent une certaine stabilité structurale de la lysozyme. Cependant, la préservation de la structure secondaire de la lysozyme après la coupure des ponts S-S n'implique pas

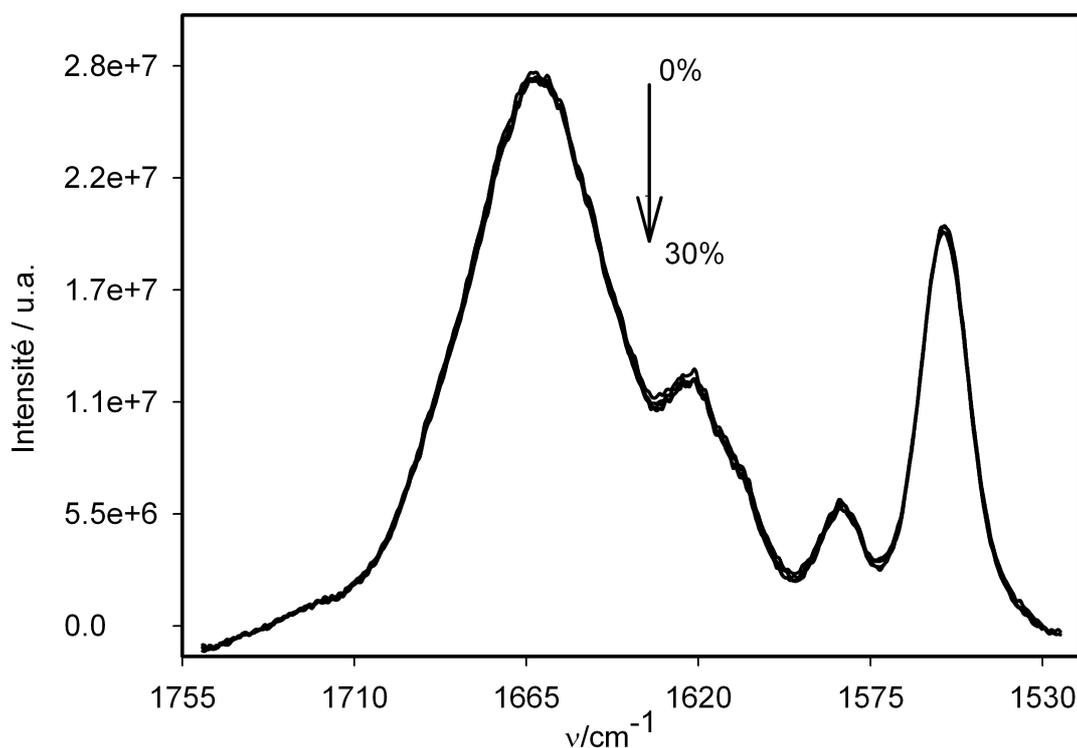


FIGURE 4.23 – Évolution de la bande Raman amide I de la lysozyme lors de la réduction d'environ 30% de ses ponts S-S par la TCEP. Les spectres ont été mesurés sur des solutions aqueuses de lysozyme à 2 mM et TCEP à 60 mM à pH 3,2.

l'absence de tout dépliement. Ce que l'on peut dire est que ce dépliement, s'il existe, ne peut pas être important. En effet, un dépliement important de la structure tertiaire de la protéine aurait certainement affecté la structure secondaire : il mènerait à une meilleure solvatation de la protéine et des changements dans le réseau de liaisons hydrogène.

La stabilité conformationnelle de la lysozyme par rapport à la réduction de ses ponts S-S a été testée aussi par des calculs de modélisation moléculaire. Nous avons ainsi calculé le potentiel des forces moyennes en fonction de la distance S-S entre les thiols d'un pont réduit. Le calcul a été effectué en utilisant la méthode d'échantillonnage dite "Umbrella Sampling", avec quatre fenêtres de simulation centrées sur 3,0, 4,0, 5,0 et 6,0 Å respectivement. La constante de force utilisée dans ces simulations a été de $10,0 \text{ kcal.mol}^{-1} \text{ \AA}^{-2}$. Les simulations de dynamique moléculaire ont été réalisées à température et pression constantes (300 K, 1 atm) sur la protéine solvatée dans une boîte $50 \times 60 \times 60 \text{ \AA}^3$, modèle TIP3P à conditions périodiques de frontières [80, 81]. Des contre-ions Cl^- ont été ajoutés afin de préserver la neutralité de la solution. La longueur de la trajectoire générée dans chaque fenêtre d'échantillonnage a été de 2 ns. Les données ainsi produites ont été analysées par la méthode WHAM d'Alan Grossfield. Les potentiels calculés pour

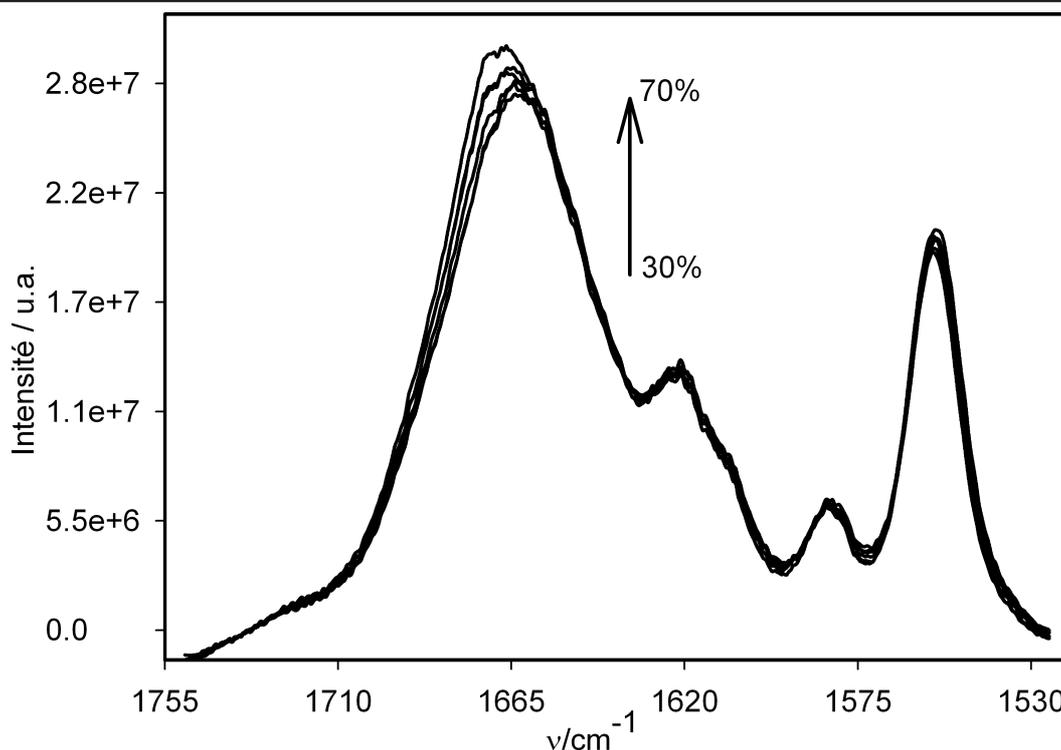


FIGURE 4.24 – Évolution de la bande Raman amide I de la lysozyme lors de la réduction de 30% à 70% de ses ponts S-S par la TCEP. Les spectres ont été mesurés sur des solutions aqueuses de lysozyme à 2 mM et TCEP à 60 mM à pH 3,2.

les quatre ponts de la lysozyme sont présentés sur la Figure 4.26. On constate que chacune de quatre courbes a un minimum à environ 3,5 Å qui correspond à peu près à la distance de Van der Waals entre deux atomes de soufre.

Ce résultat peut être interprété dans le sens que la structure protéique favorise le contact des thiols de chaque pont. Cela facilite la formation (ou la reconstitution) de ces ponts : pour une protéine en conformation native, les thiols d'un pont restent en contact d'une manière naturelle.

Conclusions

L'analyse des cinétiques des réactions de réduction des ponts S-S dans la lysozyme et les albumines sériques met en évidence l'hétérogénéité de ces cinétiques. Cette hétérogénéité a été attribuée aux différences de réactivité entre les ponts S-S d'une protéine.

Les différences entre les constantes de vitesse de réduction de ponts S-S dans les protéines et celle de la réduction d'un pont S-S modèle tel le GSSG ont été discutées dans le cadre d'un modèle de réaction en deux étapes. Dans ce modèle, la première étape de réaction est un dépliement local de la protéine qui rend l'accès au pont possible. Les

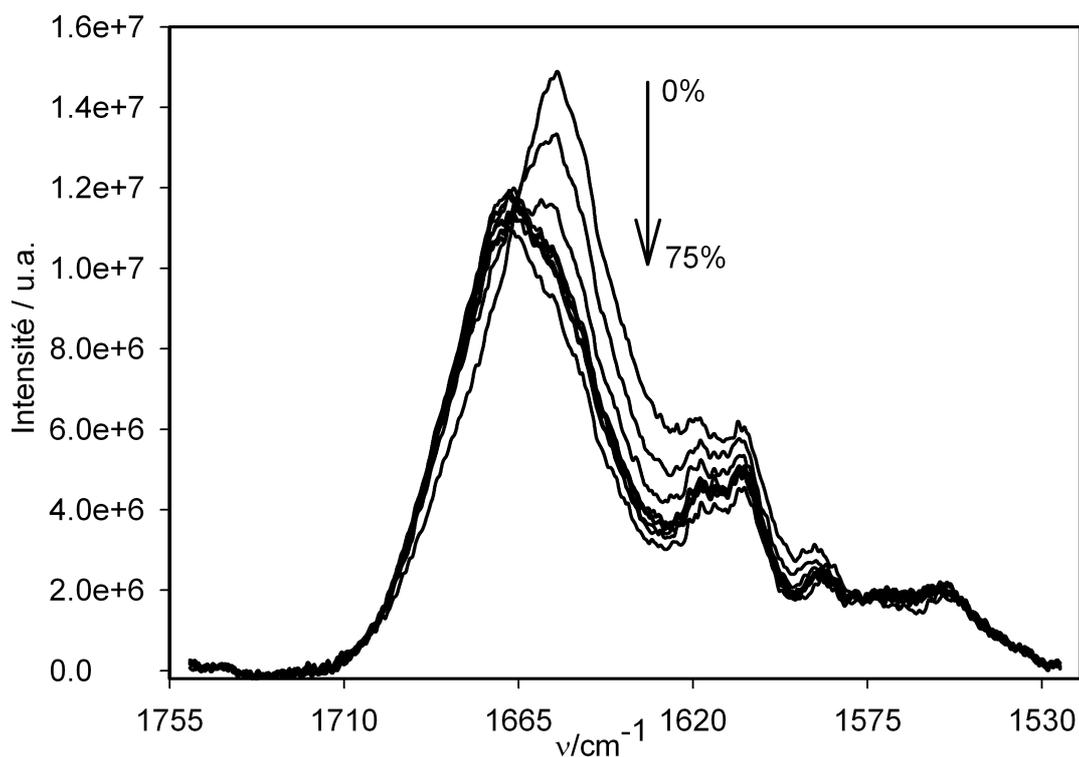


FIGURE 4.25 – Évolution de la bande Raman amide I de la BSA lors de la réduction de 0% à 75% de ses ponts S-S par le DTT. Les spectres ont été mesurés sur des solutions aqueuses de BSA à 1 mM et DTT à 17 mM à pH 7,8.

données expérimentales nous ont permis d'évaluer l'enthalpie libre associée à ce dépliement local. Cependant, les différences entre les constantes de vitesse de réduction des ponts S-S dans les protéines et celle de réduction du GSSG sont relativement faibles. On peut alors se demander si la différence d'accessibilité des ponts est la seule source de ces différences de réactivité.

Les calculs moléculaires qui sont présentés dans le chapitre suivant essaient de répondre à cette question dans un cas précis : celui des ponts S-S de la lysozyme.

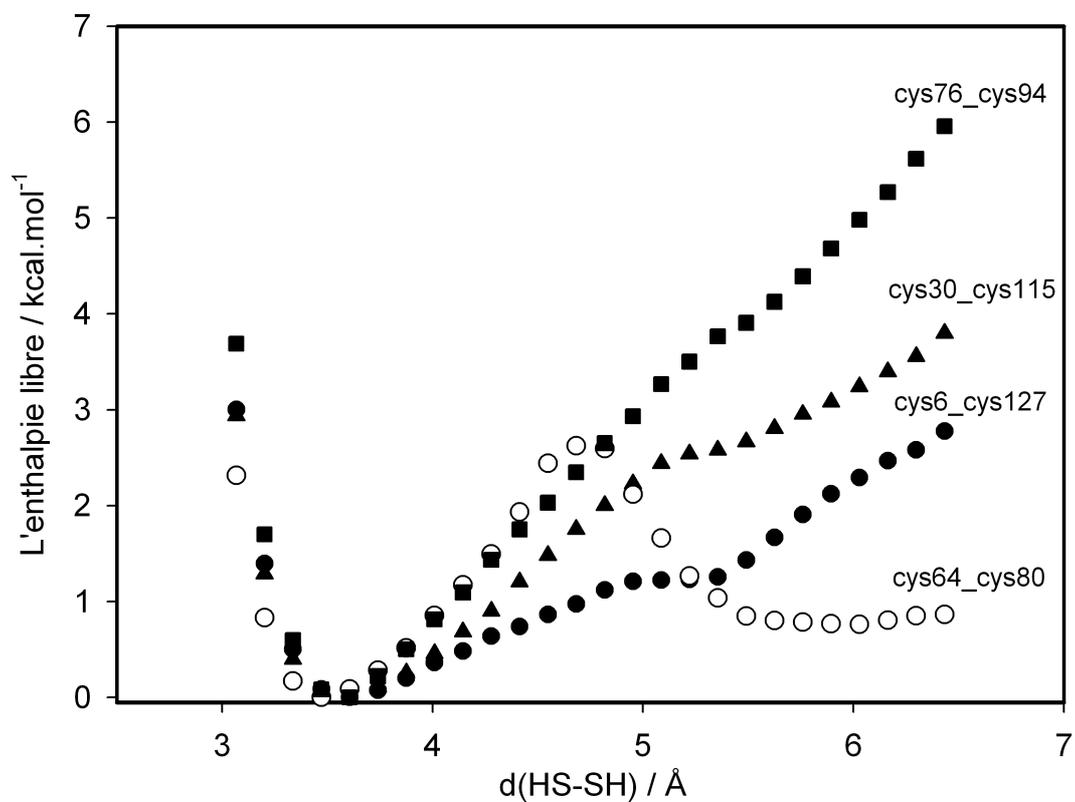


FIGURE 4.26 – Potentiel de force moyenne calculé en fonction de la distance S-S entre les thiols des ponts réduits de la lysozyme.

Chapitre 5

Modélisation moléculaire de la réaction de réduction des ponts S-S dans la lysozyme par la TCEP

Sommaire

5.1	Introduction	118
5.2	Principe de la méthode	120
5.3	Méthodes de calcul	122
5.4	Résultats	128
5.5	Discussion	137

5.1 Introduction

Dans l'étude théorique des réactions chimiques en environnement protéique il est d'usage de séparer le système en deux parties : les atomes actifs (directement impliqués dans la réaction) et l'environnement (le reste de la protéine et/ou le solvant). Pour une configuration donnée du système, les méthodes hybrides QM/MM utilisant une méthode *ab initio* pour le niveau de théorie QM permettent de calculer avec précision l'énergie du groupe d'atomes actifs (le système modèle) tout en prenant en compte son interaction avec l'environnement. Cependant, pour les systèmes moléculaires de grande taille, l'énergie calculée pour une seule configuration pourrait ne pas être significative en tant que paramètre décrivant l'état du système. Des calculs d'énergie libre sont alors nécessaires, ce qui implique l'évaluation de l'énergie du système pour un très grand nombre de configurations. En raison de son coût en temps de calcul, cette approche se place hors de la portée d'une méthode QM/MM dont le niveau QM de la méthode serait basé sur des calculs *ab initio*. Néanmoins, l'échantillonnage de l'espace configurationnel peut être effectué en utilisant des méthodes semi-empiriques à la place des méthodes *ab initio* [52, 121–124]. Les résultats peuvent encore être améliorés en effectuant des corrections par rapport aux méthodes *ab initio* [58, 125]. Cependant, l'applicabilité des méthodes semi-empiriques reste très limitée, en raison de leurs insuffisances intrinsèques.

L'alternative proposée par Jorgensen et ses collaborateurs [126, 127] consiste à limiter l'échantillonnage de l'espace de configurations du système aux coordonnées des atomes MM seules. Une trajectoire de réaction est d'abord déterminée pour le système modèle (QM) par des calculs *ab initio* en phase gazeuse. Des configurations du système modèle sont extraites de cette trajectoire et mises en interaction avec l'environnement réel. L'énergie libre d'interaction avec l'environnement est calculée en échantillonnant l'espace de configurations des atomes MM pour chaque configuration des atomes QM maintenue fixe. L'énergie libre totale du système est ensuite obtenue en additionnant l'énergie QM à l'énergie libre d'interaction avec l'environnement. C'est pourquoi la méthode est désignée comme la méthode QM-FE (quantum mechanics free energy method). Il est évident que l'on ne peut traiter de cette manière que les réactions dont les trajectoires en phase gazeuse et, respectivement, dans l'environnement naturel, sont très proches. De plus, avec cette méthode, les éventuelles distorsions du système modèle induites par l'environnement ne peuvent pas être prises en compte. La solution optimum pour dépasser cette insuffisance est d'obtenir la trajectoire de réaction directement par une méthode QM/MM. Il est vrai que dans ce type de calcul, l'énergie totale du système calculée au long de la trajectoire n'est pas fiable à cause de l'instabilité configurationnelle du système. Cependant, l'énergie

du système QM et l'évolution de la configuration de ce système au long de la trajectoire de réaction sont beaucoup plus stables et peuvent servir de base pour définir cette trajectoire.

Une fois la trajectoire de réaction déterminée, le calcul de l'énergie libre d'interaction du système modèle avec l'environnement s'effectue de la même manière que dans la méthode de Jorgensen.

Le calcul de la trajectoire de réaction par une méthode QM/MM doit se faire en utilisant l'assemblage électrostatique ("electronic embedding") entre le système modèle et son environnement. Dans ces conditions, l'optimisation de la géométrie du système par la minimisation de son énergie devient délicate. Dans certains cas où le couplage entre le système modèle et l'environnement est fort, les procédures standard d'optimisation peuvent ne pas aboutir. Cet échec est dû au caractère itératif de ces procédures : on optimise d'abord l'environnement pour une configuration du système QM fixée, ensuite on fixe la configuration de l'environnement et on optimise celle du système QM. Ces opérations sont répétées jusqu'à ce que la convergence serait atteinte.

La réaction de réduction des ponts S-S des protéines par la TCEP est caractérisée par un fort couplage entre le système modèle et l'environnement. En effet, le produit de réaction est très polaire et l'environnement très polarisable. De ce fait, une première conséquence est que la réaction n'est pas possible dans le vide, donc la méthode de Jorgensen n'est pas utilisable dans sa forme originale. Dans nos essais, l'optimisation du produit de réaction par une méthode QM/MM basée sur l'assemblage électrostatique n'a pas convergé. Il n'a donc pas été possible de déterminer la trajectoire de réaction directement par des calculs QM/MM. La solution que nous avons trouvée part de l'observation qu'au moment de la réaction un pont S-S doit être bien exposé au solvant, son environnement étant assez proche de celui d'un pont S-S modèle réagissant avec la TCEP en solution aqueuse. Cela suggère fortement que la trajectoire de réaction en milieu protéique et celle en solution aqueuse sont très proches. Nous avons donc déterminé la trajectoire de réaction en effectuant des calculs *ab initio* pour un système modèle en solution aqueuse. Dans ces calculs, le solvant a été traité en utilisant le modèle du solvant continu, PCM.

Un autre point original de notre méthode est la prise en compte des distorsions du système modèle induites par l'environnement. Ces distorsions ont été simulées pour un assemblage mécanique entre le système modèle et son environnement : pour chaque configuration du système modèle extraite de la trajectoire de réaction une optimisation QM/MM a été effectuée en environnement protéique. Dans cette optimisation, la coordonnée de réaction a été fixée ainsi que les charges atomiques partielles du système modèle. L'énergie quantique calculée pour le système ainsi optimisé a ensuite été utilisée pour le calcul de

l'énergie libre totale du système.

Dans ce travail, la variation d'énergie libre au long de la trajectoire de réaction a été calculée pour la réduction des quatre ponts S-S de la lysozyme, ainsi que pour la réduction du GSSG. Le but de ces calculs a surtout été de comparer les réactivités de ces systèmes et d'analyser l'importance des effets d'environnement dans cette réaction.

5.2 Principe de la méthode

L'énergie d'un système hybride QM/MM s'exprime de la manière suivante (dans le cadre du schéma QM/MM soustractif) :

$$E(X_{R-M}, X_M) = E_{QM}^M(X_M) + E_{MM}^R(X_{R-M}, X_M) - E_{MM}^M(X_M) \quad (5.1)$$

où X_{R-M} représente l'ensemble des coordonnées des atomes MM, X_M l'ensemble des coordonnées des atomes QM (le système modèle), $E_{QM}^M(X_M)$ l'énergie quantique du système modèle, $E_{MM}^R(X_{R-M}, X_M)$ l'énergie MM du système réel et $E_{MM}^M(X_M)$ l'énergie MM du système modèle.

Soit ξ la coordonnée de réaction dont on suppose qu'elle n'est dépendante que des coordonnées du système modèle. On peut définir une fonction de partition partielle liée à la coordonnée de réaction par l'expression :

$$Z(\xi_i) = \int \delta(\xi - \xi_i) \int e^{-\beta E(X_{R-M}, X'_M, \xi)} d\xi dX_{R-M} dX'_M = \int e^{-\beta E(X_{R-M}, X'_M, \xi_i)} dX_{R-M} dX'_M \quad (5.2)$$

Dans l'équation 5.2, $\beta = 1/k_B T$ et X'_M représente l'ensemble des coordonnées du système modèle, moins la coordonnée ξ . À partir de $Z(\xi_i)$ on peut calculer une énergie libre partielle :

$$F(\xi_i) = -\frac{1}{\beta} \ln Z(\xi_i) \quad (5.3)$$

Si l'on néglige les fluctuations de $E_{QM}^M(X_M)$ autour de sa valeur la plus probable sur la trajectoire de réaction qui est $E_{QM}^M(X_{M,min}(\xi_i))$, on obtient une expression simplifiée de $F(\xi_i)$ [128] :

$$F(\xi_i) \cong E_{QM}^M(X_{M,min}(\xi_i)) + F_{env}(\xi_i) \quad (5.4)$$

où :

$$F_{env}(\xi_i) = -\frac{1}{\beta} \ln \int e^{-\beta E_{env}(X_{R-M}, X'_M, \xi_i)} dX_{R-M} dX'_M \quad (5.5)$$

et :

$$E_{env}(X_{R-M}, X'_M, \xi_i) = E_{MM}^R(X_{R-M}, X'_M, \xi_i) - E_{MM}^M(X'_M, \xi_i) \quad (5.6)$$

Les coordonnées $X_{M,min}(\xi_i)$ sont celles qui minimisent l'énergie QM/MM pour $\xi = \xi_i$. Par la suite, nous utiliserons la notation $E_{QM}^M(\xi_i)$ pour désigner $E_{QM}^M(X_{M,min}(\xi_i))$. Ici F_{env} représente l'énergie libre d'interaction entre le système modèle et l'environnement. La variation de F_{env} lors du passage de ξ_i à ξ_{i+1} s'écrit :

$$\Delta F_{env}^{i \rightarrow i+1} = F_{env}(\xi_{i+1}) - F_{env}(\xi_i) \quad (5.7)$$

Dans la méthode QM-FE standard, $\Delta F_{env}^{i \rightarrow i+1}$ est calculée en appliquant la méthode des perturbations. Cette méthode exige que, dans l'équation 5.5, les coordonnées du système modèle soient fixées aux valeurs $X_{M,min}(\xi_i)$. Cela revient à négliger la variation de E_{env} lors des variations des coordonnées du système modèle. De plus, la méthode présente une certaine difficulté technique : afin d'évaluer l'énergie de perturbation, les configurations sur la trajectoire de simulation doivent être modifiées en déplaçant les coordonnées du système modèle de $X_{M,min}(\xi_i)$ à $X_{M,min}(\xi_{i+1})$. Afin que ce déplacement n'entraîne pas de distorsions artefactuelles, des schémas ont été proposés où, à part X_M , les coordonnées de certains atomes MM sur la frontière QM-MM sont elles aussi modifiées [124].

Nous proposons ici une approche différente qui évite ces difficultés tout en prenant en compte les variations des coordonnées du système modèle dans les calculs du type MM. Le premier pas est de remplacer $E_{MM}^M(X'_M, \xi_i)$ dans l'équation 5.6 par sa valeur moyenne calculée sur l'ensemble de configurations généré pour $\xi = \xi_i$:

$$E_{env}(X_{R-M}, X'_M, \xi_i) \cong E_{MM}^R(X_{R-M}, X'_M, \xi_i) - \left\langle E_{MM}^M(X'_M, \xi_i) \right\rangle_{\xi_i} \quad (5.8)$$

Cette approximation est cohérente avec l'omission des fluctuations de l'énergie quantique du système modèle $E_{QM}^M(X_{M,min}(\xi_i))$ (équation 5.4). De plus, les effets des fluctuations des coordonnées X'_M sur l'énergie $E_{MM}^R(X_{R-M}, X'_M, \xi_i)$ sont maintenant prises en compte. L'énergie libre partielle du système devient alors :

$$F(\xi_i) = E_{QM}^M(\xi_i) - \left\langle E_{MM}^M(X'_M, \xi_i) \right\rangle_{\xi_i} + F_{MM}(\xi_i) \quad (5.9)$$

où :

$$F_{MM}(\xi_i) = \int e^{-\beta E_{MM}^R(X_{R-M}, X'_M, \xi_i)} d\xi dX_{R-M} dX'_M \quad (5.10)$$

Il est intéressant de remarquer que $F_{MM}(\xi_i)$ représente l'énergie libre partielle du système réel calculée au niveau de théorie MM. Sa variation $\Delta F_{MM}^{i \rightarrow i+1}$ peut être déterminée par d'autres méthodes que la méthode des perturbations. Dans ce travail nous l'avons calculée par la méthode TDI. Dans ce but, $F_{MM}(\xi_i)$ est remplacée par une énergie libre partielle modifiée :

$$\bar{F}_{MM}(\xi_i) = -\frac{1}{\beta} \ln \int e^{-\beta \frac{k}{2} (\xi - \xi_i)^2} \int e^{-\beta E_{MM}^R(X_{R-M}, X'_M, \xi)} d\xi dX_{R-M} dX'_M \quad (5.11)$$

Cela revient à remplacer le système à coordonnée de réaction fixée par un système dont la variation de la coordonnée de réaction est restreinte par une restriction harmonique.

Il est évident que pour des valeurs de la constante de force k suffisamment élevées :

$$F_{MM}(\xi_{i+1}) - F_{MM}(\xi_i) \cong \bar{F}_{MM}(\xi_{i+1}) - \bar{F}_{MM}(\xi_i) \quad (5.12)$$

L'énergie effective du système modifié est :

$$\bar{E}_{MM}^R(X_{R-M}, X'_M, \xi, \xi_i) = E_{MM}^R(X_{R-M}, X'_M, \xi) + \frac{k}{2}(\xi - \xi_i)^2 \quad (5.13)$$

Selon la procédure standard de la méthode TDI on va définir une énergie hybride :

$$\bar{E}_{MM}^R(\lambda) = (1 - \lambda)\bar{E}_{MM}^R(\xi_i) + \lambda\bar{E}_{MM}^R(\xi_{i+1}) \quad (5.14)$$

et l'énergie libre hybride :

$$F_{MM}(\lambda) = -\frac{1}{\beta} \ln \int e^{-\beta\bar{E}_{MM}^R(\lambda)} d\xi dX_{R-M} dX'_M \quad (5.15)$$

Il faut préciser que $\bar{E}_{MM}^R(\xi_i)$ est une notation abrégée pour $\bar{E}_{MM}^R(X_{R-M}, X'_M, \xi, \xi_i)$ calculée avec les charges atomiques partielles qui correspondent à ξ_i .

La variation d'énergie libre s'exprime alors :

$$\Delta F_{env}^{i \rightarrow i+1} = \int_0^1 \frac{\partial \bar{F}_{MM}}{\partial \lambda} d\lambda \quad (5.16)$$

D'autre part :

$$\frac{\partial \bar{F}_{MM}}{\partial \lambda} = \langle \bar{E}_{MM}^R(\xi_{i+1}) - \bar{E}_{MM}^R(\xi_i) \rangle_\lambda \quad (5.17)$$

La notation $\langle \rangle_\lambda$ signifie que l'ensemble de configurations sur lequel s'effectue le calcul de la moyenne correspond à un système dont l'énergie est $\bar{E}_{MM}^R(\lambda)$.

La moyenne qui apparaît dans le membre de droite de l'équation 5.17 peut être évaluée en utilisant une trajectoire de simulation de dynamique moléculaire générée avec $\bar{E}_{MM}^R(\lambda)$ comme énergie du système. Par conséquent, l'application de la méthode est assez simple.

5.3 Méthodes de calcul

Trajectoire préliminaire de réaction

La trajectoire de réaction a été déterminée en effectuant des calculs QM-PCM sur le système modèle en solution aqueuse. Les différentes configurations sur la trajectoire de

réaction ont été obtenues en fixant la distance S-P à des différentes valeurs et en optimisant la géométrie du système par rapport aux autres coordonnées. La coordonnée de réaction, ξ , est une combinaison des distances S-P et S-S. C'est pourquoi chaque configuration extraite de la trajectoire de réaction et utilisée ensuite dans les calculs QM/MM et MM a été caractérisée par ces deux distances. Pour chacune de ces configurations, nous avons aussi calculé les charges atomiques partielles du système modèle en utilisant la méthode NPA (natural population analysis) [129].

Charges atomiques partielles

Dans les calculs QM/MM et MM nous avons utilisé les charges NPA pour les cystéines participant aux ponts S-S et pour la TCEP, et les charges ESP [77], implémentées dans AMBER9, pour le reste du système. Les charges NPA du système modèle varient en fonction de la coordonnée de réaction. Elles ont été maintenues fixes dans les simulations effectuées pour une valeur donnée de cette coordonnée de réaction.

Dans la méthode NPA, les charges atomiques partielles sont obtenues en analysant les populations électroniques qui occupent les différentes orbitales atomiques (OA). Dans ce cas on utilise les OA naturelles qui, à la différence des orbitales canoniques, sont orthogonales les unes par rapport aux autres. Dans la méthode ESP, les charges atomiques partielles sont obtenues en modélisant par des charges ponctuelles le potentiel électrostatique quantique sur une surface choisie.

La raison pour laquelle nous avons préféré les charges NPA au charges ESP réside dans le manque de fiabilité de ces dernières dans le cas du système étudié. Ce manque de fiabilité est indiqué par :

- les différences significatives qui apparaissent entre les charges ESP calculées pour des atomes du système modèle chimiquement équivalents
- la variation discontinue de ces charges au long de la trajectoire de réaction
- l'instabilité des charges ESP par rapport aux changements de configuration du système modèle

Ces insuffisances viennent du fait que les charges ESP ont une signification physique limitée, étant plutôt des paramètres de modélisation. Ainsi, elles reproduisent d'une manière correcte les potentiels électrostatiques globaux des fragments considérés. Cependant, les potentiels partiels attribués aux groupes d'atomes dans un fragment ne sont pas fiables.

Atomes \ d(S-P)Å	2,12	2,17	2,3	2,4	2,6	2,8	3,0	3,4	3,8	4,5
C2	-0,396	-0,380	-0,406	-0,481	-0,373	-0,272	-0,319	-0,282	-0,240	-0,278
C1	0,089	0,059	0,091	0,109	0,055	0,020	0,082	0,085	0,154	0,179
P	0,269	0,278	0,253	0,111	0,204	0,089	0,026	-0,028	-0,334	-0,476
C1'	-0,185	-0,179	-0,151	-0,122	0,008	-0,082	-0,096	-0,068	0,278	0,182
C2'	-0,102	-0,089	-0,113	-0,217	-0,200	-0,150	-0,132	-0,156	-0,301	-0,218
C1''	-0,087	-0,088	-0,051	-0,240	-0,155	-0,108	-0,100	-0,079	-0,086	0,374
C2''	-0,095	-0,091	-0,101	-0,069	-0,142	-0,195	-0,191	-0,270	-0,237	-0,422
C4'	-0,082	-0,089	-0,126	-0,572	-0,327	-0,286	-0,257	-0,238	-0,315	-0,136
C3'	0,027	0,010	0,079	0,234	0,037	-0,003	-0,025	0,080	0,024	0,351
S1	-0,133	-0,147	-0,194	0,016	-0,027	-0,012	-0,020	-0,091	-0,091	-0,190
C3	-0,162	-0,155	-0,189	-0,253	-0,250	-0,237	-0,222	-0,188	-0,267	-0,105
C4	0,253	0,228	0,228	0,109	0,097	0,059	0,081	0,099	0,134	-0,029
S1'	-1,136	-1,119	-0,958	-0,768	-0,474	-0,376	-0,337	-0,289	-0,269	-0,225

Tableau 5.1 – Charges atomiques partielles du système modèle calculées par la méthode ESP (voir la Figure 5.4).

Atomes \ d(S-P)Å	2,12	2,17	2,3	2,4	2,6	2,8	3,0	3,4	3,8	4,5
C2	-0,585	-0,585	-0,584	-0,583	-0,579	-0,578	-0,578	-0,578	-0,578	-0,579
C1	-0,720	-0,720	-0,716	-0,704	-0,693	-0,693	-0,695	-0,695	-0,695	-0,695
P	1,335	1,334	1,259	1,133	0,914	0,846	0,812	0,780	0,755	0,740
C1'	-0,719	-0,719	-0,716	-0,709	-0,695	-0,694	-0,693	-0,691	-0,691	-0,690
C2'	-0,588	-0,587	-0,585	-0,583	-0,580	-0,579	-0,579	-0,579	-0,579	-0,580
C1''	-0,709	-0,711	-0,712	-0,705	-0,697	-0,697	-0,696	-0,694	-0,694	-0,693
C2''	-0,588	-0,588	-0,586	-0,585	-0,583	-0,582	-0,581	-0,581	-0,581	-0,581
C4'	-0,595	-0,595	-0,594	-0,594	-0,594	-0,593	-0,593	-0,594	-0,594	-0,595
C3'	-0,495	-0,497	-0,508	-0,505	-0,502	-0,506	-0,510	-0,513	-0,512	-0,513
S1	0,017	0,014	-0,004	-0,010	0,018	0,041	0,053	0,060	0,055	0,047
C3	-0,585	-0,585	-0,584	-0,588	-0,591	-0,593	-0,594	-0,595	-0,595	-0,596
C4	-0,498	-0,499	-0,503	-0,506	-0,509	-0,509	-0,509	-0,507	-0,507	-0,507
S1'	-0,778	-0,769	-0,585	-0,378	-0,102	-0,029	0,001	0,025	0,039	0,046

Tableau 5.2 – Charges atomiques partielles du système modèle calculées par la méthode NPA (voir la Figure 5.4).

Cela pose de gros problèmes dans le calcul des interactions qui impliquent des atomes de frontière liés par des liaisons covalentes. D'autre part, les charges NPA reproduisent un peu moins bien les potentiels électrostatiques globaux des fragments, mais elles sont directement liées à la densité électronique réelle du système. De ce fait elles sont plus stables et suivent d'une manière fiable les changements chimiques du système. Une comparaison entre les deux types de charges pour les principaux atomes du système modèle est donnée dans les Tableaux 5.1 et 5.2.

Il est nécessaire de mentionner que l'utilisation des charges NPA pour le système modèle n'est pas incompatible avec l'attribution des charges ESP aux autres fragments moléculaires du système réel. En effet, les moments dipolaires du système modèle calculés avec les charges NPA sont suffisamment proches des moments dipolaires ESP (voir la Figure 5.1).

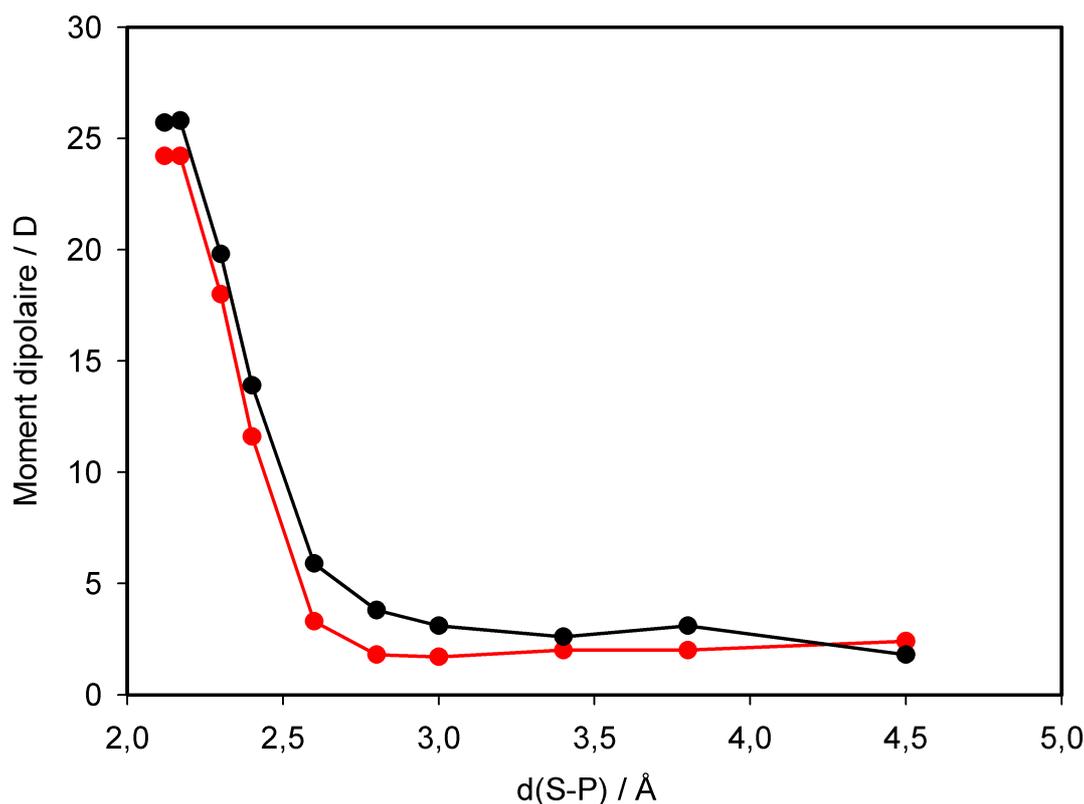


FIGURE 5.1 – Comparaison entre les moments dipolaires ESP (noir) et NPA (bleu) du système modèle calculés au long de la trajectoire de réaction.

Calcul de l'énergie quantique du système modèle

Chaque configuration du système modèle extraite de la trajectoire de réaction préliminaire et qui correspond à une valeur fixe de la coordonnée de réaction ξ_i a été transférée dans le système QM/MM. Des optimisations de géométrie ont été ensuite effectuées pour chaque valeur ξ_i maintenue fixe, et avec les charges atomiques partielles du système modèle correspondantes. Par ce calcul nous avons obtenu l'énergie $E_{QM}(\xi_i)$ qui inclut les distorsions imposées au système modèle par l'environnement.

Calcul de la contribution de l'environnement à l'énergie libre

La contribution de l'environnement à l'énergie libre totale a été déterminée en échantillonnant l'espace de configurations du système réel au niveau de théorie MM. L'échantillonnage a été réalisé par des simulations de dynamique moléculaire.

La variation d'énergie libre lorsque la coordonnée de réaction varie de ξ_i à ξ_{i+1} a été déterminée par la méthode de l'intégration thermodynamique (voir le Chapitre 2). Pour chaque valeur du paramètre de mixage λ , une trajectoire de 0,5 ns a été générée en utilisant l'énergie hybride suivante :

$$E_i(\lambda) = (1 - \lambda)E(\xi_i) + \lambda E(\xi_{i+1}) \quad (5.18)$$

L'énergie $E(\xi_i)$ est calculée en utilisant les charges atomiques partielles du système modèle qui correspondent à la configuration i sur la trajectoire de réaction. Les longueurs d'équilibre imposées aux liaisons S-S et S-P correspondent également à la configuration i . Ces valeurs ne sont pas fixes, mais restreintes par des restrictions harmoniques de constante de force $k = 240 \text{ kcal.mol}^{-1} \text{ \AA}^{-2}$.

Étant donné que les configurations i et $i + 1$ sont très proches, la variation d'énergie libre a été calculée en utilisant une seule valeur de λ . Conformément aux équations 2.42 et 2.43, on obtient dans ce cas :

$$\Delta F_{i,MM} = \bar{F}_{MM}(\xi_{i+1}) - \bar{F}_{MM}(\xi_i) = \langle \bar{E}_{MM}^R(\xi_{i+1}) - \bar{E}_{MM}^R(\xi_i) \rangle_{0,5} \quad (5.19)$$

La variation d'énergie libre totale a été calculée en ajoutant la contribution quantique et en soustrayant la contribution MM moyenne du système modèle :

$$\Delta F_i = \Delta F_{i,MM} + E_{QM}^M(\xi_{i+1}) - E_{QM}^M(\xi_i) + \langle E_{MM}^M \rangle_{\xi_i} - \langle E_{MM}^M \rangle_{\xi_{i+1}} \quad (5.20)$$

Il est évident que dans ce cas les variations d'énergie libre et d'enthalpie libre (ΔG_i) sont égales puisque :

$$G_i = F_i + RT \quad (5.21)$$

Détails sur les méthodes de calcul

Les calculs QM et QM/MM ont été effectués avec le logiciel Gaussian 03. Nous avons utilisé la méthode B3LYP/6-311(d,p) pour le niveau de théorie QM. Dans les calculs QM/MM nous avons utilisé la version 1996 du champ de force AMBER, implémentée dans Gaussian 03.

Les calculs de mécanique moléculaire ont été effectués en utilisant le programme AMBER9 et le champ de force AMBER 1999, particulièrement adapté à la description des

protéines. Les paramètres de force pour la TCEP sont ceux fournis par GAFF (Generalised Amber Force Field). La géométrie de départ de la lysozyme a été obtenue à partir de sa structure de diffraction des rayons X [104]. Les calculs MM ont été effectués sur la protéine solvatée dans une boîte d'eau parallélépipédique $50 \times 60 \times 60 \text{ \AA}^3$ contenant environ 4690 molécules d'eau TIP3P (Figure 5.2). Des conditions périodiques aux frontières ont été imposées et les interactions électrostatiques ont été décrites à l'aide de la méthode d'Ewald. Afin de neutraliser la charge de la protéine, des ions Cl^- ont été ajoutés.

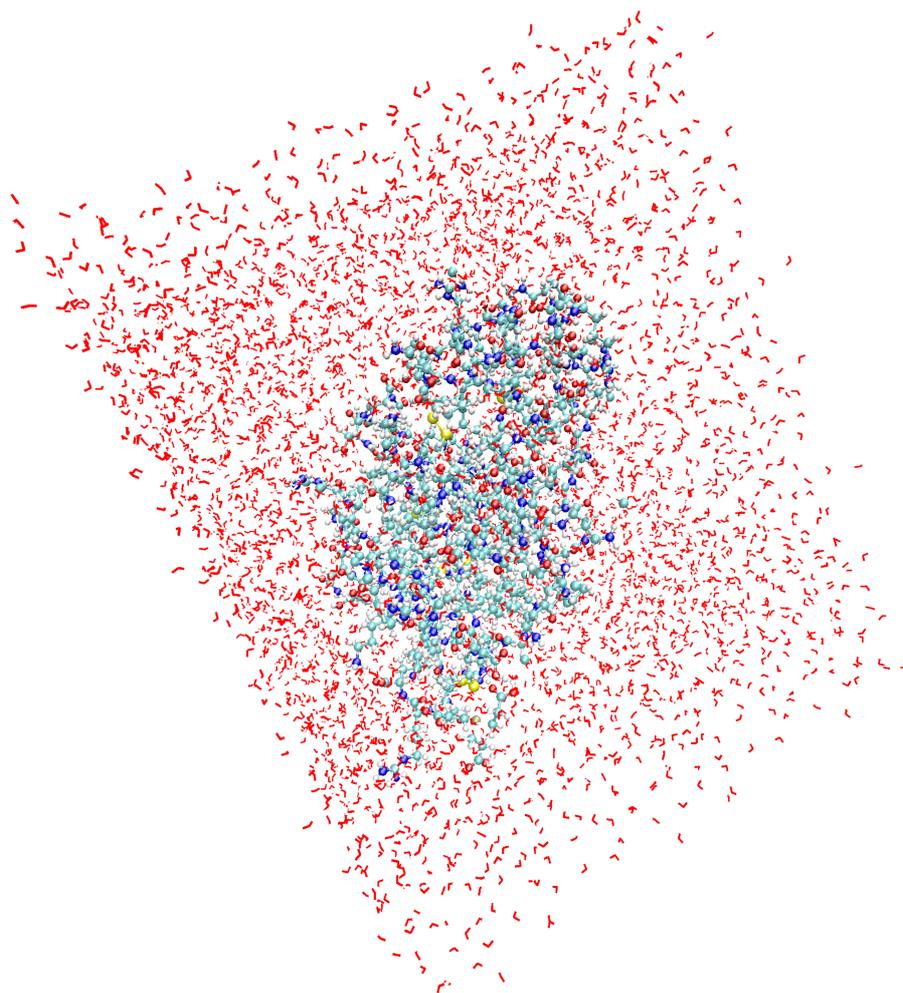


FIGURE 5.2 – Système réel étudié au niveau MM.

Les optimisations QM/MM ont été effectuées sans conditions périodiques. Les structures de départ ont été similaires à celles utilisées dans les calculs MM avec la mention que seulement les molécules d'eau situées à moins de 10 \AA de l'atome de soluté le plus proche ont été retenues (Figure 5.3). Le sous-système QM du système QM/MM est détaillé sur la Figure 5.4.

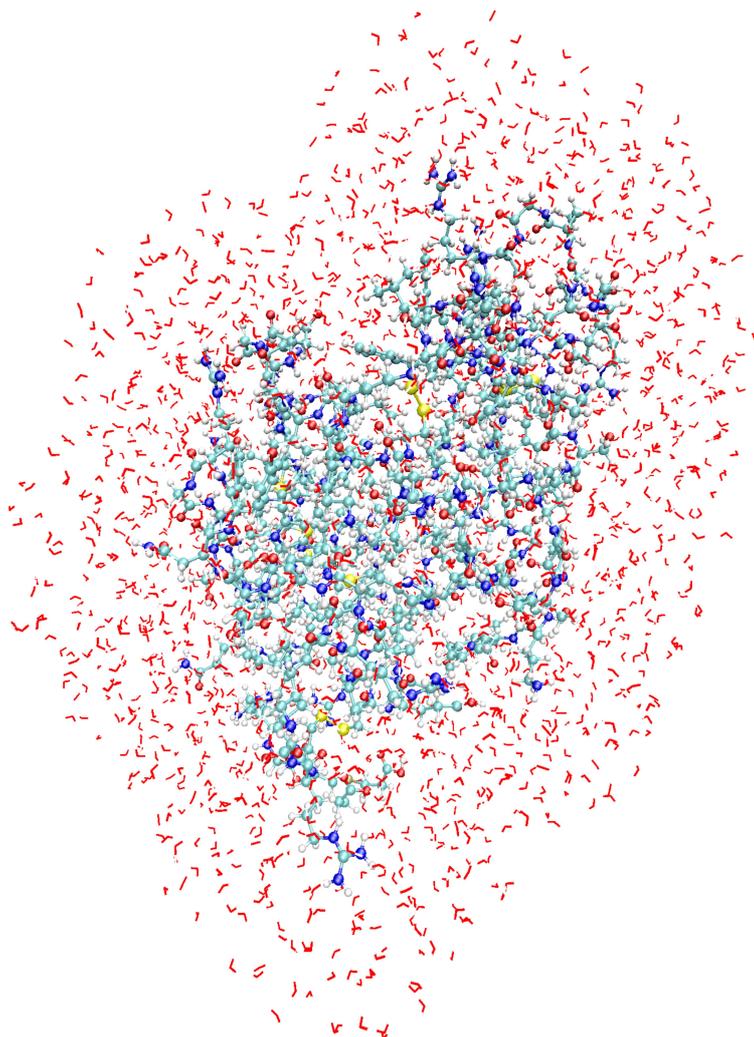


FIGURE 5.3 – Système réel étudié au niveau QM/MM : le pont S-S et la TCEP en présence de molécules d'eau se trouvant à moins de 10 Å de l'atome de soluté le plus proche.

5.4 Résultats

La méthode décrite ci-dessus a été appliquée à l'étude de la première étape de la réaction de réduction des ponts S-S des protéines par la TCEP. Cette première étape (Figure 5.5) est celle qui limite la vitesse de réaction [23, 109]. Elle consiste en la substitution d'un atome S de la liaison S-S par l'atome P avec la formation d'un produit intermédiaire de réaction.

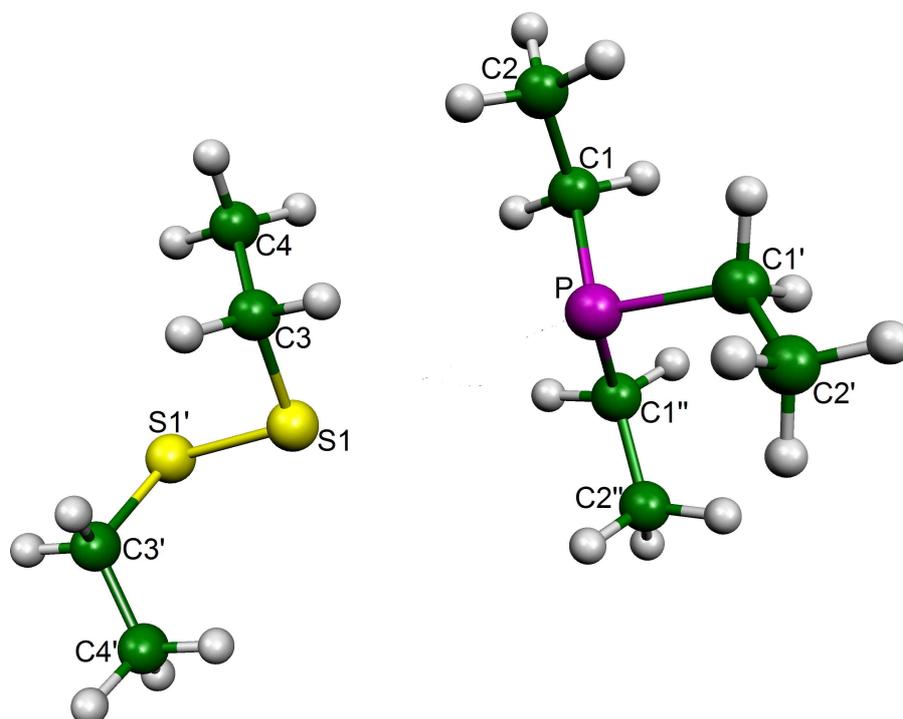


FIGURE 5.4 – Système modèle pour la réaction de réduction des ponts S-S par la TCEP.

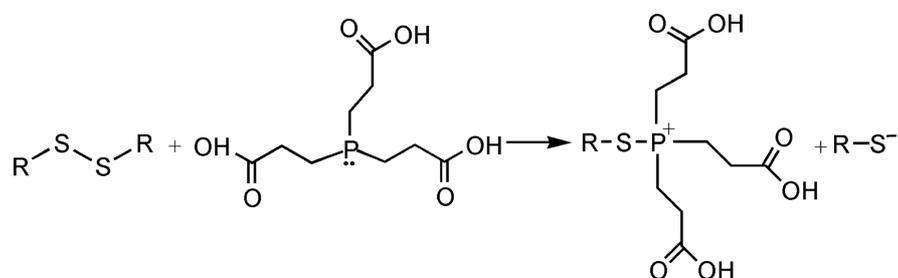


FIGURE 5.5 – Première étape de la réaction de réduction des ponts S-S par la TCEP.

Calculs QM sur le système modèle

Les optimisations partielles (en maintenant la distance S-P fixe) effectuées dans le vide au niveau de théorie B3LYP/6-311+G(d,p) montrent que, lors du rapprochement de la TCEP par rapport au pont S-S, l'énergie du système augmente d'une manière continue (Figure 5.6). Même si pour une distance S-P de 2,0 Å la liaison S-S est affaiblie, elle résiste et il n'y a pas de formation de produit intermédiaire stable.

Une comparaison avec les résultats obtenus en utilisant la méthode semi-empirique

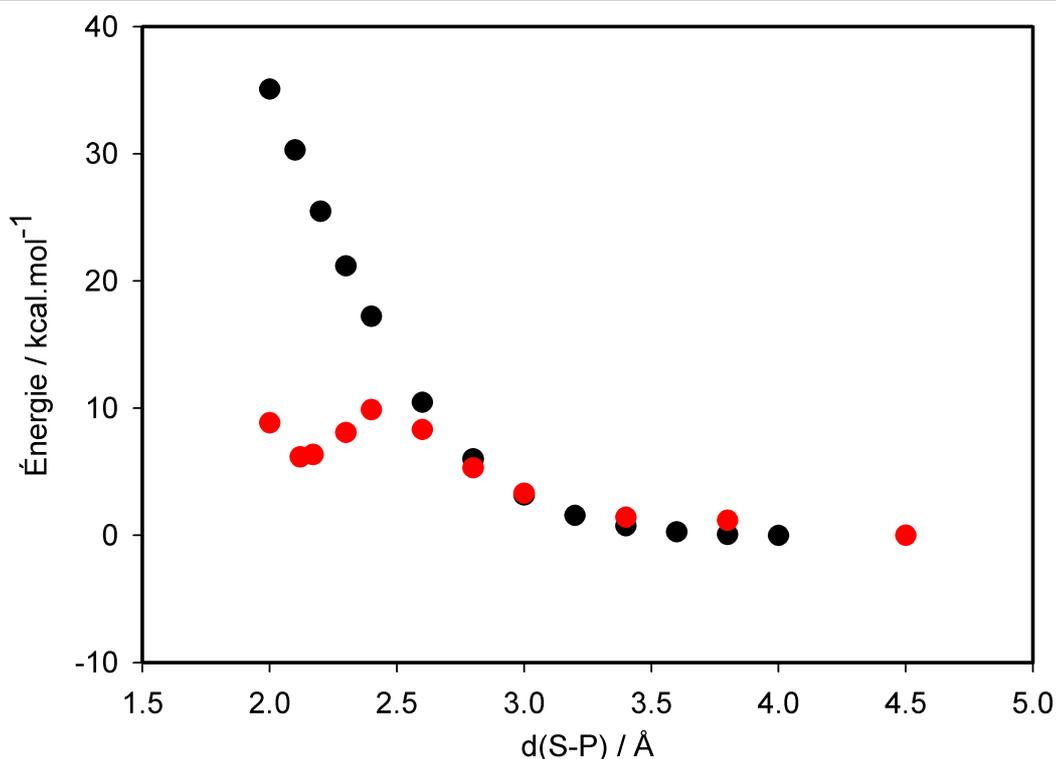


FIGURE 5.6 – Courbes d’énergie pour l’interaction du pont S-S avec la TCEP (système modèle) : en noir l’énergie quantique dans le vide et en rouge l’énergie libre en solution aqueuse (modèle de solvation : PCM).

PM3 montre l’inadéquation de cette dernière méthode à l’étude de cette réaction. En effet, contrairement à la méthode DFT, la méthode PM3 prédit la formation d’un produit de réaction très stable avec un minimum d’énergie qui correspond à une distance S-P de 2,15 Å (Figure 5.7). La liaison S-S persiste dans le produit formé, ce qui est encore plus invraisemblable.

Le calcul au niveau de théorie B3LYP/6-311+G(d,p), en solution aqueuse (modèle de solvation : PCM) indique la présence d’un maximum d’énergie libre pour une distance S-P de 2,4 Å ainsi que la formation d’un produit intermédiaire qui correspond à une distance S-P de 2,1 Å (Figure 5.6). Une réoptimisation sans contrainte de la structure qui correspond à $d(\text{S-P})=2,4$ Å montre que cette structure est très proche d’un état de transition (ET) caractérisé par une fréquence négative. La structure de cet état est présentée sur la Figure 5.8.

La barrière d’énergie de réaction ainsi prédite est de $9,8 \text{ kcal.mol}^{-1}$. Le produit de réaction n’est pas très stable car la barrière prédite pour la réaction inverse est de 4 kcal.mol^{-1} seulement. Les résultats sont en général en accord avec ceux rapportés par Cline et al. [23]. Cependant ces auteurs n’ont pas pu localiser l’état ET en utilisant

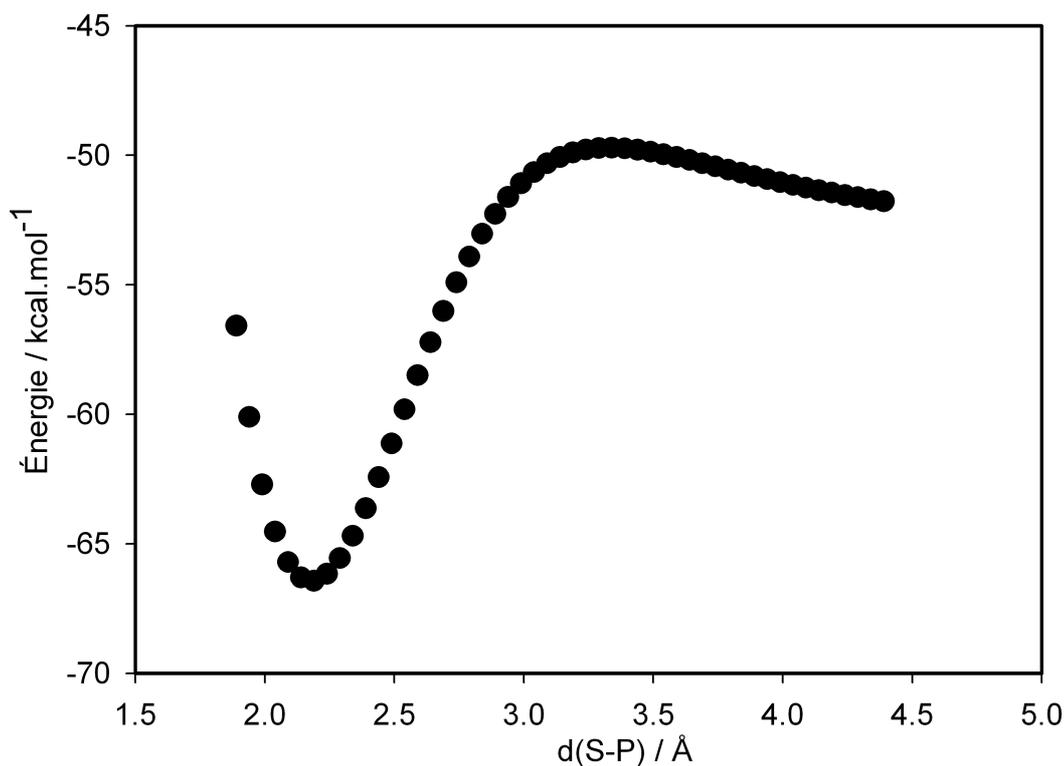


FIGURE 5.7 – Étude de l’interaction du pont S-S avec la TCEP dans le vide par la méthode semi-empirique PM3.

simplement le modèle de solvation du continuum diélectrique, mais seulement après l’ajout de 1-2 molécules d’eau explicites en interaction avec les réactifs.

La trajectoire de réaction déterminée par la méthode QM-PCM a servi ensuite comme trajectoire préliminaire de réaction dans les calculs QM-FE.

Calcul QM-FE sur la réduction du GSSG

Dans le cas du GSSG, le pont S-S est très bien solvatoé et on s’attend à ce que la courbe d’énergie libre soit proche de celle du système modèle. L’étude de la réduction du GSSG par la TCEP présente un double intérêt. D’une part elle sert à valider la méthode, d’autre part, la réduction du pont S-S dans le GSSG représente un terme de comparaison naturel pour la réduction des ponts S-S dans les protéines.

La courbe d’énergie calculée en utilisant la méthode QM-FE décrite précédemment (équation 5.20) est présentée sur la Figure 5.9. Nous avons suivi l’interaction entre les deux réactifs jusqu’à une distance maximum de 7,0 Å. Étant donné qu’à cette distance la courbe d’énergie devient plate (dans la limite de la dispersion accidentelle des points), nous avons choisi la valeur de l’énergie libre du système à 7,0 Å comme origine de l’échelle

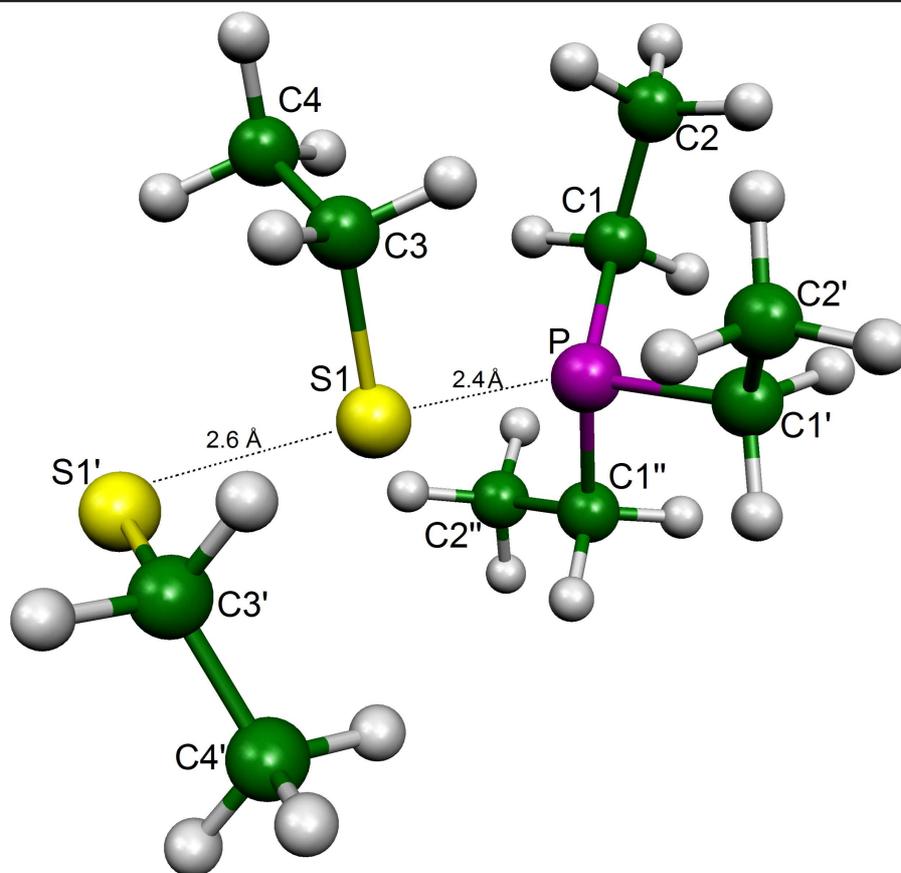


FIGURE 5.8 – État de transition (ET) de la réaction de réduction d'un pont S-S par la TCEP (système modèle) en solution aqueuse. Les calculs ont été effectués au niveau de théorie B3LYP/6-311+G(d,p) en utilisant le modèle de solvatation PCM.

d'énergie libre.

Par rapport aux résultats obtenus par la méthode QM-PCM appliquée au système modèle, la barrière d'énergie est plus haute d'environ $5,0 \text{ kcal.mol}^{-1}$ et déplacée à $d(\text{S-P})=2.6 \text{ \AA}$. La différence la plus importante entre les deux méthodes, QM-FE et QM-PCM, est que le produit de réaction est beaucoup plus stable dans le cas de la méthode QM-FE. Cette stabilité est en accord avec les données expérimentales qui indiquent la première étape de réaction comme étape limitante de la vitesse de réaction. Cela signifie que le produit formé après cette étape ne pourrait pas être instable par rapport aux réactifs.

Malgré leurs différences, les deux courbes d'énergie ne sont pas incompatibles, ce qui confirme *a posteriori* le choix approprié de la trajectoire de réaction QM-PCM comme trajectoire préliminaire pour les calculs QM-FE.

Une analyse de l'origine des différences entre les deux courbes montre que les différences de stabilité du produit sont dues à l'énergie d'interaction avec l'environnement. Les deux

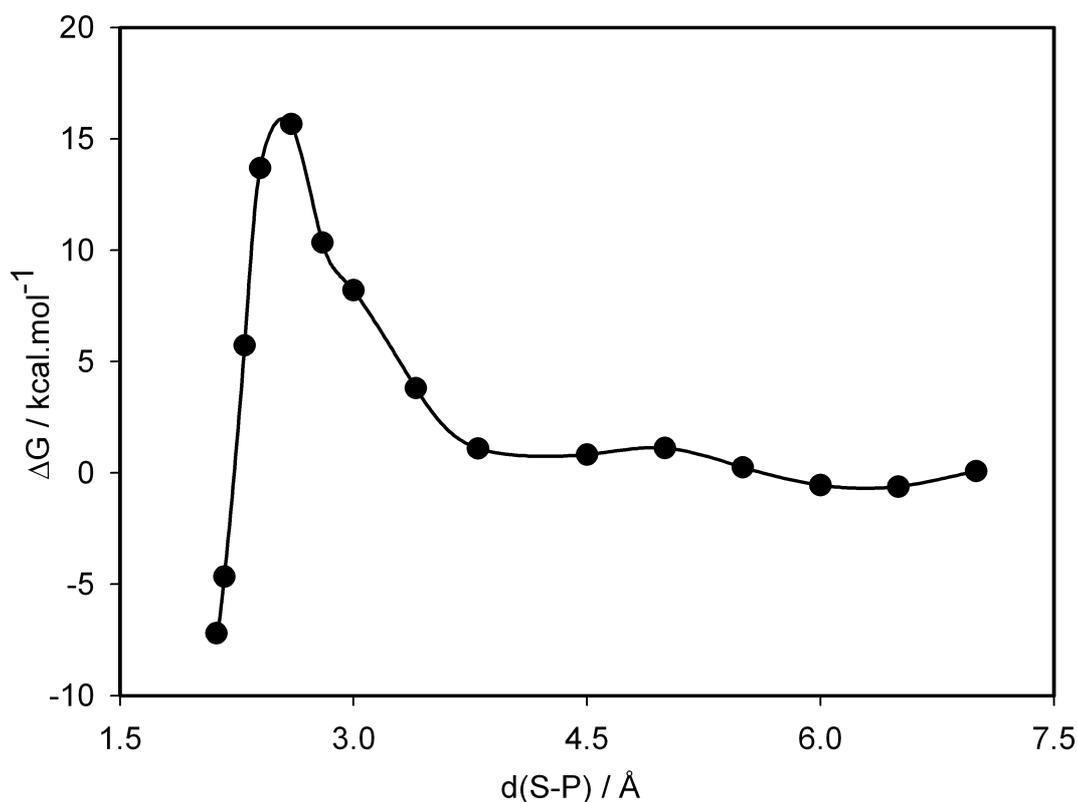


FIGURE 5.9 – Courbe d’enthalpie libre de la réaction de réduction du GSSG par la TCEP, calculée par la méthode QM-FE.

méthodes prédisent des valeurs différentes pour cette énergie. Ainsi, dans le calcul QM-FE sur le GSSG, l’énergie d’interaction du produit avec l’environnement est plus basse que celle des réactifs de 62 kcal.mol^{-1} . D’autre part, dans le calcul QM-PCM cette différence est de 48 kcal.mol^{-1} seulement.

Cependant, la différence entre les barrières d’énergie de réaction calculées dans les deux cas est due surtout aux contributions quantiques. Évidemment, les différences entre ces contributions quantiques ne proviennent pas des différences entre les méthodes de type QM, car elles sont les mêmes, mais des légères différences entre les géométries du système modèle dans les deux cas. Ces différences de géométrie sont certainement dues, elle aussi, à l’interaction avec l’environnement.

Calculs QM-FE sur la réduction des quatre ponts S-S dans la lysozyme

Les courbes d’énergie libre calculées pour la réduction des quatre ponts S-S de la lysozyme sont présentées sur la Figure 5.10. Pour chaque courbe, l’origine de l’énergie libre

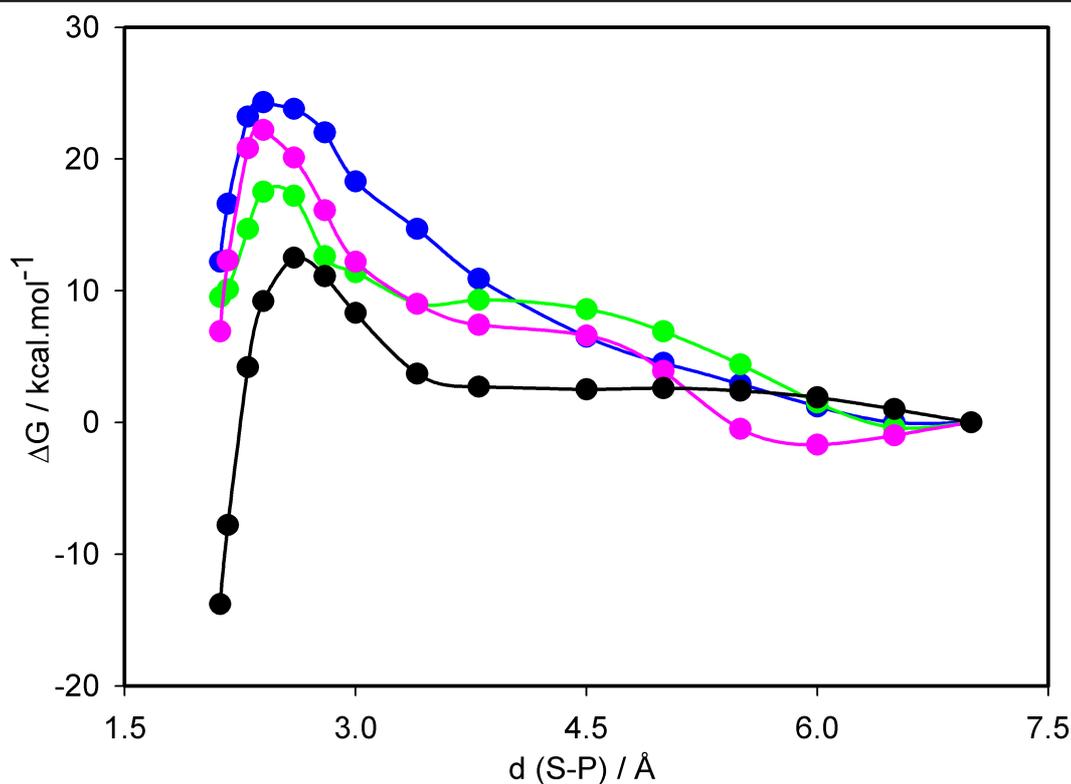


FIGURE 5.10 – Courbes d’enthalpie libre de la réaction de réduction des ponts S-S dans la lysozyme par la TCEP calculées en utilisant la méthode QM-FE : cys6-cys127 (noir), cys30-cys115 (bleu), cys64-cys80 (vert) et cys76-cys94 (rose).

a été choisie au point $d(S-P)=7,0$ Å. On constate que dans le cas des ponts S-S de la lysozyme, les barrières d’énergie varient entre $12,6 \text{ kcal.mol}^{-1}$, pour le pont cys6-cys127 et $24,3 \text{ kcal.mol}^{-1}$, pour le pont cys30-cys115. Trois de ces barrières sont supérieures à celle calculée pour le GSSG, tandis que la barrière calculée pour le pont le plus exposé, cys6-cys127, est inférieure d’environ $3,0 \text{ kcal.mol}^{-1}$. On constate aussi que les différences d’accessibilité du site réactif sont visibles dans la montée des courbes entre $7,0$ et $4,0$ Å. Dans le cas du pont cys6-cys127 on note l’accord entre la grande accessibilité du site réactif indiquée par l’absence de cette montée et le degré d’exposition du pont au solvant.

Un premier facteur qui pourrait expliquer les différences entre les quatre courbes est la distorsion du site actif soumis aux contraintes par son environnement. Cette distorsion affecte l’énergie quantique E_{QM} . La variation de cette énergie en fonction de la distance S-P est présentée sur la Figure 5.11 pour les quatre ponts de la lysozyme et pour le GSSG. Pour toutes les courbes, l’origine de l’énergie E_{QM} correspond à l’énergie E_{QM} du produit

de réaction.

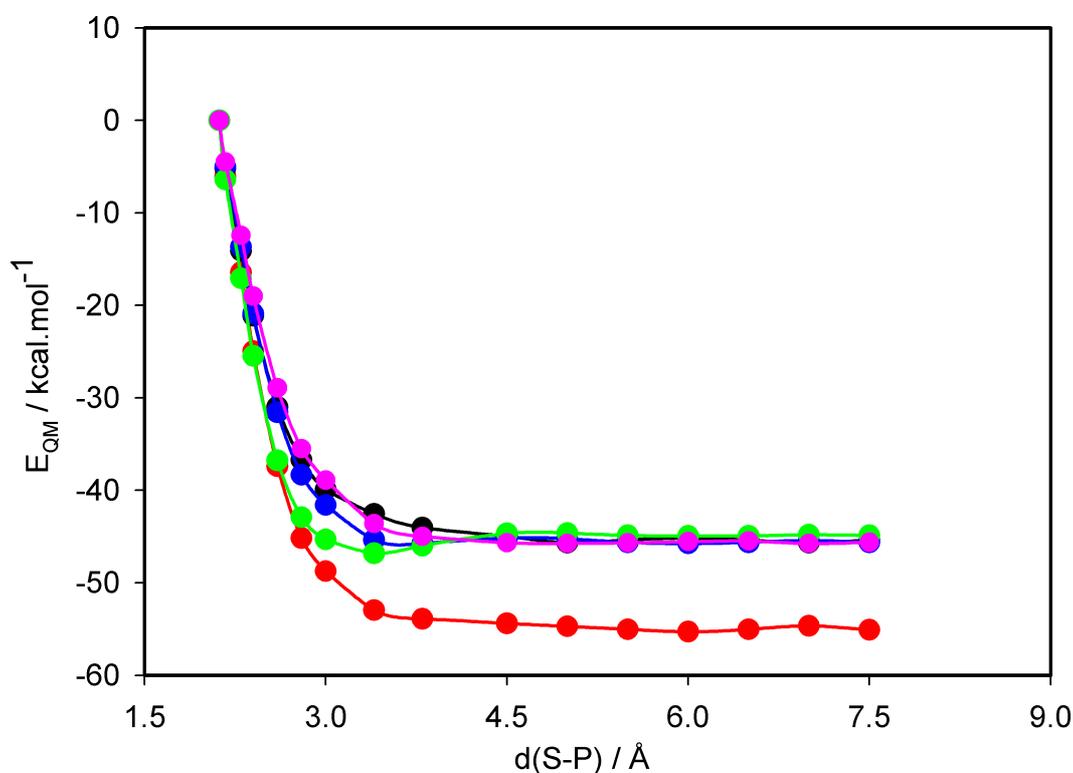


FIGURE 5.11 – Variation de l'énergie E_{QM} en fonction de la distance $d(S-P)$: cys6-cys127 (noir), cys30-cys115 (bleu), cys64-cys80 (vert), cys76-cys94 (rose) et GSSG (rouge).

Dans le cas du GSSG, la différence d'énergie E_{QM} entre les réactifs et le produit est d'environ $-55,0 \text{ kcal.mol}^{-1}$, tandis que pour les quatre ponts de la lysozyme cette différence est de $-45,0 \text{ kcal.mol}^{-1}$ seulement. Cette augmentation relative de l'énergie des réactifs dans la protéine est certainement due à la distorsion des ponts S-S. La distorsion des ponts affecte aussi d'une manière significative les barrières d'énergie de réaction. Nous avons représenté dans le Tableau 5.3 les contributions quantiques respectives à la différence d'enthalpie libre entre les points $d(S-P)=2,4 \text{ Å}$ et $d(S-P)=7,0 \text{ Å}$. Ces contributions sont plus faibles dans le cas des ponts S-S de la protéine. Cela est en accord avec l'hypothèse de l'existence d'une distorsion des ponts S-S qui fait monter l'énergie des réactifs par rapport à l'état de transition. En effet, la distorsion serait relaxée partiellement dans l'état de transition et totalement dans le produit. Cette relaxation semble avancer plus vite dans le cas des ponts cys64-cys80 pour lequel la contribution quantique à la barrière d'énergie est plus basse.

Pont S-S \ Contributions	$\Delta E_{QM}/kcal.mol^{-1}$	$\Delta F_{env}/kcal.mol^{-1}$
GSSG	30,5	-16,0
cys6-cys127	24,7	-14,0
cys30-cys115	24,5	1,1
cys64-cys80	19,0	-4,2
cys76-cys94	26,7	-5,0

Tableau 5.3 – Contributions à la barrière d'énergie de réaction. Sont représentées les différences ΔE_{QM} et ΔF_{env} entre $d(S-P)=2,4 \text{ \AA}$ et $d(S-P)=7,0 \text{ \AA}$.

Une deuxième contribution à prendre en compte dans l'analyse des courbes d'enthalpie libre (Figure 5.10) est l'énergie d'interaction entre le centre actif (le système modèle) et l'environnement. Les courbes de l'enthalpie libre ΔF_{env} , sont présentées sur la Figure 5.12.

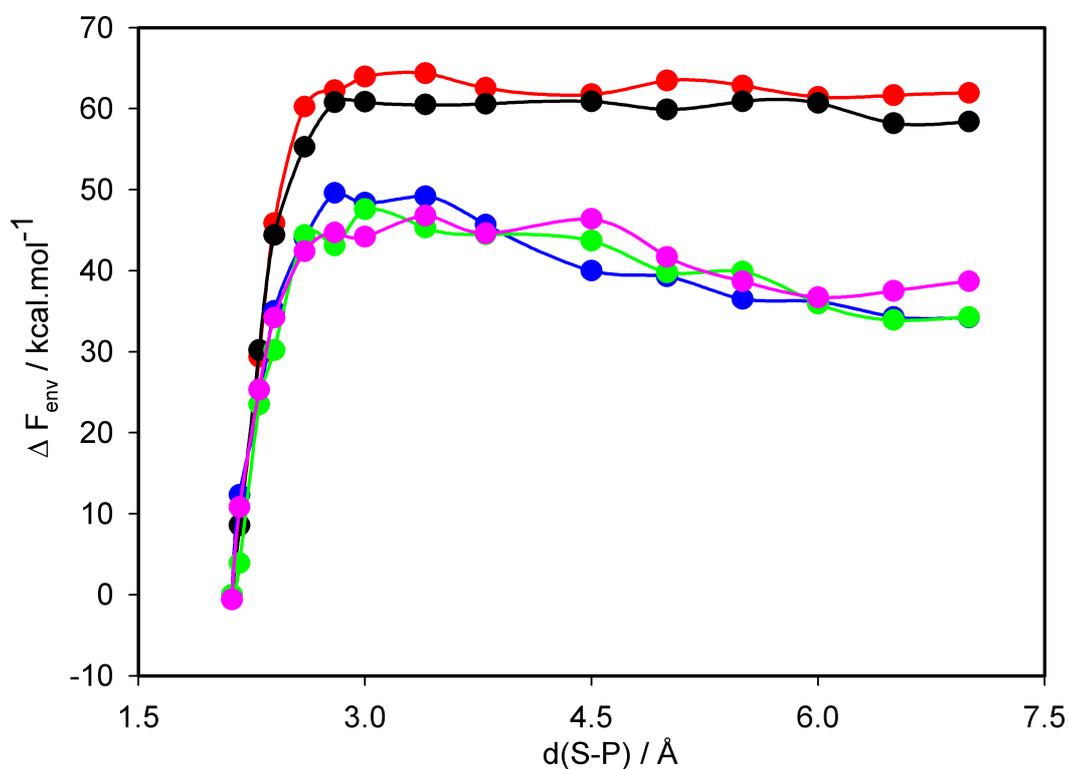


FIGURE 5.12 – Variation de l'énergie libre d'interaction avec l'environnement en fonction de la distance $d(S-P)$: cys6-cys127 (noir), cys30-cys115 (bleu), cys64-cys80 (vert), cys76-cys94 (rose) et GSSG (rouge).

On remarque que la courbe qui correspond au pont S-S exposé au solvant, cys6-cys127, est très proche de celle du GSSG, tandis que les courbes des trois autres ponts se ressemblent beaucoup. D'une manière générale, l'enthalpie d'interaction avec l'environnement dépend de deux facteurs : l'accessibilité du pont et la polarisabilité de cet environnement. Il est difficile de séparer les contributions des deux facteurs au long de la trajectoire de réaction. Cependant, entre $d(S-P)=7,0$ et $d(S-P)=4,0$ Å, la polarité du centre actif change peu et on peut attribuer la montée des courbes dans cet intervalle au premier facteur : c'est à dire, à l'opposition de l'environnement protéique au rapprochement des réactifs. Cette opposition est nulle dans le cas du pont cys6-cys127. Dans le cas des trois autres ponts elle contribue en moyenne d'environ 10 kcal.mol^{-1} à la montée de l'enthalpie libre.

Entre $d(S-P)=4,0$ et $d(S-P)=2,0$ Å, la polarité du centre actif change beaucoup et l'effet de la polarisabilité de l'environnement sur les courbes d'enthalpie libre devrait être important. De plus, l'accessibilité du pont et la polarisabilité de son environnement doivent être corrélés d'une certaine manière : plus le pont est exposé au solvant, plus son environnement est polarisable.

Globalement, l'interaction avec l'environnement fait monter l'enthalpie libre des réactifs par rapport à celle du produit d'environ 60 kcal.mol^{-1} dans le cas du GSSG et du pont cys6-cys127 et d'environ 35 kcal.mol^{-1} dans le cas des autres ponts (Figure 5.12). Elle a aussi des contributions importantes aux barrières d'énergie de réaction (Tableau 5.3). On observe une différence remarquable entre le GSSG et le pont cys6-cys127, d'une part, et les autres ponts, d'autre part. Conformément à ces résultats, l'interaction avec l'environnement fait baisser fortement la barrière d'énergie de réaction dans le cas du GSSG et du pont cys6-cys127.

5.5 Discussion

Avant d'analyser plus loin la signification des résultats présentés ci-dessus, il est utile d'examiner le rapport entre ces résultats et les résultats expérimentaux rapportés dans le Chapitre 4. Ainsi, pour la réduction du GSSG par la TCEP nous avons trouvé une constante de vitesse de $15,2 \text{ M}^{-1}\text{s}^{-1}$. En supposant que la première étape de cette réaction (Figure 5.5) est celle qui détermine la vitesse de réaction et en appliquant la formule d'Eyring (équation 2.29), on trouve une barrière d'énergie d'environ $16,5 \text{ kcal.mol}^{-1}$.

De la même manière, dans le cas de la réduction des ponts S-S dans la lysozyme, on trouve,

à partir des constantes de vitesse $k_f^{(1)} = 0,45M^{-1}s^{-1}$ et $k_f^{(2)} = 0,052M^{-1}s^{-1}$, deux barrières d'énergie : de 18 kcal.mol^{-1} et de $19,5 \text{ kcal.mol}^{-1}$, respectivement. Ces valeurs ne sont pas à comparer directement avec les barrières d'énergie dans les Figures 5.9 et 5.10. En effet, ces dernières valeurs ne représentent pas les vraies barrières d'énergie de la réaction considérée puisque l'état de référence choisi n'est pas celui des réactifs séparés, mais celui des réactifs se trouvant à une distance fixe de $7,0 \text{ \AA}$. De plus, ces énergies libres sont des énergies libres partielles calculées avec une coordonnée gelée. Dans cette situation, il nous reste à comparer les valeurs relatives de ces barrières d'énergie.

Les barrières calculées par la méthode QM-FE sont les suivantes : $12,5 \text{ kcal.mol}^{-1}$ pour le pont cys6-cys127, $17,5 \text{ kcal.mol}^{-1}$ pour le pont cys64-cys80, $22,3 \text{ kcal.mol}^{-1}$ pour le pont cys76-cys94 et $24,5 \text{ kcal.mol}^{-1}$ pour le pont cys30-cys115. Dans le cas du GSSG, la barrière de réaction théorique est de $15,5 \text{ kcal.mol}^{-1}$.

Étant donné que les deux valeurs de barrière d'énergie expérimentales sont proches l'une de l'autre, on tire la conclusion que les quatre ponts S-S de la lysozyme doivent avoir des barrières d'énergie assez similaires pour leur réduction. En prenant en compte l'incertitude probable dans l'évaluation des valeurs expérimentales ($\pm 1 \text{ kcal.mol}^{-1}$), on conclut que les barrières d'énergie pour la réduction de quatre ponts S-S devrait se situer dans l'intervalle $17\text{-}21 \text{ kcal.mol}^{-1}$.

En termes relatifs, les barrières d'énergie pour la réduction des ponts S-S dans la lysozyme devraient être supérieure à celle du GSSG de $0,5$ à $4,5 \text{ kcal.mol}^{-1}$. On constate que la barrière calculée pour le pont cys6-cys127 est inférieure de 3 kcal.mol^{-1} par rapport à celle calculée pour le GSSG. D'autre part, la barrière calculée pour le pont cys30-cys115 dépasse celle du GSSG de 9 kcal.mol^{-1} . Ces valeurs sortent de l'intervalle de variation attendu. Cependant, ces déviations tombent dans le domaine d'incertitude des méthodes utilisées dans le calcul présent. En conclusion, l'accord entre les valeurs calculées et les valeurs expérimentales peut être considéré comme satisfaisant.

Un aspect intéressant à examiner à la lumière des résultats théoriques est l'adéquation du modèle de déplieement local à la description de la réaction de réduction des ponts S-S des protéines. Dans l'analyse des données expérimentales, nous avons attribué les différences entre les constantes de vitesse de réduction des ponts S-S de la lysozyme et du GSSG au facteur accessibilité. Cependant, les résultats théoriques montrent qu'il existe d'autres facteurs qui contribuent à ces différences. Parmi ces facteurs, on compte la distorsion des ponts en interaction avec leur environnement et la polarisabilité de cet environnement.

L'effet de l'accessibilité des ponts sur l'enthalpie libre est visible sur la Figure 5.10

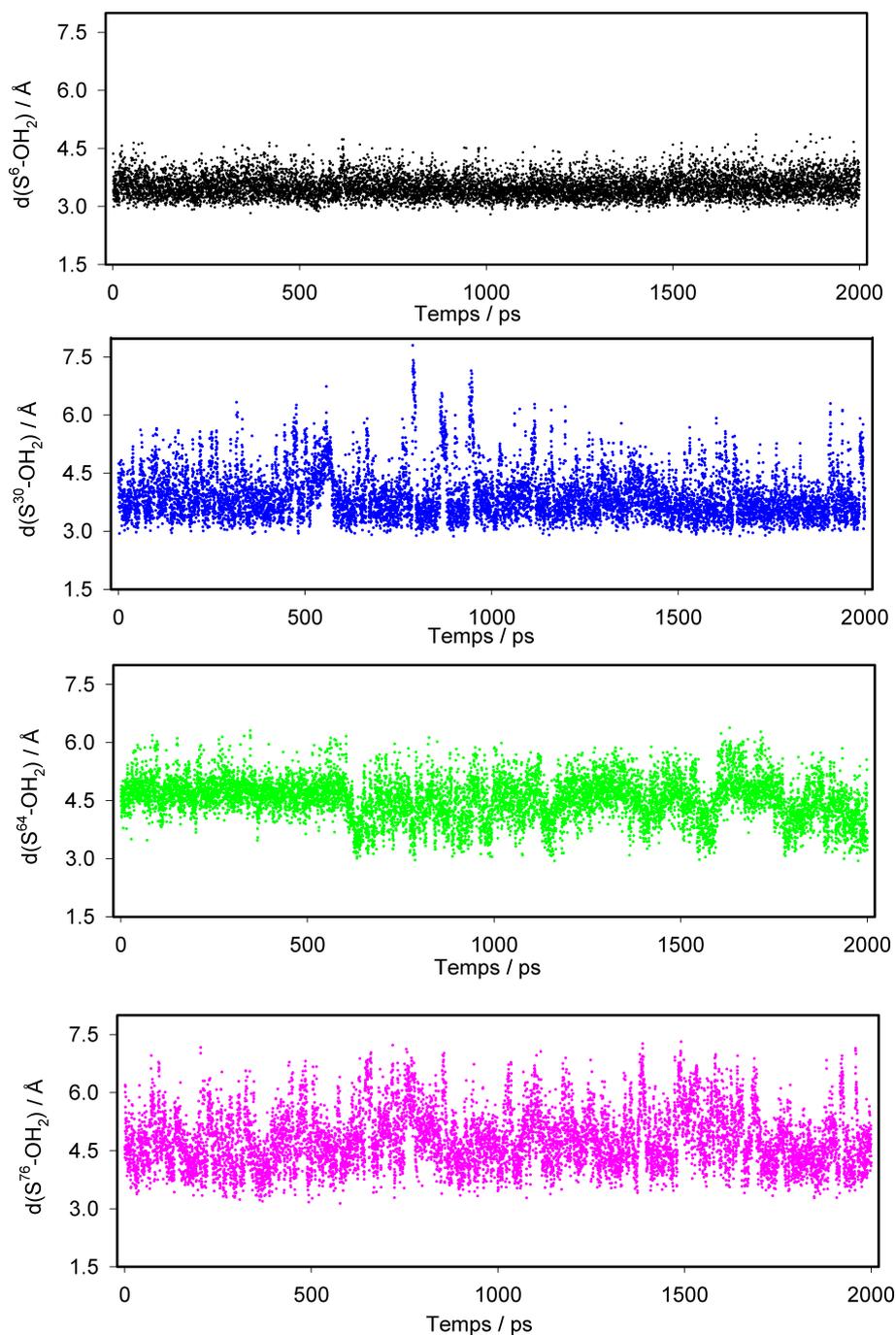


FIGURE 5.13 – Variation de la distance entre l'atome de soufre extérieur du pont et la molécule d'eau la plus proche au long de la trajectoire simulée de la lysozyme native : cys6 (noir), cys30(bleu), cys64(vert) et cys76 (rose).

dans la montée des courbes entre 7,0 et 4,0 Å. Quant au processus par lequel l'agent réducteur arrive en contact avec le pont S-S protégé par la structure protéique, il n'est pas évident de conclure s'il s'agit d'un dépliement local ou d'une simple diffusion. Nous

avons essayé de répondre à cette question en effectuant des simulations supplémentaires. Le paramètre choisi pour cette analyse est la distance S-O entre l'atome de soufre le plus exposé du pont S-S et la molécule d'eau la plus proche (Figure 5.13).

La distribution des distances obtenue dans le cas du pont cys6-cys127 montre que l'atome de soufre le plus exposé du pont est en contact permanent avec le solvant, la distance moyenne étant d'environ 3,8 Å.

Dans le cas du pont cys64-cys80, la distribution des distances met en évidence une certaine périodicité qui pourrait correspondre à un dépliement local de la protéine. Ainsi, dans l'intervalle 0-600 ps, la distance minimum entre l'atome de soufre 80 et la molécule d'eau la plus proche est limitée à 4,0 Å, tandis qu'entre 600-1200 ps la distance minimum descend à 3,0 Å. D'autres intervalles d'alternance plus petits suivent.

Dans le cas des ponts cys30-cys115 et cys76-cys94 on n'observe pas de périodicité dans la distribution des distances S-O. D'autre part, la distance S-O moyenne dans le cas du pont cys76-cys94 est d'environ 4,5 Å, tandis que dans le cas du pont cys30-cys115 cette distance est de 4,0 Å. Cela indique une meilleure accessibilité du pont cys30-cys115, malgré sa barrière de réaction supérieure. Cette contradiction apparente peut être résolue si on accepte, en accord avec la discussion précédente, que l'accessibilité n'est pas le seul facteur dont la barrière de réaction dépend.

Les résultats présentés ci-dessus indiquent que le modèle de dépliement local est valable au moins dans le cas du pont cys64-cys80. Pour les trois autres ponts, l'analyse de la distribution des distances S-O au long d'une trajectoire de 2 ns n'a pas permis de mettre en évidence un tel dépliement. Des analyses supplémentaires d'accessibilité plus sophistiquées pourraient être envisagées afin de mieux clarifier cet aspect.

Conclusions

Nous avons réalisé une analyse détaillée, expérimentale et théorique, des mécanismes de réduction des ponts disulfure dans les albumines sériques et le lysozyme par deux agents réducteurs très utilisés dans les expériences sur le repliement et le dépliement des protéines. L'analyse a été orientée vers la compréhension du rôle que l'environnement joue dans cette réaction.

Dans cette étude, une méthodologie originale a été mise au point et employée avec succès. Cette méthodologie est basée sur l'association étroite de la spectroscopie Raman et du calcul moléculaire.

La spectroscopie Raman nous a permis de déterminer directement l'avancement de la réaction de réduction des ponts en détectant les changements des bandes Raman S-S. D'autre part, les changements de conformation qui accompagnent la coupure des ponts S-S ont été suivis en détectant la bande Raman amide I. Cette démarche a été rendue possible par un travail systématique d'analyse et d'attribution des bandes S-S dans les spectres Raman des protéines.

Les résultats expérimentaux ainsi obtenus ont été soumis à un traitement complexe qui nous a donné accès à des paramètres physico-chimiques de grande importance dans la compréhension des mécanismes étudiés. Ainsi, nous avons déterminé les enthalpies libre de réaction et les enthalpies libre du dépliement local de la structure protéique. En combinant les données expérimentales et les calculs *ab initio* nous avons évalué la variation de l'enthalpie conformationnelle globale de la protéine qui accompagne la réduction des ponts S-S dans les albumines sériques. Les analyses de cinétique de réaction nous ont permis de déterminer avec précision les constantes de vitesse de réaction et de mettre en évidence les différences de réactivité entre les ponts S-S.

Le calcul moléculaire a ensuite été utilisé afin de compléter et d'affiner les résultats expérimentaux. Il a été appliqué à l'étude de la réduction des quatre ponts S-S du lysozyme par la TCEP. Les calculs ont été basés sur l'application dans une forme originale de la

méthode QM-FE. Les barrières d'énergie de réaction ainsi calculées sont en bon accord avec les résultats expérimentaux. De plus, la modélisation moléculaire nous a permis de séparer les différentes contributions à la barrière d'énergie, ce qui nous a facilité l'analyse des mécanismes de réaction.

De cette étude se dégage un modèle de réaction de réduction des ponts S-S dans des protéines dans lequel le rôle de l'environnement protéique est précisément défini. Il apparaît ainsi que :

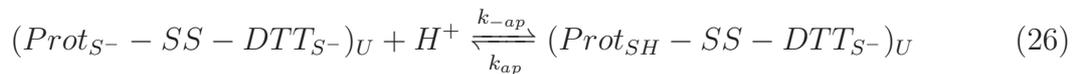
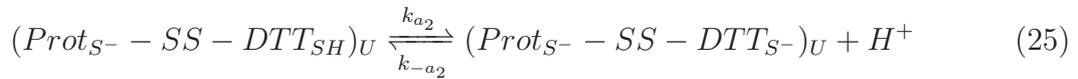
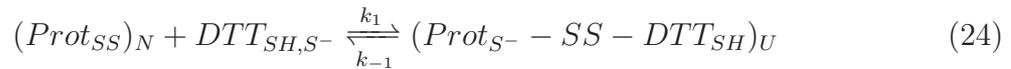
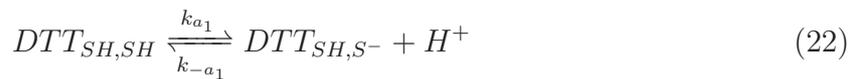
- les contraintes imposées par cet environnement au centre actif peuvent induire des distorsions du pont S-S en déstabilisant d'une manière significative les réactifs par rapport au produit.
- la polarisabilité de l'environnement, directement liée à l'exposition du pont S-S au solvant, joue un rôle essentiel dans la baisse de la barrière de réaction.
- l'accessibilité du pont S-S par l'agent réducteur est le troisième facteur ayant une contribution importante à la réactivité des ponts S-S dans les protéines. Un dépliement local peut intervenir à ce point en rendant le pont S-S plus accessible.

En conclusion, notre travail marque un progrès significatif dans la compréhension de la réactivité des ponts S-S dans la protéine. La complexité du sujet fait que certains aspects traités dans cette étude devraient être développés ultérieurement. Il s'agit, par exemple, de l'analyse des bandes S-S dans les systèmes modèles ou de la nature du dépliement qui accompagne la réduction des ponts S-S dans les protéines. Au niveau de la modélisation moléculaire, des progrès techniques sont attendus qui permettraient le calcul de la trajectoire de réaction préliminaire au niveau QM/MM (par l'assemblage électronique). Il mérite aussi d'être souligné que l'application des méthodes que nous avons mises au point dans cette étude peut être étendue à d'autres problématiques concernant la réactivité en milieu biologique.

Annexe

Modèle détaillé de la réaction de réduction des ponts S-S par le DTT

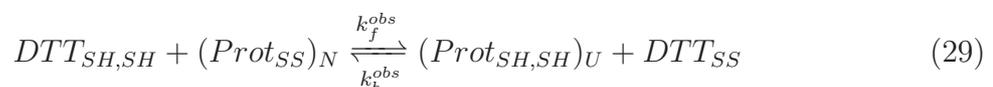
Les étapes de réaction :



Où : k_{a_1} et k_{a_2} sont les deux constantes d'acidité du DTT réduit, k_{ap} est la constante d'acidité des thiols de la protéine après la réduction, k_u et k_n sont les constantes des vitesses de dépliement et de repliement local de la protéine, respectivement. L'indice u indique le dépliement local qui précède la réduction des ponts tandis que l'indice U indique le dépliement induit par la réduction d'un pont.

On suppose aussi que la coupure des ponts S-S et le dépliement qui l'accompagne sont pratiquement des processus simultanés.

La réaction globale de réduction d'un pont S-S par le DTT dans les protéines est exprimée par l'équation suivante :



Réduction des ponts S-S par le DTT : relations valables à l'équilibre

La constante d'équilibre de la réaction exprimée par l'équation 29 est donnée par la relation :

$$K_{eq}^{obs} = \frac{[(Prot_{SH,SH})_U]_{eq}[DTT_{SS}]_{eq}}{[DTT_{SH,SH}]_{eq}[(Prot_{SS})_N]_{eq}} \quad (30)$$

En supposant que la concentration de ponts mixtes ($Prot_{SS} - DTT_{SS}$) est négligeable par rapport aux autres espèces, on peut écrire :

$$[(Prot_{SH,SH})_U] + [(Prot_{SH,S^-})_U] + [(Prot_{S^-,S^-})_U] + [(Prot_{S^-,SH})_U] = [DTT_{SS}] \quad (31)$$

Vue que :

$$[(Prot_{SH,S^-})_U] = [(Prot_{S^-,SH})_U] = [(Prot_{SH,SH})_U] \times 10^{(pH-pK_{ap})} \quad (32)$$

et que

$$[(Prot_{S^-,S^-})_U] = [(Prot_{SH,SH})_U] \times 10^{2(pH-pK_{ap})} \quad (33)$$

on obtient :

$$[(Prot_{SH,SH})_U] = \frac{[DTT_{SS}]}{1 + 2 \times 10^{(pH-pK_{ap})} + 10^{2(pH-pK_{ap})}} \quad (34)$$

Ici pK_{ap} est la constante d'acidité des thiols dans la protéine après réduction. D'une manière similaire on obtient :

$$[DTT_{SH,SH}] = \frac{[DTT_{SH,SH}]_I - [DTT_{SS}]}{1 + 2 \times 10^{(pH-pK_{a1})} + 10^{(2pH-pK_{a1}-pK_{a2})}} \quad (35)$$

et :

$$[(Prot_{SS})_N] = [(Prot_{SS})_N]_I - [DTT_{SS}] \quad (36)$$

où pK_{a1} et pK_{a2} sont les constantes d'acidité du DTT, $[DTT_{SH,SH}]_I$ est la concentration initiale du DTT réduit (ou la concentration totale de DTT) et $[(Prot_{SS})_N]_I$ est la concentration initiale de ponts S-S de la protéine. Après la substitution des équations 34, 35 et 36 dans l'équation 30, l'expression de la constante d'équilibre devient :

$$K_{eq}^{obs} = \frac{\alpha [DTT_{SS}]_{eq}^2}{\left([DTT_{SH,SH}]_I - [DTT_{SS}]_{eq}\right) \left([(Prot_{SS})_N]_I - [DTT_{SS}]_{eq}\right)} \quad (37)$$

où :

$$\alpha = \frac{1 + 2 \times 10^{(pH-pK_{a1})} + 10^{(2(pH-pK_{a1}))}}{1 + 2 \times 10^{(pH-pK_{ap})} + 10^{2(pH-pK_{ap})}} \quad (38)$$

La résolution de l'équation 37 par rapport au produit $[DTT_{SS}]$ nous permet d'obtenir :

$$[DTT_{SS}]_{eq} = \left\{ K_{eq}^{obs} \left([(Prot_{SS})_N]_I + [DTT_{SH,SH}]_I \right) - \left((K_{eq}^{obs})^2 \left([(Prot_{SS})_N]_I + [DTT_{SH,SH}]_I \right)^2 + 4K_{eq}^{obs} \alpha [(Prot_{SS})_N]_I [DTT_{SH,SH}]_I \right)^{1/2} \right\} / 2(K_{eq}^{obs} - \alpha) \quad (39)$$

C'est la fonction modèle qui a été utilisée pour traiter les données expérimentales concernant l'équilibre de la réaction 29.

Réduction des ponts S-S par le DTT : traitement de la cinétique de réaction

L'équation de vitesse pour la réduction du DTT oxydé peut être exprimée en utilisant des couples des réactifs alternatifs :

$$\frac{d[DTT_{SS}]}{dt} = k_f^{obs}[DTT_{SH,SH}][(Prot_{SS})_N] - k_b^{obs}[DTT_{SS}][(Prot_{SH,SH})_U] \quad (40)$$

$$\frac{d[DTT_{SS}]}{dt} = k_f[DTT_{SH,S^-}][(Prot_{SS})_N] - k_b[DTT_{SS}][(Prot_{SH,S^-})_U] \quad (41)$$

$$\frac{d[DTT_{SS}]}{dt} = k_f^0[DTT_{SH,S^-}][(Prot_{SS})_u] - k_b[DTT_{SS}][(Prot_{SH,S^-})_U] \quad (42)$$

Ici k_f^{obs} et k_b^{obs} sont les constantes apparentes de vitesse de la réaction directe et inverse, respectivement, k_f et k_b sont les constante intrinsèques de vitesse de la réaction directe et inverse, respectivement, et k_f^0 est la constante intrinsèque de vitesse de la réaction directe de réduction d'un pont complètement accessible. Elle est égale à la constante de réaction de réduction d'un pont S-S modèle, tel que GSSG.

À l'équilibre $\frac{d[DTT_{SS}]}{dt} = 0$. Cela implique :

$$\frac{k_f^{obs}}{k_b^{obs}} = K_{eq}^{obs} \quad (43)$$

En comparant les équations 39 et 41, on obtient :

$$k_f^{obs} = k_f \times 2 \times (10^{(pH-pKa_1)}) \quad (44)$$

et :

$$k_b^{obs} = 2k_b \times 10^{(pH-pKap)} \quad (45)$$

En écrivant les équations 44 et 45 nous avons tenu compte du fait que, pour chaque pont réduit, il y a deux positions de déprotonation possibles.

La substitution des équations 34, 35 et 36 conduit à :

$$\frac{d[DTT_{SS}]}{dt} = \frac{k_f^{obs}}{1 + 2f1 + f1f2} \left([DTT_{SH,SH}]_I - [DTT_{SS}] \right) \left([(Prot_{SS})_N]_I - [DTT_{SS}] \right) - \frac{\alpha}{K_{eq}^{obs}} [DTT_{SS}]^2 \quad (46)$$

où : $f1 = 10^{pH-pKa_1}$ et $f2 = 10^{pH-pKa_2}$.

L'intégration de l'équation 46 a pour résultat :

$$d[DTT_{SS}]dt = \frac{X_1 X_2 \exp[k't(X_2 - X_1)] - 1}{X_2 \exp[k't(X_2 - X_1)] - X_1} \quad (47)$$

Ici :

$$k' = \frac{k_f^{obs}}{1 + 2f1 + f1f2} \left(1 - \frac{\alpha}{K_{eq}^{obs}}\right) \quad (48)$$

X_1 et X_2 sont les solutions de l'équation suivante :

$$\left([DTT_{SH,SH}]_I\right) \left([(Prot_{SS})_N]_I - x\right) - \frac{\alpha}{K_{eq}^{obs}} x^2 = 0 \quad (49)$$

L'équation 47 représente la fonction modèle utilisée pour la modélisation des données expérimentales concernant la cinétique de la réaction de réduction des ponts S-S par le DTT.

La concentration de ponts S-S mixtes $[(Prot_{S^-} - SS - DTT_{SH})_U]$ qui provient de $Prot_{SS}$ est négligeable par rapport à la concentration des ponts réduits $[Prot_{SH,SH}]$. Cela est une conséquence du fait que dans l'état déplié U , la distance entre le pont S-S mixtes et un autre thiolate RS^- de la protéine est plus grande que celle par rapport au thiolate du DTT. Cette hypothèse conduit à :

$$k_f^0 = k_1 \quad (50)$$

Conformément aux équations 23, 24 et 50, l'équation de vitesse pour la réduction des ponts S-S exposés au solvant est la suivante :

$$\frac{d[(Prot_{SS})_u]}{dt} = k_u[(Prot_{SS})_N] - k_n[(Prot_{SS})_u] - k_f^0 \left([DTT_{SH,S^-}] + [DTT_{S^-,SH}] \right) \left([(Prot_{SS})_u] + k_{-1}[(Prot_{S^-} - SS - DTT_{SH})_U] \right) \quad (51)$$

En se plaçant dans des conditions expérimentales où le flux inverse, $k_1[(Prot_{S^-} - SS - DTT_{SH})_U]$, est négligeable et la concentration de l'intermédiaire $(Prot_{SS})_u$ quasi-stationnaire ($\frac{d[(Prot_{SS})_u]}{dt} = 0$), on peut écrire la relation suivante :

$$[(Prot_{SS})_u] = \frac{k_u[(Prot_{SS})_N]}{k_n + k_f^0 \left([DTT_{SH,S^-}] + [DTT_{S^-,SH}] \right)} \quad (52)$$

En comparant les équations 41, 42 et 52 on obtient :

$$k_f = \frac{k_u}{k_n + k_f^0 \left([DTT_{SH,S^-}] + [DTT_{S^-,SH}] \right)} k_f^0 \quad (53)$$

Puisque $k_n \gg k_f^0 [DTT_{SH,S^-}]$ (le processus de repliement est un processus très rapide), l'équation 53 devient :

$$k_f = \frac{k_u}{k_n} k_f^0 \quad (54)$$

Cette approximation a été aussi proposée dans la littérature [50, 112].

Bibliographie

- [1] Z. Feng, M.C. Butler, S.L. Alam, and S.N. Loh. *J. Mol. Biol.*, 314 :153–166, 2001.
 - [2] J.M. Wilson, R.J. Bayer, and D.J. Hupe. *J. Am. Chem. Soc.*, 99(24) :7922–7926, 1977.
 - [3] D.M. Rothwarf and H.A. Scheraga. *Proc. Natl. Acad. Sci. USA*, 89 :7944–7948, 1992.
 - [4] W.J. Lees and G.M. Whitesides. *J. Org. Chem.*, 58 :642–647, 1993.
 - [5] B. Wilkinson, R. Xiao, and H.F. Gilbert. *J. Biol. Chem.*, 280(12) :11483–11487, 2005.
 - [6] F. Vinci, M. Ruoppolo, P. Pucci, R.B. Freedman, and G. Marino. *Protein Sci.*, 9 : 525–535, 2000.
 - [7] D.P. Goldenberg, L.S. Bekeart, D.A. Laheru, and J.D. Zhou. *Biochem.*, 32 :2835–2844, 1993.
 - [8] M. Karlsson, L.G. Martensson, C. Karlsson, and U. Carlsson. *Biochem.*, 44 :3487–3493, 2005.
 - [9] W.J. Wedemeyer, E. Welker, M. Narayan, and H.A. Scheraga. *Biochem.*, 39(15) : 4207–4216, 2000.
 - [10] D.C. Vaz, J.R. Rodrigues, W. Sebald, C.M. Dobson, and R.M.M. Brito. *Protein Sci.*, 15 :33–44, 2006.
 - [11] T.E. Creighton. *Biol. Chem.*, 378 :731–744, 1997.
 - [12] M. Langenhof, S.S.J. Leong, L.K. Pattenden, and A.P.J. Middelberg. *J. Chromatographie A*, 1069 :195–201, 2005.
-

-
- [13] T.V. DeCollo and W.J. Lees. *J. Org. Chem.*, 66 :4244–4249, 2001.
- [14] E. Li-Chan and S. Nakai. *J. Agric. Food Chem.*, 39 :1238–1245, 1991.
- [15] R.C.K. Reddy, H. Lilie, R. Rudolph, and C. Lange. *Protein Sci.*, 14 :929–935, 2005.
- [16] S.J. Burton, A.V. Quirk, and P.C. Wood. *Eur. J. Biochem.*, 179 :379–387, 1989.
- [17] D.L. Rabenstein and K.H. Weaver. *J. Org. Chem.*, 61 :7391–7397, 1996.
- [18] G.H. Snyder, M.J. Cennerazzo, A.J. Karalis, and D. Field. *Biochem.*, 20(23) :6509–6518, 1981.
- [19] K. Kuwajima, M. Ikeguchi, T. Sugawara, Y. Hiraoka, and S. Sugai. *Biochem.*, 29 :8240–8249, 1990.
- [20] R.K. Owusu-Apenten, C. Chee, and O.P. Hwee. *Food Chem.*, 83 :541–545, 2003.
- [21] P.G. Vilaro, A. Scaloni, P. Amodeo, C. Barsotti, I. Cecconi, M. Cappiello, B. Lopez Mendez, R. Rullo, M. Dal Monte, A. Del Corso, and U. Mura. *Biochem.*, 40 :11985–11994, 2001.
- [22] A. Sawicka, P. Skurski, R.R. Hudgins, and J. Simons. *J. Phys. Chem. B*, 107 :13505–13511, 2003.
- [23] O. Dmitrenko, C. Thorpe, and R.D. Bach. *J. Org. Chem.*, 72 :8298–8307, 2008.
- [24] S. Antonello, K. Daasbjerg, H. Jensen, F. Taddei, and F. Maran. *J. Am. Chem. Soc.*, 125 :14905–14916, 2003.
- [25] J. Berges, E. Kassab, D. Conte, E. Adjadj, and C. Houe-Levin. *J. Phys. Chem. A*, 101 :7809–7817, 1997.
- [26] J. Berges, G. Rickards, A. Rauk, and C. Houee-Levin. *Chem. Phys. Lett.*, 421 :63–67, 2006.
- [27] S.U. Sane, S.M. Cramer, and T.M. Przybycien. *Anal. Biochem.*, 269 :255–272, 1999.
- [28] Z. Chi, X.G. Chen, J.S.W. Holtz, and S.A. Asher. *Biochem.*, 37 :2854–2864, 1998.
- [29] N.C. Maiti, M.M. Apetri, M.G. Zagorski, P.R. Carey, and V.E. Anderson. *J. Am. Chem. Soc.*, 126(8) :2399–2408, 2004.
- [30] H.E. Van Wart and H.A. Scheraga. *Proc. Natl. Acad. Sci.*, 74 :13–17, 1977.
-

-
- [31] H. Sugeta, A. Go, and T. Miyazawa. *B. Chem. Soc. Jap.*, 46 :3407–3411, 1973.
- [32] N. Nogami, H. Sugeta, and T. Miyazawa. *B. Chem. Soc. Jap.*, 48 :2417–2420, 1975.
- [33] N. Nogami, H. Sugeta, and T. Miyazawa. *Chem. Lett.*, pages 147–150, 1975.
- [34] B.A. Katz and A. Kossiakoff. *J. Biol. Chem.*, 261 :15480–15485, 1986.
- [35] H.E. van Wart and H.A. Scheraga. *J. Phys. Chem.*, 80(16) :1812–1822, 1976.
- [36] H.E. van Wart and H.A. Scheraga. *J. Phys. Chem.*, 80(16) :1823–1832, 1976.
- [37] H. Sugeta, A. Go, and T. Miyazawa. *Chem. Lett.*, pages 83–86, 1972.
- [38] W. Qian and S. Krimm. *Biopolymers*, 32 :321–326, 1992.
- [39] A. Moran and S. Mukamel. *Proc. Natl. Acad. Sci.*, 101(2) :506–510, 2004.
- [40] J.H. Choi, S. Ham, and M. Cho. *J. Phys. Chem. B*, 103 :9132–9138, 2003.
- [41] S. Krimm and Y. Abe. *Proc. Nat. Acad. Sci. USA*, 69(10) :2788–2792, 1972.
- [42] R. Schweitzer-Stenner. *Vibr. Spec.*, 42(1) :98–117, 2006.
- [43] G. Rauhut and P. Pulay. *J. Phys. Chem.*, 99(10) :3093–3100, 1995.
- [44] M.D. Halls and H.B. Schlegel. *J. Chem. Phys.*, 111(19) :8819–8824, 1999.
- [45] G.K. Schenter and B.C. Garrett. *J. Chem. Phys.*, 113(13) :5171–5178, 2000.
- [46] S.S. Iyengar. *J. Chem. Phys.*, 123(8) :084309–084309(9), 2005.
- [47] Z. Zhu and M.E. Tuckerman. *J. Phys. Chem. B*, 106 :8009–8018, 2002.
- [48] M.A. Mroginiski, F. Mark, W. Thiel, and P. Hildebrandt. *Biophys. J.*, 93 :1885–1894, 2007.
- [49] H. Zhang, M. Le, and G.E. Means. *Bioorg. Chem.*, 26 :356–364, 1998.
- [50] Y-J. Li, D.M. Rothwarf, and H.A. Scheraga. *Nat. Struct. Biol.*, 2 :489–494, 1995.
- [51] H. Lin and D.G. Truhlar. *Theor. Chem. Acc.*, 117 :185–199, 2007.
- [52] M.R. Reddy and M.D. Erion. *J. Am. Chem. Soc.*, 129(30) :9296–9297, 2007.
- [53] T. Vreven and K. Morokuma. *Theor. Chem. Acc.*, 109 :125–132, 2003.
-

-
- [54] L. Ridder, J.N. Harvey, I.M.C.M. Rietjens, J. Vervoort, and A.J. Mulholland. *J. Phys. Chem. B*, 107(9) :2118–2126, 2003.
- [55] H. Lin, J.C. Schneboom, S. Cohen, S. Shaik, and W. Thiel. *J. Phys. Chem. B*, 108(28) :10083–10088, 2004.
- [56] P.A. Kollman, B. Kuhn, O. Donini, M. Perakyla, R. Stanton, and D. Bakowies. *Acc. Chem. Res.*, 34(1) :72–79, 2001.
- [57] W.P. Jorgensen, J. Chandrasekhar, and J.P. Madura. *J. Chem. Phys.*, 79 :926–935, 1983.
- [58] J.J. Ruiz-Perna, E. Silla, I. Tun, S. Mart, and V. Moliner. *J. Phys. Chem. B*, 108(24) :8427–8433, 2004.
- [59] E.B. Wilson Jr., J.C. Decius, and P.C. Cross. *Molecular vibrations. The theory of Infrared and Raman vibrational Spectra*. Dover Publications, Inc., New York, 1995.
- [60] A.C. Albrecht. *J. Chem. Phys*, 34 :1476–1484, 1961.
- [61] I.R. Lewis and H.G.M Edwards. *Practical spectroscopy(Handbook of Raman Spectroscopy)*. Marcel Dekker, Inc., New York, 2001.
- [62] M. Born and J.R. Oppenheimer. *Ann. Physik*, 874 :457, 1927.
- [63] P. Hohenberg and W. Kohn. *Phys. Rev.*, 136 :B864–B871, 1964.
- [64] W. Kohn and L.J. Sham. *Phys. Rev.A*, 140 :1133–1138, 1965.
- [65] A.D. Becke. *J. Chem. Theor. Comput.*, 98 :5648–5652, 1993.
- [66] C. Lee, W. Yang, and R.G. Parr. *Phys. Rev. B*, 37 :785–789, 1988.
- [67] S.H. Vosko, L. Wilk, and M. Nusair. *Can. J. Phys.*, 58 :1200–1211, 1980.
- [68] P.J. Stephens, F.J. Devlin, C.F. Chabalowski, and M.J. Frisch. *J. Phys. Chem.*, 98 : 11623–11627, 1994.
- [69] C.C.J. Roothaan. *Ev. Mod. Phys.*, 23 :69–89, 1951.
- [70] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheese-
man, J. A. Montgomery Jr, T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S.
Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A.
-

- Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople. Gaussian 03, Revision C.02. Gaussian, Inc., Wallingford, CT, 2004.
- [71] M. Cossi, G. Scalmani, N. Rega, and V. Barone. *J. Chem. Phys.*, 117 :43–54, 2002.
- [72] S. Glasstone, K.J. Laidler, and H. Eyring. *Theory of rate processes*. MacGraw-Hill, New York, 1941.
- [73] J.P. Ryckaert, G. Ciccoti, and H.J.C. Berendsen. *J. Comput. Phys.*, 23 :327–341, 1977.
- [74] J.P. Ryckaert. *Mol. Phys.*, 55 :549–556, 1985.
- [75] D.A. Case, T.A. Darden, T.E. Cheatham III, C.L. Simmerling, J. Wang, R.E. Duke, R.Luo, K.M. Merz, D.A. Pearlman, M. Crowley, R.C. Walker, W. Zhang, B. Wang, S.Hayik, A. Roitberg, G. Seabra, K.F. Wong, F. Paesani, X. Wu, S. Brozell, V. Tsui, H.Gohlke, L. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, P. Beroza, D.H. Mathews, C.Schafmeister and W.S. Ross, , and P.A. Kollman. Amber 9,. 2006, University of California, San Francisco.
- [76] P.A. Kollman, R. Dixon, W. Cornell, T. Fox, C. Chipot, and A. Pohorille. *Computer simulation of biomolecular systems, Vol. 3*. Elsevier, New York, 1997.
- [77] U.C. Singh and P.A. Kollman. *J. Comp. Chem.*, 5 :129–145, 1984.
- [78] B. Roux. *Comput. Phys. Commun.*, 91 :275–282, 1995.
- [79] G.M. Torrie and J.P. Valeau. *J. Comput. Phys.*, 23(2) :187–199, 1977.
- [80] C. Bartels and M. Karplus. *J. Comput. Chem.*, 18 :1450–1462, 1997.
- [81] S. Kumar, D. Bouzida, R. Swendsen, P. Kollman, and J. Rosenberg. *J. Comput. Chem.*, 13 :1011–1021, 1992.
-

-
- [82] A. Warshel and M. Levitt. *J. Mol. Biol.*, 103 :227–249, 1976.
- [83] U.C. Singh and P.A. Kollman. *J. Comp. Chem.*, 7(6) :718–730, 1986.
- [84] M.J. Field, P.A. Bash, and M. Karplus. *J. Comp. Chem.*, 11(6) :700–733, 1990.
- [85] T. Vreven, K.S. Byun, I. Komromi, S. Dapprich, J.A. Montgomery, K. Morokuma, and M.J. Frisch. *J. Chem. Theor. Comput.*, 2(3) :815–826, 2006.
- [86] F. Maseras and K. Morokuma. *J. Comp. Chem.*, 16 :3654–3661, 1995.
- [87] M. Svensson, S. Humbel, R.D.J. Froese, T. Matsubara, S. Sieber, and K. Morokuma. *J. Phys. Chem.*, 100 :19357–19363, 1996.
- [88] T. Vreven, B. Mennucci, C.O. da Silva, K. Morokuma, and J. Tomasi. *J. Phys. Chem.*, 115 :62–72, 2001.
- [89] S. Humbel, S. Sieber, and K. Morokuma. *J. Chem. Phys.*, 105 :1959–1967, 1996.
- [90] M. Svensson, S. Humbel, and K. Morokuma. *J. Chem. Phys.*, 105 :3654–3661, 1996.
- [91] M.C. Chen and R.C. Lord. *J. Am. Chem. Soc.*, 18 :990–992, 1976.
- [92] K. Nakamura, S. Era, Y. Ozaki, M. Sogami, T. Hayashi, and M. Murakami. *FEBS Lett.*, 417 :375–378, 1997.
- [93] H.E. Van Wart, A. Lewis, H.A. Scheraga, and F.D. Saeva. *P.N.A.S.*, 70 :2619–2623, 1973.
- [94] E.J. East, R.C.C. Chang, N.T. Yu, and J.F.R. Kuck Jr. *J. Biol. Chem.*, 253 :1436–1441, 1978.
- [95] S.W. Ellepola, S.M. Choi, D.L. Phillips, and C.Y. Ma. *J. Cereals Sci.*, 43 :85–93, 2006.
- [96] K. Yamamoto, Y. Mizutani, and T. Kitawaga. *Biophys. J.*, 79 :485–495, 2000.
- [97] W. Chen, S. Nie, J.F.R. Kuck Jr., and N.T. Yu. *Biophys. J.*, 60 :447–455, 1991.
- [98] C. David, S. Foley, C. Mavon, and M. Enescu. *Biopolymers*, 89 :623–634, 2008.
- [99] C. David, S. Foley, and M. Enescu. *Phys. Chem. Chem. Phys.*, 11 :1–11, 2009.
- [100] H.E. Van Wart and H.A. Scheraga. *Proc. Natl. Acad. Sci.*, 83 :3064–3067, 1986.
-

-
- [101] T. Kitagawa, T. Azuma, and K. Hamaguchi. *Biopolymers*, 18 :451–465, 1978.
- [102] S. Sugio, A. Kashima, S. Mochizuki, M. Noda, and K. Kobayashi. *Prot. Eng.*, 12 : 439–446, 1999.
- [103] S.C. Wang and C. Ted Lee Jr. *J. Phys. Chem. B*, 110 :16117–16123, 2006.
- [104] J. Jakoncic, M. Aslantas, V. Honkimaki, M. Dimichel, and V. Stojanoff. *J. Appl. Crystallogr.*, 39 :831–841, 2006.
- [105] A.B. Kudryavtsev, S.B. Mirov, L.J. DeLucas, C. Nicolete, M. van der Woerd, T.L. Bray, and T.T. Basiev. *Acta Cryst.*, D54 :1216–1229, 1998.
- [106] Z. Deng, P.L. Polavarapu, S.J. Ford, L. Hecht, L.D. Barron, C.S. Ewing, and K. Jalkanen. *J. Phys. Chem.*, 100 :2025–2034, 1996.
- [107] W.G. Han, K.J. Jalkanen, M. Elstner, and S. Suhai. *J. Phys. Chem. B*, 102 :2587–2602, 1998.
- [108] C.F. Weise and J.C. Weisshaar. *J. Phys. Chem. B*, 107 :3265–3277, 2003.
- [109] D.J. Cline, S.E. Redding, S.G. Brohawn, J.N. Psathas, J.P. Schneider, and C. Thorpe. *Biochem.*, 43 :15195–15203, 2004.
- [110] W.K. Surewicz, M.A. Moscarello, and H.H. Mantsch. *J. Biol. Chem.*, 262 :8598–8602, 1987.
- [111] B.M. Smith and S. Franzen. *S. Anal. Chem.*, 74 :4076–4080, 2002.
- [112] J. Fontecha, J. Bellanato, and M. Juarez. *J. Dairy Sci.*, 76 :3303–3309, 1993.
- [113] C.T. Reimann, I. Velazquez, and O. Tapia. *J. Phys. Chem. B*, 102 :2277–2283, 1998.
- [114] A.P. Hinck, D.M. Truckses, and J.L. Markley. *Biochem.*, 35 :10328–10338, 1996.
- [115] J. Zheng and H.F. Gilbert. *J. Biol. Chem.*, 276(19) :15747–15752, 2001.
- [116] G. Xu, M. Narayan, I. Kurinov, D.R. Ripoll, E. Welker, M. Khalili, S.E. Ealick, and H.A. Scheraga. *J. Am. Chem. Soc.*, 128 :1204–1213, 2006.
- [117] L.-C. Ma and S. Anderson. *Biochem.*, 36 :3728–3736, 1997.
- [118] R.E. Hansen, H. Østergaard, and J.R. Winther. *Biochem.*, 44 :5899–5906, 2005.
- [119] A. Holmgren. *Annu. Rev. Biochem.*, 54 :237–271, 1985.
-

- [120] M. Narayan, G. Xu, D. Ripoll, H. Zhai, K. Breuker, C. Wanjalla, H.J. Leung, A. Navon, E. Welker, F.W. McLafferty, and H.A. Scheraga. *J. Mol. Biol.*, 338 :795–809, 2004.
- [121] H. Gunaydin, O. Acevedo, W.L. Jorgensen, and K.N. Houk. *J. Chem. Theor. Comput.*, 3 :1028–1035, 2007.
- [122] G. Li, X. Zhang, and Q. Cui. *J. Phys. Chem. B*, 107 :8643–8653, 2003.
- [123] P.L. Cummins, I.V. Rostov, and J.E. Gready. *J. Chem. Theor. Comput.*, 3 :1203–1211, 2007.
- [124] J. Kastner, H.M. Senn, S. Thiel, N. Otte, and W. Thiel. *J. Chem. Phys.*, 2 :452–461, 2006.
- [125] E. Rosta, M. Klhn, and A. Warshel. *J. Phys. Chem. B*, 110(6) :2934–2941, 2006.
- [126] J. Chandrasekhar, S.F. Smith, and W.L. Jorgensen. *J. Am. Chem. Soc.*, 107 :154–162, 1985.
- [127] J. Chandrasekhar and W.L. Jorgensen. *J. Am. Chem. Soc.*, 107 :2974–2975, 1985.
- [128] Y. Zhang, H. Liu, and W. Yang. *J. Chem. Phys.*, 112(8) :3483–3492, 2000.
- [129] A.E. Reed, R.B. Weinstock, and F. Weinhold. *J. Chem. Phys.*, 83(15) :735–746, 1985.
-