

THESIS / THÈSE

MASTER EN SCIENCES CHIMIQUE À FINALITÉ APPROFONDIE

Étude mécanistique de la phosphosérine phosphatase humaine, une enzyme impliquée dans la chimiorésistance au 5-Fluorouracile dans le cancer du côlon approches expérimentale et computationnelle

Mirgaux, Manon

Award date:
2019

Awarding institution:
Universite de Namur

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



Université de Namur
Faculté des Sciences

ÉTUDE MÉCANISTIQUE DE LA PHOSPHOSÉRINE PHOSPHATASE HUMAINE, UNE ENZYME IMPLIQUÉE DANS LA CHIMIORÉSISTANCE AU 5-FLUOROURACILE DANS LE CANCER DU CÔLON : APPROCHES EXPÉRIMENTALE ET COMPUTATIONNELLE

Mémoire présenté pour l'obtention

du grade académique de Master Chimie «Chimie du Vivant et des Nanomatériaux»: Finalité Approfondie

Manon MIRGAUX

Janvier 2019

UNIVERSITE DE NAMUR
Faculté des Sciences
Secrétariat du Département de Chimie
Rue de Bruxelles 61 – 5000 NAMUR
Téléphone : +32(0)81 72.54.44 – Téléfax : +32(0)81 72.54.40
E-mail : enseignement.chimie@unamur.be - www.unamur.be/sciences

Étude mécanistique de la phosphosérine phosphatase humaine, une enzyme impliquée dans la chimiorésistance au 5-Fluorouracile dans le cancer du côlon : approches expérimentale et computationnelle

MIRGAUX Manon

Résumé

À l'heure actuelle, le cancer est une des causes majeures de décès au monde. En Belgique, parmi les cancers les plus courants, le cancer du côlon est à la deuxième position chez la femme et à la troisième chez l'homme. Le traitement des patients touchés par ce cancer résulte en une combinaison de chirurgie et de chimiothérapie à l'aide d'un médicament nommé le 5-Fluorouracile (5-FU), médicament causant un stress oxydatif aux cellules cancéreuses. Toutefois, de plus en plus de cas de résistances aux traitements par 5-FU sont enregistrés. Cette résistance, dont les causes sont multiples, peut être due à la surexpression d'une enzyme dans les cellules cancéreuses, la phosphosérine phosphatase humaine (HPSP).

HPSP est une enzyme magnésio-dépendante impliquée dans la dernière étape de la voie de la sérine. La surexpression de HPSP entraîne une surproduction de L-sérine, pouvant être convertie en glutathion, un antioxydant régulant le stress oxydatif et diminuant l'effet du 5-FU. L'inhibition de HPSP pourrait aider à lutter contre la chimiorésistance dans le cas des cancers colorectaux. Cependant, HPSP est une cible thérapeutique peu étudiée dont le mécanisme réactionnel reste relativement méconnu. La structure de l'enzyme n'a d'ailleurs jamais été résolue en présence de magnésium. Une meilleure compréhension de l'enzyme pourrait aider à concevoir de nouveaux médicaments contre la chimiorésistance.

Dans le but d'élucider le mécanisme de HPSP, la protéine a été cristallisée avec des ligands afin de déterminer la structure de ceux-ci au sein du site actif et de mettre en lumière les interactions clés entre les ligands et les acides aminés de ce site actif. Une hypothèse mécanistique a dès lors pu être élaborée. Cette hypothèse a ensuite pu être appuyée par une analyse computationnelle des paramètres thermodynamiques et cinétiques du système par simulation du site actif dans des modèles simplifiés, de complexités croissantes. Un mécanisme final a ainsi pu être proposé.

Mémoire de master en Sciences Chimiques (Orientation Chimie du Vivant et des Nanomatériaux)
Finalité Approfondie

Janvier 2019

Promoteurs : Prof. J. Wouters et Prof. B. Champagne

Encadrants : M. Haufroid et P. Beaujean

*" C'est de la force des convictions que dépend la réussite,
pas du nombre de partisans."*

**Rémus Lupin,
Harry Potter et les reliques de la mort, partie deux**

Lors de la réalisation de ce mémoire, beaucoup de personnes m'ont soutenues à leur manière. Pour commencer, je souhaiterais remercier mes deux promoteurs, Professeur J. Wouters et Professeur B. Champagne. Merci pour l'écoute, la patience et le suivi apporté durant cette année, les longues réunions (mais toujours instructives) dans un but de rendre ma recherche scientifique toujours plus précise et réaliste.

*En pensant à mes deux promoteurs me vient directement à l'esprit mes deux encadrants, Marie Haufroid et Pierre Beaujean. Marie, merci pour cette bonne humeur contagieuse qui te caractérise, cette volonté d'être toujours attentive à la moindre question de tes mémorants, cette patience quand il s'agit de **R**elire mainte et mainte fois ce mémoire. Merci pour tes **C**onseils, toujours pertinents, mais surtout, je te remercie pour tout ce qu'il y a "à côté » du laboratoire. Tu resteras **I**ndiscutablement l'encadrante à qui ont peut parler de tout (ou presque) autour d'un verre, d'un cupcake ou tout ce qui peut être assimilé par un estomac. À l'heure où j'écris ces remerciements, j'ai devant les deux trois post'its de toi : sur deux d'entre eux, il y a les mots glace et brownies. Je n'oublierais pas ces escapades au cercle VT ou en terrasse, pour se détendre après une journée difficile.*

Vient alors mon deuxième encadrant, mais pas des moindres : Pierre. Comment décrire cette année sous ton encadrement ? Amis tout d'abord, collègue du KapForScience, j'ai appris à travailler avec toi. Tu m'as apporté énormément en bagage scientifique, n'hésitant jamais à réexpliquer les bases pour comprendre la théorie derrière mes manipulations. Qu'importe le jour (où l'heure parfois) de la semaine, toujours là pour répondre à mes angoisses les plus profondes (Un "Oh Pieeeeeerre, il faut tout recommencer, je crois" met toujours de bonne humeur à 9h du matin le dimanche). Merci d'avoir relu, sans peine, mes parties expérimentales, toujours prêt pour donner un conseil ou deux. Un jour, on inventera un pc avec deux souris séparées, à ta plus grande joie. Rien que nos prises de têtes et petites embrouilles (parce qu'on aime tous les deux avoir raison) manqueront sûrement à mes journées.

Je tiens à remercier le CBS : Laurie, Mégane, Thomas, Sebastien, Quentin, Elise, Andréas et Kalina, toujours là pour aider. Et particulièrement, Laurie et Andreas pour leur aide lors de mon analyse de la CSD ainsi que Seb pour ses recommandations lors de la lecture de mon introduction et son intérêt pour mon mémoire. Merci aussi à mon comparse, Nef, pour cette année à partager la paillasse d'"à côté". Toujours au rendez-vous pour papoter, c'est avec plaisir que j'ai appris à te connaître. Tout cela m'amène à Fanny Zola, stagiaire particulière du CBS devenue amie. Etonnement (ou pas), je suis sur que les cupcakes et longues discussions accompagnantes perdureront longtemps entre nous deux.

Après cela, merci aux Nombreux intervenants externes lors de mon mémoire : L'URBM pour la mise à disposition du matériel, Étienne De Plaen pour les souches de HPSP, le CECI et Frédérique Wautelet pour les ressources computationnelles, Laurent Demelene pour l'aide informatique ainsi que Loic Jeanmart pour les conseils organiques. Je N'oublie pas Vincent Liegois, toujours prêt pour la mise à jour de Drawmol (DrawProt bientôt). Je remercie le synchrotron SOLEIL et ces employés pour leur accueil, sans oublier la sauce des boulettes Et les McFleury à 23h sur le retour de Paris.

Enfin, je remercie mes amis, spécialement Osi, Tom, Damien, Loraine, Mathilde, Martin et Charlotte. Merci pour ces cinq années, tous ces souvenirs marqués à jamais. Merci à Charlotte d'avoir été le bureau du paysager où aller pour papoter, rigoler ou se faire consoler, au choix. Merci d'avoir été l'oreille à qui parler pendant tout ce temps. Je remercie également mes parents pour leur soutien. Merci de m'avoir laissé faire les études dont je rêvais et de n'avoir jamais douté de moi. Je vous remercie aussi pour la relecture, la correction de l'orthographe et le soutien lors de l'écriture. Je remercie également ma sœur et mon petit frère pour l'accueil du week-end depuis 5 ans et tout ce qu'ils ont pu m'apporter depuis des années. Pour conclure, je remercie Octave pour sa présence, toujours à mes côtés, depuis deux ans, jouant le rôle à merveille de conseiller, de piller, d'ami, d'amoureux et bien encore au gré des jours.

Liste des abbréviations

Liste des acides aminés

Tableau 1 – Liste et classe des acides aminés, codes à une et trois lettres.

| Nom | Classe | Fonction terminale | Code 3 lettres | Code 1 lettre |
|---------------|---------------------|-------------------------------------|----------------|---------------|
| Alanine | Hydrophobe | $-\text{CH}_3$ | Ala | A |
| Cystéine | Hydrophobe | $-\text{SH}$ | Cys | C |
| Glycine | Hydrophobe | $-\text{H}$ | Gly | G |
| Isoleucine | Hydrophobe | $-\text{CH}_3$ | Ile | I |
| Leucine | Hydrophobe | $-\text{CH}_3$ | Leu | L |
| Méthionine | Hydrophobe | $-\text{SCH}_3$ | Met | M |
| Proline | Hydrophobe | $-\text{C}_4\text{H}_8\text{N}$ | Pro | P |
| Valine | Hydrophobe | $-\text{CH}_3$ | Val | V |
| Asparagine | Polaire | $-\text{CONH}_2$ | Asn | N |
| Glutamine | Polaire | $-\text{CONH}_2$ | Gln | Q |
| Sérine | Polaire | $-\text{OH}$ | Ser | S |
| Thréonine | Polaire | $-\text{OH}$ | Thr | T |
| Aspartate | Chargé négativement | $-\text{COO}^-$ | Asp | D |
| Glutamate | Chargé négativement | $-\text{COO}^-$ | Glu | E |
| Arginine | Chargé positivement | $-\text{NHC}(\text{NH}_2)_2^+$ | Arg | R |
| Histidine | Chargé positivement | $-\text{C}_3\text{H}_4\text{N}_2^+$ | His | H |
| Lysine | Chargé positivement | $-\text{NH}_3^+$ | Lys | K |
| Phénylalanine | Aromatique | $-\text{C}_6\text{H}_6$ | Phe | F |
| Tryptophane | Aromatique | $-\text{C}_8\text{H}_7\text{N}$ | Trp | W |
| Tyrosine | Aromatique | $-\text{C}_6\text{H}_5\text{OH}$ | Tyr | Y |

Acronymes

3PG 3-Phosphoglycerate

3PHP 3-Phosphohydroxypyruvate

5-FU 5-Fluorouracile

ADN Acide désoxyribonucléique

α -KG α -ketoglutarate

AP3 Acide L-(+)-2-Amino-3-phosphonopropionic

APS Persulfate d'ammonium

ARN Acide ribonucléique

CSD *Cambridge Structural Data Bank*

DFT *Density functional theory*

DMSO Dimethyl sulfoxide

DTT Dithiothreitol

E. coli *Escherichia coli*

EDTA Acide éthylènediaminetetraacetic

FdUMP Fluorodeoxyuridine monophosphate

FdUTP Fluorodeoxyuridine triphosphate

FPLC *Fast Purification Liquid Chromatography*

FUTP Fluorouridine triphosphate

GGA *Generalized Gradient Approximation*

HAD Haloacid dehalogenase

HEPES Acide 4-(2-Hydroxyethyl)-piperazine-1-ethanesulfonic

HOM Acide homocystéique

HPSP Phosphosérine phosphatase humaine

IEF-PCM *Integral equation formalism -polarizable continuum model*

IMAC *Immobilized Metal-Affinity Chromatography*

IPTG Isopropyl- β -D-galactopyranoside

IRC *Intrinsic reaction coordinates*

LB Lysogeny Broth

LCAO *Linear Combination of Atomic Orbital*

LDA *Local Density Approximation*

mGGA *meta-Generalized Gradient Approximation*

MJPSP Methanococcus jannashii phosphoserine phosphatase

NAD Nicotinamide adenine dinucléotide (oxidé)

NADH Nicotinamide adenine dinucléotide (réduit)

NADPH Nicotinamide adenine dinucléotide phosphate hydrogen

NTA Acide nitrilotriacétique

PDB *Protein Data Bank*

PEG Polyéthylène glycol

PHGDH Phosphoglycérate déhydrogénase

Pi Phosphate inorganique

PLS Phospho-L-sérine (3-Phosphosérine)

PRDX2 Peroxiredoxin-2

PSAT1 Phosphosérine aminotransférase

RMSD *Root-mean-square-deviation*

ROS *Reactive oxygen species*

RX Rayon X

SAS *Solvatation accesible surface*

SCF *Self Consistent Field*

SDS-PAGE *Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis*

SES *Solvent Excluded Surface*

SHMT Sérine hydroxyméthyltransferase

SN2 Substitution nucléophile d'ordre 2

TEMED Tétraméthylethylenediamine

Tris-HCl Tris(hydroxyméthyl)-aminométhane-hydrochloride

TS Thymidylate synthase

UFF *Universal Force Field*

UV Ultra violet

Table des matières

| | | |
|-----------|--|-----------|
| I | Introduction | 1 |
| 1 | Le cancer | 2 |
| 1.1 | Description de la maladie | 3 |
| 1.2 | Le cancer du côlon | 4 |
| 1.3 | Traitements | 5 |
| | La chirurgie | 5 |
| | La chimiothérapie | 5 |
| | La radiothérapie | 5 |
| | Thérapies nouvelles | 5 |
| 1.4 | Traitement du cancer du côlon par le 5-Fluorouracile | 6 |
| 2 | HPSP, une cible thérapeutique peu décrite | 10 |
| 2.1 | Identification de HPSP comme cible thérapeutique | 10 |
| 2.1.1 | Voie de la sérine | 10 |
| 2.1.2 | Implication de HPSP dans la chimiorésistance | 12 |
| 2.2 | Caractéristiques connues de HPSP | 13 |
| 2.2.1 | Connaissances mécanistiques et structurales sur les HADs | 13 |
| 2.2.2 | Connaissances mécanistiques et structurales sur HPSP | 16 |
| 3 | Objectifs et stratégies | 20 |
| II | Approche cristallographique | 22 |
| 4 | Expression, production et purification | 23 |

| | | |
|----------|---|-----------|
| 4.1 | Expression et production de HPSP | 23 |
| 4.1.1 | Culture bactérienne et induction | 23 |
| 4.1.2 | Lyse par sonification | 25 |
| 4.2 | Purification | 25 |
| 4.2.1 | Principe de la purification par IMAC | 26 |
| 4.2.2 | Résultats de la purification par IMAC | 27 |
| 4.3 | Suivi de purification | 29 |
| 4.3.1 | Suivi de purification par SDS-PAGE | 29 |
| 4.3.2 | Suivi de purification par test d'activité | 30 |
| 4.4 | Conclusions intermédiaires | 31 |
| 5 | Cristallisation de HPSP | 32 |
| 5.1 | Principe de la cristallisation | 32 |
| 5.2 | Cristallisation en présence de phospho-L-sérine | 33 |
| 5.3 | Cristallisation en présence d'un analogue de substrat | 35 |
| 5.4 | Conclusions intermédiaires | 37 |
| 6 | Étude structurale de HPSP | 38 |
| 6.1 | Principe de l'affinement de structure protéique | 38 |
| | Le <i>R-factor</i> | 39 |
| | Les RMSD | 39 |
| | Les cartes de Ramachandran et de rotamères | 39 |
| 6.2 | Structure de HPSP en présence de phospho-L-sérine | 40 |
| 6.3 | Structure de HPSP en présence de L-sérine et de phosphate | 48 |
| 6.4 | Structure de HPSP en présence d'un analogue de substrat | 54 |
| 6.5 | Étude structurale de la boucle dynamique de HPSP | 58 |
| 6.6 | Hypothèse mécanistique | 59 |
| 6.7 | Conclusions intermédiaires | 63 |

| | | |
|------------|--|------------|
| III | Approche computationnelle | 65 |
| 7 | Étude computationnelle du mécanisme | 66 |
| 7.1 | Réaction d'attaque nucléophile par l'Asp20 | 66 |
| 7.1.1 | Aspects structuraux | 67 |
| 7.1.2 | Aspects thermodynamiques et cinétiques | 71 |
| 7.2 | Protonation de la L-sérine par l'Asp22 | 76 |
| 7.3 | Hydrolyse de l'Asp20 | 78 |
| 7.3.1 | Activation à partir de la Lys158 | 79 |
| 7.3.1.1 | Aspects structuraux | 79 |
| 7.3.1.2 | Aspects thermodynamiques et cinétiques | 82 |
| 7.3.2 | Activation à partir de l'Asp183 | 82 |
| 7.3.2.1 | Aspect structuraux | 82 |
| 7.3.2.2 | Aspects thermodynamiques et cinétiques | 86 |
| 7.3.3 | Élimination du phosphate | 88 |
| 7.3.3.1 | Aspect structuraux | 88 |
| 7.3.3.2 | Aspect thermodynamiques et cinétiques | 90 |
| 7.3.4 | Conclusions de la réaction d'hydrolyse | 92 |
| 7.4 | Système protéique | 93 |
| 7.5 | Conclusions intermédiaires | 95 |
| IV | Conclusions | 98 |
| 8 | Conclusions et perspectives | 99 |
| V | Matériels et méthodes | 103 |
| 9 | Notions théoriques abordées | 104 |
| 9.1 | Méthodes liées à la diffraction de rayon X | 104 |
| 9.2 | Méthodes calculatoires | 106 |
| 9.2.1 | Bases de la chimie quantique | 106 |

| | | |
|-----------|---|------------|
| 9.2.2 | Principe de la DFT | 108 |
| 9.2.3 | Fonctionnelles d'échange-corrélation | 109 |
| 9.2.4 | Optimisation de géométrie | 110 |
| 9.2.5 | Détermination des états de transition et des complexes intermédiaires | 110 |
| 9.2.6 | Thermodynamique statistique | 111 |
| 9.2.7 | La solvatation | 112 |
| 10 | Matériel et méthodes | 113 |
| 10.1 | Production et purification | 113 |
| 10.1.1 | Préparation des solutions | 113 |
| 10.1.2 | Préculture et culture | 114 |
| 10.1.3 | Induction | 114 |
| 10.1.4 | Lyse | 114 |
| 10.1.5 | Purification par IMAC | 115 |
| 10.1.5.1 | Conditionnement de la colonne | 115 |
| 10.1.5.2 | Élution de la protéine | 115 |
| 10.2 | Suivi de purification | 115 |
| 10.2.1 | Électrophorèse SDS-PAGE | 115 |
| 10.2.2 | Test d'activité | 117 |
| 10.2.3 | Dosage enzymatique | 118 |
| 10.3 | Cristallisation en présence de ligands | 118 |
| 10.3.1 | Préparation des solutions pour la cristallisation | 118 |
| 10.3.2 | Préparation de la solution enzymatique : PD10 et aliquots | 119 |
| 10.3.3 | Compositions des puits / co-cristallisation avec la phospho-L-sérine | 120 |
| 10.3.4 | Compositions des puits/ co-cristallisation avec de l'analogue | 120 |
| 10.4 | La diffraction au synchrotron | 121 |
| 10.4.1 | Cryo-protection | 121 |
| 10.4.2 | Collecte des données de diffraction | 121 |
| 10.5 | Analyse computationnelle | 121 |

| | |
|-------------------------|------------|
| VI Bibliographie | 122 |
| VII Annexes | 127 |

Première partie

Introduction

Chapitre 1

Le cancer

Le cancer se définit comme une division anormale, anarchique et incontrôlée de cellules ayant acquis une ou plusieurs mutations.¹⁻³ En 2015, la mortalité causée par cette maladie s'est élevée à 8,8 millions de morts, soit un décès sur six, ce qui en fait une des causes majeures de décès dans le monde. En Belgique particulièrement, on dénombre 67 087 diagnostics de cancer pour cette même année. Les trois cancers les plus fréquents sont le cancer du sein, de la prostate et du côlon (Figure 1.1).⁴ Cette maladie touche souvent les personnes âgées de 50 ans et plus, mais n'épargne pas forcément les jeunes adultes ou encore les enfants.² Différents facteurs à risques peuvent mener au cancer, tels que l'obésité, une alimentation peu variée, un mode de vie sédentaire, le tabagisme, des infections (hépatite B, hépatite C ou encore le papillomavirus humain) ou encore un abus d'alcool.^{2,3}

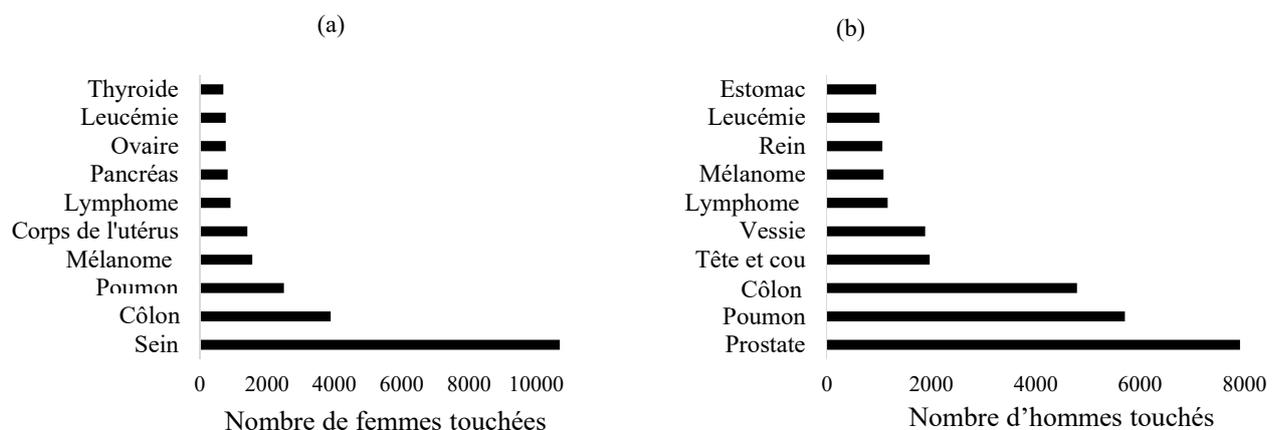


Figure 1.1 – Statistiques des cancers les plus fréquents en 2015, en Belgique, par genre. (a) Chez la femme (b) Chez l'homme.⁴

1.1 Description de la maladie

Il s'agit d'une maladie génétique, dans le sens où elle est causée par des mutations dans le patrimoine génétique des cellules (Figure 1.2, A).^{1,2} Plusieurs facteurs sont en cause, tels qu'une transformation physique (due, par exemple, aux ultra violets ou aux rayons ionisants), chimique (due à l'utilisation de produits chimiques cancérigènes), biologique (due à des agents infectieux) ou tout simplement le vieillissement et l'apparition de mutations ponctuelles de plus en plus fréquentes lors de la division cellulaire.³ Trois types de gènes sont souvent la cible de mutations, à savoir les gènes oncogènes^A, les gènes suppresseurs de tumeurs^B et les gènes ayant un rôle sur la réparation des mutations dans l'ADN.^{1,2} Il n'est pas anodin d'observer des mutations génétiques dans le génome des cellules composant le corps humain lors de la division cellulaire. Cependant, dans le cas d'une cellule normale, une mutation peut soit être corrigée, soit l'apoptose cellulaire est induite si la réparation est impossible.^{1,2} Dans le cas d'une cellule cancéreuse, les mécanismes de réparations ou d'apoptose sont altérés. Les cellules continuent donc leur prolifération en conservant cette mutation (Figure 1.2, B).¹ Par la mutation d'un ou plusieurs de ses gènes, la cellule cancéreuse acquiert des caractéristiques primordiales, telles que la résistance à la mort cellulaire ou encore la capacité de suractiver certaines cascades métaboliques. Cela entraîne une modification du métabolisme standard des cellules leur permettant d'échapper au système immunitaire et au système inflammatoire. Ces caractéristiques garantissent la survie cellulaire et font du cancer une maladie complexe et difficile à traiter.¹

Au fur et à mesure que les cellules se multiplient, elles s'accumulent et forment un amas, nommé tumeur (Figure 1.2, C). Il en existe deux sortes : les tumeurs bénignes et les tumeurs malignes. Les premières grandissent lentement, ne se métastasent pas et sont constituées de cellules fortement similaires aux cellules non cancéreuses. Elles sont relativement peu dangereuses sauf si leur taille et localisation deviennent gênantes pour le patient.^{1,2} À l'inverse, les tumeurs malignes croissent plus vite, s'éparpillent dans les tissus et y causent des dommages. De plus, elles peuvent voyager dans d'autres parties du corps par le biais du sang et former des tumeurs secondaires nommées métastases.^{1,2} Ces métastases sont très dangereuses et mènent à un pronostic vital engagé.

A. Gènes favorisant l'arrivée du cancer en suractivant la division cellulaire.

B. Gènes antioncogène permettant d'empêcher ou de ralentir la division cellulaire.

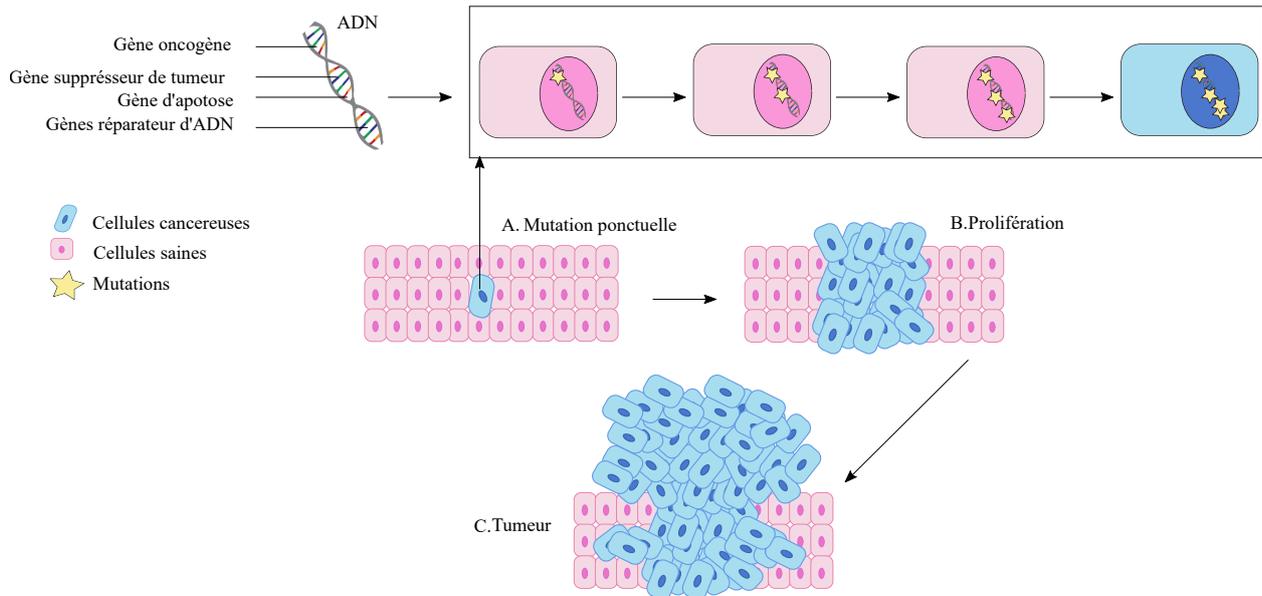


Figure 1.2 – Évolution et prolifération d’une cellule mutée en une tumeur potentiellement cancéreuse. Adapté de Cancer Research (UK) et de Larichesse.^{2,5}

Il existe six types principaux de cancer, à savoir le carcinome (origine : la peau ou les tissus), le sarcome (cancer débutant dans les supports de tissus : os, cartilage, muscles, vaisseaux sanguins), la leucémie (cancer de la moelle osseuse), le lymphome (origine dans le système lymphatique), le myélome (origine : les cellules du plasma et du système immunitaire) ainsi que le cancer du cerveau et du système nerveux central.²

1.2 Le cancer du côlon

Parmi les cancers les plus courants,⁶⁻⁸ le cancer du côlon trône à la troisième position, causant 774 000 décès dans le monde en 2015.³ En Belgique, celui-ci se classe au rang du deuxième cancer le plus fréquent chez la femme et du troisième chez l’homme.^{1,9} Il s’agit d’un carcinome touchant souvent les personnes de plus de 50 ans.^{9,10}

Parmi les différents patients atteints de ce cancer, différents stades sont observés, à savoir le stade zéro ou un dans le cas d’un cancer précoce touchant la couche muqueuse de la paroi intestinale, un stade de type deux si la tumeur traverse la paroi intestinale sans toucher d’autres organes, un stade de type trois lorsque la propagation s’étend au ganglion lymphatique voisin et,

enfin, un stade de type quatre si l'on est face à une forme évoluée de la maladie avec présence de métastases.¹¹ Ces métastases se situent fréquemment dans le foie, les poumons ou les os.¹² À titre d'exemple, 44% des personnes mourant du cancer du côlon ont des métastases au foie.¹² Sans traitement, les patients avec des métastases au foie ont un taux de survie allant de 5 à 12 mois alors que des patients sans traitement et sans métastases peuvent survivre durant une période s'étendant d'une année (77%) à trois années (14% à 23%).¹²

1.3 Traitements

De par le caractère létal du cancer, de nombreuses recherches ont été réalisées et ont permis de mettre au point différents traitements contre cette maladie. Les plus connus sont la chirurgie, la chimiothérapie et la radiothérapie.¹³ Toutefois, d'autres traitements, décrits ci-dessous, existent et se développent au cours de ces dernières années.^{2,9}

La chirurgie consiste à retirer le tissu cancéreux du corps. Son application dépend du type de cancer, de la taille, de la position de celui-ci et de l'état de santé général du patient. Si le cancer est étendu, cette opération n'est pas envisageable. Il en est de même pour le cas des leucémies ou des lymphomes.²

La chimiothérapie est un traitement systémique utilisant des médicaments anticancéreux. Étant donné que les cellules cancéreuses se divisent anormalement par rapport aux cellules saines, le médicament s'attaque principalement à celles-ci. Il est possible de combiner la chimiothérapie à d'autres types de traitements (bithérapie) afin d'augmenter l'efficacité de celle-ci.

La radiothérapie utilise des rayons ionisants pour détruire les cellules cancéreuses en abimant leur ADN, directement à l'intérieur des cellules. Les cellules normales sont aussi affectées, mais elles possèdent des mécanismes de réparation à l'instar des cellules cancéreuses.

Thérapies nouvelles De plus en plus d'autres thérapies sont en train de voir le jour. En guise d'exemple, on peut citer la thérapie aux hormones, ou hormonothérapie. Cette technique, utilisée pour certains cancers du sein et de la prostate,⁹ permet de traiter des cancers hormono-dépendants

avec des thérapies ciblant les récepteurs des hormones. La technique utilise des médicaments empêchant l'arrivée de ces hormones de croissance. On peut aussi citer l'immunothérapie qui stimule les cellules immunitaires pour combattre le cancer en reconnaissant les cellules cancéreuses.²

1.4 Traitement du cancer du côlon par le 5-Fluorouracile

Spécifiquement, dans le cas du cancer du côlon, le traitement résulte en une combinaison de deux approches, à savoir la chirurgie et la chimiothérapie.⁷ La chirurgie est une étape obligatoire soit pour enlever la tumeur ou soit éviter une obstruction intestinale causée par le cancer. La chimiothérapie est ensuite administrée pour éviter la récurrence et les métastases.^{9,13}

Afin d'effectuer la chimiothérapie, un des médicaments, administré par voie intraveineuse, est le 5-Fluorouracile (5-FU), découvert dans les années cinquante (Figure 1.3, (c)).¹⁴⁻¹⁶ Il s'agit d'un analogue d'uracile (Figure 1.3, (a)) avec un fluor en C-5 à la place d'un hydrogène. Il soigne la plupart des patients de stade trois, mais en cas de cancer de type avancé, le taux de guérison avoisine les 15%. Des méthodes existent afin d'améliorer l'efficacité de la chimiothérapie dont l'utilisation de médicaments complémentaires. C'est le cas du traitement le plus courant, le FOLFIRINOX, mêlant le 5-FU à d'autres médicaments.¹⁷ Dans ce cas, le taux de guérison est compris entre 40 et 50 %.^{9,14}

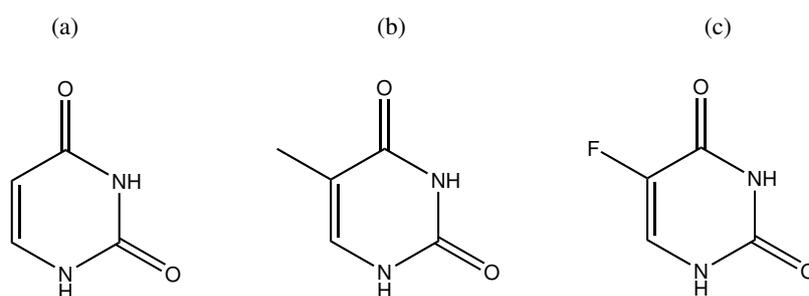


Figure 1.3 – Structure (a) de l'uracile présent dans l'ARN (b) de la thymine présente dans l'ADN (c) du 5-fluorouracile (5-FU) utilisé pour la chimiothérapie.

Le mode d'action du 5-FU se traduit en une double stratégie. D'une part, sa ressemblance avec l'uracile va permettre au médicament de jouer le rôle d'antimétabolite^C. Les cellules cancéreuses se reproduisant plus vite que les cellules normales, elles ont besoin d'une grande quantité de bases

C. Substance perturbant le métabolisme des cellules. Le métabolisme est un ensemble de processus chimiques nécessaires au fonctionnement, à la croissance et à la reproduction des cellules.⁹

nucléiques pour former l'ADN et l'ARN.^{14,18} Pour satisfaire ce besoin, le 5-FU entre dans les cellules à l'aide d'une diffusion passive et d'un transport actif.¹⁵ Une fois entré dans le cytosol, il va être converti en une série de métabolites (fluorodeoxyuridine monophosphate (FdUMP), fluorodeoxyuridine triphosphate (FdUTP) et fluorouridine triphosphate (FUTP)¹⁴) perturbant la synthèse de l'ADN, de l'ARN ainsi que l'action d'une enzyme, la thymidylate synthase (TS) (Figure 1.4). Cette enzyme ajoute un phosphate sur l'extrémité 5' du désoxyribose de l'ARN et réaliser la synthèse *de novo* de la thymidine-5'-phosphate, constituant l'ADN.¹⁸

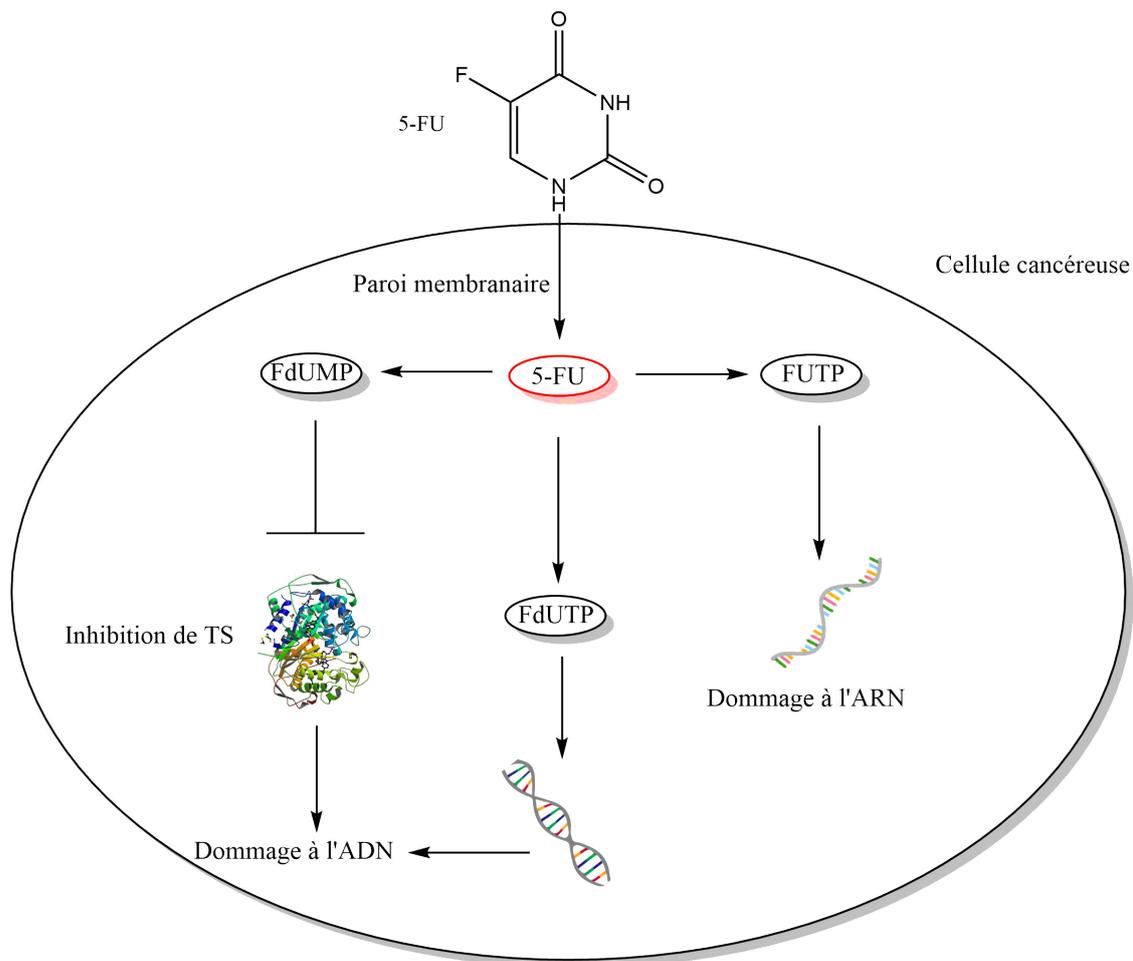


Figure 1.4 – Métabolisme du 5-FU dans la cellule tumorale. Adapté de Longley *et al.* Code de la thymidylase synthase (TS) : 1HVY.^{14,19}

Au vu des deux moyens de transport pour faire entrer le 5-FU dans la cellule, une plus grande disponibilité dans le cytosol par rapport à l'uracile et la thymine est observée. Comme elles ont un plus grand besoin en bases nucléiques, les cellules cancéreuses vont être moins sélectives et confondre le 5-FU avec ces bases afin de l'incorporer dans l'ARN et l'ADN de manière plus fréquente que dans les cellules saines.^{14,16} D'autre part, l'inhibition de la TS cause un stress méta-

bolique puisqu'elle fait partie de la voie métabolique dite "à un carbone" permettant de synthétiser des molécules biologiques essentielles pour la prolifération cellulaire.^{15,20} Le second mode d'action du 5-FU est d'induire la mort cellulaire en augmentant fortement la concentration en espèce oxygénée radicalaire (*Reactive oxygen species*, ROS) dans la cellule.^{6,21} L'accroissement de radicaux libres permet de provoquer un stress oxydatif causant des dommages aux nucléotides, aux lipides ou encore aux protéines contenues dans les cellules tumorales. La prolifération cellulaire étant empêchée, ce mode d'action est aussi responsable de la mort de ces cellules.²¹ Ce mécanisme d'action a pu être mis en évidence dans le cas d'autres médicaments tels que le trioxyde d'arsenic qui interfère avec les mitochondries et provoque la libération de ROS²² ou encore des agents anticancéreux, comme le cisplatine, ou les irradiations.²³

Pendant longtemps, le 5-FU a été considéré comme un médicament très puissant de part ces nombreux modes d'action. Cependant, au cours des années, le patient peut développer des résistances aux traitements.²³ Parmi les moyens de résistance les plus connus, la cellule peut s'adapter au traitement en changeant la cible de ce dernier par des mutations, en instaurant un nombre plus important de pompes à efflux afin de faire sortir le médicament de la cellule ou encore en augmentant le nombre de cibles afin que le médicament ne puisse plus les toucher. Dans le cas particulier du traitement au 5-FU, d'autres mécanismes de résistance ont pu être mis en évidence. Par exemple, les cellules cancéreuses apprennent à bloquer l'incorporation du 5-FU dans l'ADN et l'ARN²⁴ ou encore acquièrent une résistance à l'inhibition de l'activité de la TS.²³ Récemment, il a pu être démontré que les cellules tumorales développent un nouveau moyen de chimiorésistance dû à l'adaptation au stress oxydatif. En effet, en augmentant leur effet antioxydant, la diminution du stress oxydatif permet la prolifération cellulaire et la résistance à l'apoptose induite par les médicaments.⁷ En 2017, Xu *et al.* ont mis en évidence le lien entre la surexpression de la peroxiredoxin-2 (PRDX2) et la résistance au traitement par 5-FU. Cette résistance s'explique par le rôle d'enzyme antioxydante de PRDX2. Une fois surexprimée, l'enzyme diminue l'efficacité du 5-FU.⁶ Un mécanisme similaire de surexpression cellulaire a pu être mis en évidence par l'équipe de Li *et al.* en 2016 et confirmé par Sato *et al.* en 2017^{7,21} dans le cas de la phosphosérine phosphatase humaine (HPSP).

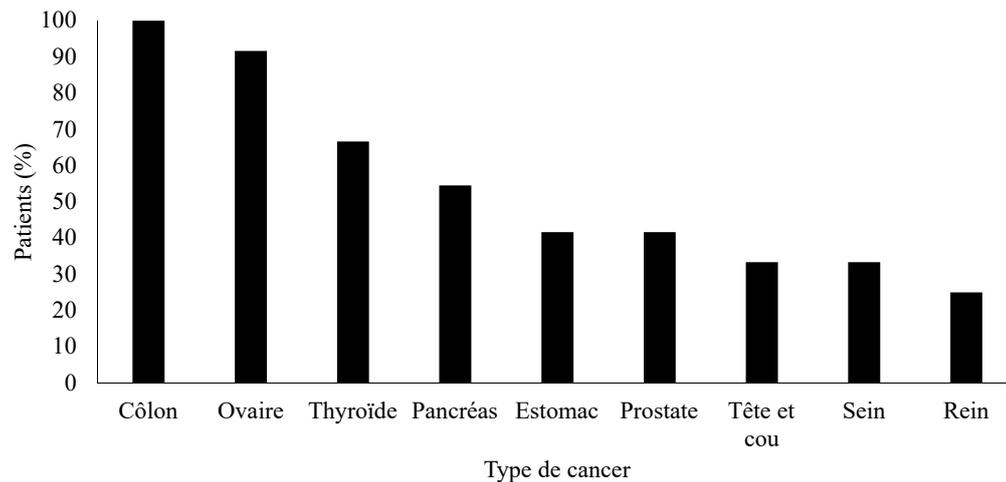


Figure 1.5 – Expression moyenne à élevée de HPSP dans les cellules cancéreuses en fonction du type de cancer.²⁵

Une augmentation de la quantité de HPSP dans les cellules tumorales a été observée dans le cas de nombreux cancers (Figure 1.5). Dans le cancer du côlon, une expression moyenne à élevée de HPSP est observée chez 100% des patients.^{21,25,26} Cette surexpression de HPSP affecte la probabilité de survie des patients. Effectivement, les courbes de Kaplan-Meier du cancer du côlon (Figure 1.6) montrent que la surexpression de HPSP diminue les chances de survie des patients. Un lien existe donc entre la surexpression de HPSP dans les cellules tumorales du cancer du côlon et la résistance au traitement.²¹ Toutefois, très peu d'études de HPSP sont réalisées afin de lutter contre cette résistance. La compréhension du mode de fonctionnement de l'enzyme est pourtant au centre de la problématique.

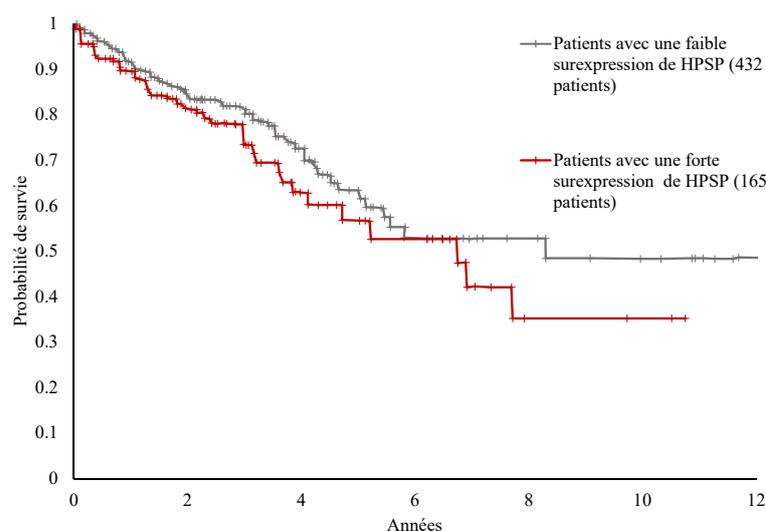


Figure 1.6 – Courbe de Kaplan-Meier du cancer du côlon. Adaptée de *The Human Protein Atlas*²⁵ via *WebPlotDigitizer*.²⁷

Chapitre 2

HPSP, une cible thérapeutique peu décrite

2.1 Identification de HPSP comme cible thérapeutique

Pour comprendre l'action de HPSP par rapport au 5-FU, il est nécessaire de comprendre son rôle dans l'organisme humain. HPSP est codée au niveau du chromosome 7 et est retrouvée habituellement dans tous les tissus humains faisant donc partie du métabolisme basal humain.^{25,28} Elle est la dernière enzyme de la voie de la sérine, où elle catalyse la formation de la L-sérine.²⁹ De cette implication dans la voie de la sérine naît la capacité de chimiorésistance induite par HPSP.

2.1.1 Voie de la sérine

La L-sérine est un acide aminé non essentiel découvert en 1865.^{30,31} Elle peut être obtenue dans notre corps de quatre manières différentes, à savoir par la nourriture, la synthèse *de novo* à partir de la glycolyse, à partir de glycine ou encore à partir de la dégradation de protéines ou de phospholipides.³¹⁻³⁵ Cet acide aminé est, par ailleurs, le point de départ de la synthèse de molécules biologiques, telles que les lipides ou les phospholipides, nécessaires pour la prolifération cellulaire via un métabolisme dit "à un carbone".^{20,30,34,36} Pour ce faire, la L-sérine est transformée en glycine par la sérine hydroxymethyltransferase (SHMT).³⁷ Le métabolisme "à un carbone" prend ensuite le relais et transforme la glycine en une série de métabolites. La L-sérine est aussi le précurseur de D-sérine, de neurotransmetteurs et de la cystéine.^{34,38}

Malgré les différents apports en L-sérine possibles, la synthèse *de novo* reste la voie principale de formation de L-sérine dans les cellules.^{35,38} Cette voie métabolique correspond à la voie dite phosphorylée^{31,39} utilisant comme réactif de départ le 3-phosphoglycérate (3PG) produit durant la glycolyse selon le chemin réactionnel décrit à la Figure 2.1. Dans cette voie, le phosphoglycérate est converti en trois étapes en L-sérine. Premièrement, la phosphoglycérate déhydrogénase (PHGDH) oxyde le substrat en 3-phosphohydroxypyruvate (3PHP). Ce dernier est ensuite transformé en phospho-L-sérine (PLS) à l'aide de la phosphosérine aminotransférase (PSAT1). Enfin, la phosphosérine phosphatase humaine (HPSP) permet la transformation de la phospho-L-sérine en L-sérine de manière irréversible.^{30,38,40}

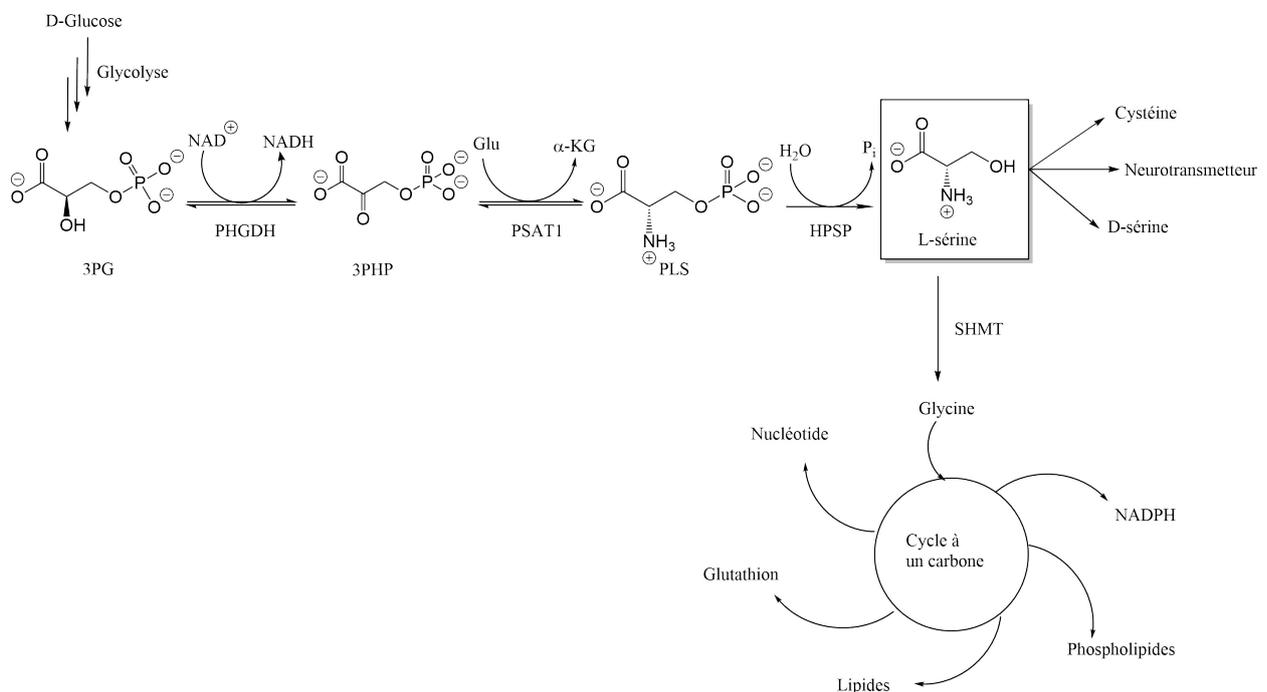


Figure 2.1 – Synthèse *de novo* de la L-sérine à partir de 3-phosphoglycérate.^{20,35,41}

L'implication de la voie de la sérine dans différents cancers est connue depuis les années 90.⁴² Des études ont d'ailleurs démontré l'importance de la L-sérine dans la régulation de la prolifération cellulaire, avec une portée cancéreuse.^{26,36,38} Cette implication est en lien avec le métabolisme "à un carbone", expliqué précédemment. À partir de la voie de la sérine, la suractivation de ce métabolisme en cas de cancer permet d'augmenter la quantité de métabolites impliqués dans la prolifération cellulaire. Les cellules cancéreuses qui suractivent cette voie prolifèrent donc plus vite. À titre d'exemple, PHGDH est surexprimée dans les cellules cancéreuses du cancer du sein et du mélanome.^{37,42} Cela active la voie de la sérine et donc le métabolisme "à un carbone" augmen-

tant le nombre de nucléotides afin que les cellules cancéreuses grandissent sans limitation. (Figure 2.1) Cette voie métabolique est une cible thérapeutique, mais est aussi montrée du doigt dans les cas de résistance à la chimiothérapie.^{20,41}

2.1.2 Implication de HPSP dans la chimiorésistance

La résistance au 5-FU au moyen de la surexpression de HPSP touche la seconde ligne d'action du 5-FU, décrite dans le chapitre précédent, à savoir la génération d'un stress oxydatif dans le but d'induire l'apoptose. Pour éviter l'effet de la chimiothérapie, le stress oxydatif généré par le 5-FU est contrebalancé par l'activité de la cellule. Pour ce faire, une suractivation de la voie de la sérine et des voies métaboliques qui en découlent est observée dans les cellules cancéreuses. HPSP étant responsable de la formation de L-sérine, la transformation de celle-ci en glycine et puis en glutathion par le cycle "à un carbone" permet à l'enzyme de lutter contre le stress oxydatif généré par le 5-FU.^{7,21} Dans le cas où HPSP est surexprimée, une surproduction de glutathions est observée et la régulation du stress oxydatif est améliorée. Ce mécanisme, décrit à la Figure 2.2, permet donc l'inactivation de l'apoptose par le 5-FU en contrebalançant le stress oxydatif généré par celui-ci, mais aussi la prolifération des cellules tumorales en augmentant la quantité de glycine convertie par la synthèse *de novo* en nucléotide. Cela fait de HPSP une cible thérapeutique de choix.

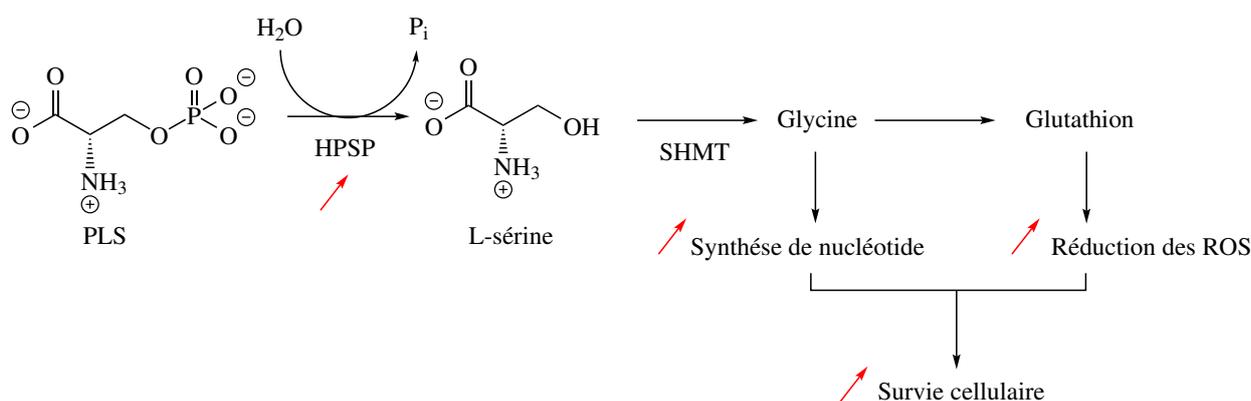


Figure 2.2 – Mécanisme de résistance au 5-FU observé dans les cellules tumorales. L'effet de la surexpression de HPSP est mis en évidence par les flèches rouges. Adapté de Sato *et al.*⁷

D'un point de vue mécanistique, HPSP est une protéine relativement méconnue. Dans le but de concevoir des inhibiteurs et d'améliorer l'efficacité des traitements, il est donc intéressant d'étudier l'enzyme.⁴¹ Le principe consiste à inhiber HPSP dans le but de supprimer la création excessive de glutathions dans les cellules cancéreuses et la perte d'activité du 5-FU. De plus, des études

ont prouvé qu'il était possible de vivre sans HPSP si un apport extérieur de L-sérine est mis en place.^{30,34,43} Une inhibition totale de l'enzyme pourrait, à terme, résoudre le problème de chimio-résistance du 5-FU. Cette stratégie est déjà utilisée dans d'autres cancers. Par exemple, dans le cas du cancer de type myélome, un traitement alliant des médicaments générant des ROS et des médicaments diminuant la capacité antioxydante cellulaire permettait d'être très efficace sur le cancer. Ce traitement, combinant le trioxyde d'arsenic et d'acide ascorbique, a ainsi été mis au point.²²

2.2 Caractéristiques connues de HPSP

HPSP (E.C : 3.1.3.3) est une enzyme dimérique cytosolique de 25 kDa comprenant 225 résidus.⁴⁴⁻⁴⁶ HPSP faisant partie de la superfamille des HAD (Haloacid dehalogenase),⁴⁷ elle possède des similarités structurales et mécanistiques avec cette superfamille protéique.³⁸ Il s'agit d'une enzyme magnésio-dépendante.^{48,49} Elle est régulée par son produit, la L-sérine.⁴⁵ Différents inhibiteurs de HPSP sont connus à l'heure actuelle : notamment, les molécules du type L-(+)-2-Amino-3-phosphonopropionic acid (AP3) sont des inhibiteurs compétitifs.⁵⁰ Les espèces telles que le BeF_3^- ou des analogues à base d'aluminium ou de magnésium possèdent aussi un rôle d'inhibiteur.^{51,52} Le remplacement du magnésium dans le site actif par un ion calcium provoque également un effet inhibant sur l'activité de la protéine.⁵³

2.2.1 Connaissances mécanistiques et structurales sur les HADs

Dans la nature, on compte environ 900 superfamilles de protéines basées sur des similitudes structurales et fonctionnelles.^{40,54,55} Les enzymes de la superfamille des HAD sont impliquées dans de nombreuses voies métaboliques primaires ou secondaires, dans l'entretien des cellules ou encore dans l'apport de nutriments.⁵⁴ Cette superfamille regroupe plus de 3000 protéines, allant des procaryotes à l'humain.^{54,56} La plupart des enzymes de cette famille catalysent des réactions de transfert de phosphate en brisant des liaisons C-Cl, C-P, CO-P ou PO-P.^{55,57} On retrouve 183 HADs chez l'humain.^{40,54,56}

Dans la grande superfamille des HAD, une sous-classe regroupe les phosphohydrolases dépendantes au magnésium dont HPSP fait partie. D'un point de vue mécanistique, comme le montre la Figure 2.3, Ridder *et al.*⁵⁸ décrivent la stabilisation du substrat dans le site actif au moyen d'une liaison de ce dernier à une sérine ou à une thréonine conservée. Le transfert de phosphate a ensuite lieu en deux étapes. Premièrement, le groupement phosphoryle est transféré à un aspartate, au moyen d'une attaque nucléophile.⁴⁰ L'hydrolyse s'effectue sur l'atome de phosphore et cela forme un intermédiaire de type aspartylphosphate.⁵⁸ L'aspartate catalytique peut être régénéré au moyen d'une molécule d'eau. Un second aspartate est observé, deux acides aminés plus loin dans la séquence. Ce motif est appelé DXDX(T/V).⁴⁰ Ce dernier possède un rôle acido-basique permettant la protonation du groupe partant lors de la première réaction et la déprotonation de la molécule d'eau lors de la seconde réaction.⁴⁰ Cependant, plusieurs possibilités sont décrites dans la littérature pour la déprotonation de la molécule d'eau. Il est aussi possible qu'elle soit réalisée au moyen d'une lysine conservée.⁵⁸ La présence du magnésium dans le site actif permet une proximité structurale des différents aspartates en écrantant l'excès de charge négative.⁵⁴

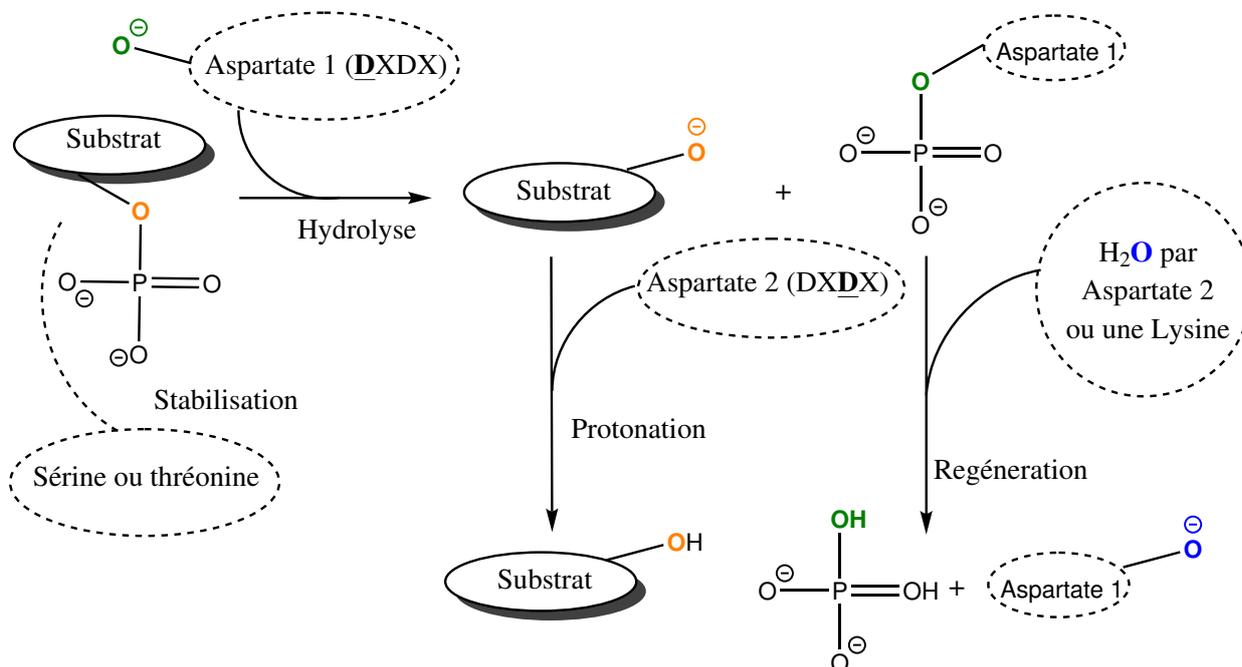


Figure 2.3 – Mécanisme général des HAD sur bases des informations issues de la littérature.^{40,58}

Les phosphohydrolases possèdent un domaine de coeur α/β^A possédant des ressemblances avec un motif de type *Rossmann fold*^B. Le motif catalytique, observable à la Figure 2.4, com-

A. Alternance de plans β reliés par des hélices α .

B. Motif structural composé d'une alternance de six feuilletts β reliés par deux paires d'hélices α .

prenant différents acides aminés hautement conservés, est composé de quatre boucles caractéristiques.⁵⁸ L'aspartate nucléophile est localisé sur la première boucle (en rouge) ainsi qu'une thréonine conservée selon le motif de type DXDX(T/V). Une sérine ou une thréonine conservée est observée sur la deuxième boucle à la fin du plan β , en vert. La troisième boucle (en bleu) contient une arginine ou une lysine totalement conservée permettant de contrecarrer l'excès de charge négative du système catalytique. La dernière boucle (en jaune) contient deux ou trois aspartates ou glutamates ayant un rôle de coordination du magnésium. Cette architecture structurale permet une orientation, une liaison et une activation du substrat afin de permettre la réaction.^{54,56,58}

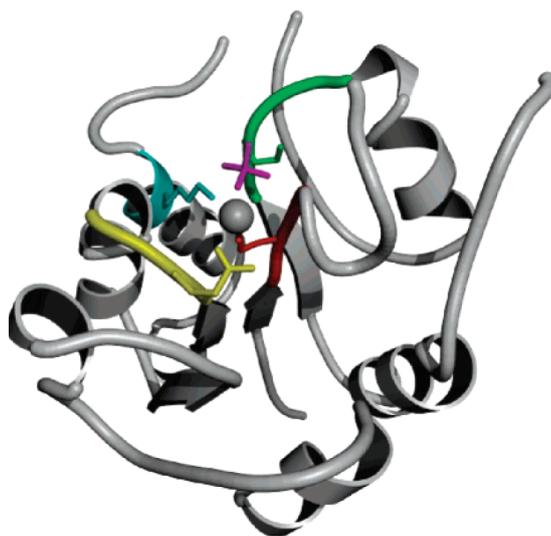


Figure 2.4 – Structure du site catalytique d'une protéine de la famille des HADs, la phosphosérine phosphatase (PDB 1L8L⁵⁰). Décrite par Lu *et al.*⁵⁶

Les HAD sont sous-divisées en 3 sous-classes en fonction de la présence ou non d'un domaine dit de coiffe.⁵⁴ Il s'agit d'un domaine dynamique, ouvrant et fermant l'accès au site actif. La première sous-catégorie possède le domaine sous une forme de motif *helix bundle*^C, la seconde sous-classe a un domaine mixte de α/β tandis que la troisième sous-catégorie n'a pas de motif coiffe. HPSP fait partie de la première catégorie en possédant une coiffe composée de 4 hélices. Dans cette sous-catégorie, une hélice dynamique possède deux conformations observées dans les HAD : une conformation ouverte et une conformation fermée. En conformation fermée, les acides aminés de l'hélice jouent un rôle dans la stabilisation des ligands et la catalyse. En conformation ouverte, le site actif est accessible au solvant et permet l'entrée du substrat et la libération du produit de la réaction.⁵⁶

C. Faisceaux d'hélices α parallèles ou antiparallèles entre elles.

2.2.2 Connaissances mécanistiques et structurales sur HPSP

Durant le début des années 2000, la structure dimérique de HPSP a été déterminée par deux laboratoires indépendants, à savoir celui de Peeraer *et al.*⁴⁴⁻⁴⁶ et celui de Kim *et al.*⁵⁰ de manière quasi simultanée. La *Protein Data Bank* (PDB)¹⁹ compte trois structures de HPSP, deux^D provenant de Kim *et al.* avec une résolution atomique de 2,51 Å et 2,80 Å ainsi qu'une^E de Peeraer *et al.*, possédant une résolution atomique de 1,53 Å. La structure obtenue par Peeraer *et al.* est reprise à la Figure 2.5. L'unité asymétrique contient deux monomères, nommés A et B ci-dessous. Les cristaux ayant permis l'obtention de cette structure ont été obtenus par la méthode de la goutte pendante. Les conditions de cristallisation comprennent une solution de CaCl₂ 0,7 M, du cacodylate 0,1 M à pH 6,0 et du PEG 1500 20%.⁴⁴ Le système cristallin orthorhombique appartient au groupe d'espace $C222_1$.

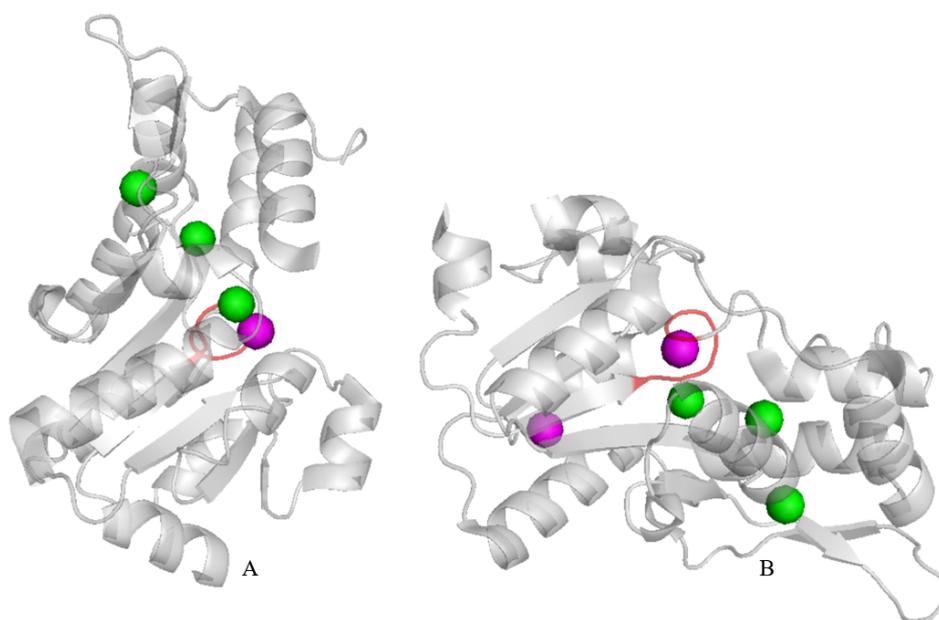


Figure 2.5 – Structure de HPSP découverte par Peeraer *et al.* possédant trois calciums (mauve), six ions chlorure (vert) et le motif DXDX(T/V) (rouge) avec une résolution atomique de 1,53 Å.⁴⁴

Après diffraction, résolution et affinement, la structure obtenue possède deux molécules dans son unité asymétrique ainsi que du chlore et du calcium. Par la présence de calcium et non de magnésium dans le site actif, la structure obtenue est une forme moins active de l'enzyme.⁵³ Il n'y a pas de substrat dans le site actif. D'un point de vue structural, la conformation ouverte de l'enzyme est observée laissant le site actif accessible au substrat. Cette conclusion a été réalisée à

D. 1L8L et 1L8O

E. 1NNL

l'aide d'informations sur l'agitation des atomes avoisinant la boucle dynamique. Cependant, aucune information sur la position de cette boucle n'existe par manque de données sur la position des atomes dans cette région. La structure proposée possède les caractéristiques typiques des protéines appartenant à la superfamille des HAD. Le cœur α/β , ressemblant à un motif de type *Rossmann fold* ainsi que le second motif de type coiffe consistant entre une interface entre les deux dimères et un *four helix bundle* sont observés. Le cœur α/β est composé de six plans β entourés de trois hélices α d'un côté et quatre de l'autre.^{45,46} Malgré la haute qualité de cette structure, peu d'informations mécanistiques peuvent en être déduites. Cependant, par comparaison avec la structure d'une autre phosphosérine phosphatase provenant de *Methanococcus jannaschii* (MJPSP)^{48,59} présentant 30% de similitude de séquence, des hypothèses mécanistiques ont été réalisées. Peeraer *et al.* propose une attaque nucléophile menée par l'Asp20 suivie d'une protonation de la L-sérine par l'Asp22. La stabilisation du phosphate pourrait être réalisée par la Lys158 et la Gly110 tandis que la L-sérine pourrait être stabilisée par le Glu29 et l'Arg65.^{45,46}

Par comparaison, les deux structures découvertes par Kim *et al.* sont renseignées aux Figures 2.6 et 2.7. Malgré la résolution atomique moyenne de ces structures, elles appuient les observations réalisées précédemment. Deux domaines sont observés, un domaine de cœur et un domaine dimérique avec un sous-domaine de type *four helix bundle*. Le dimère est sous une orientation différente en comparaison à la structure de Peeraer *et al.*, car les conditions de cristallisation et le groupe d'espace sont différents (*I4*). Le nombre de plans β et d'hélice α est identique à la structure de Peeraer *et al.* La position de la boucle dynamique est différente entre les deux molécules de l'unité asymétrique. La molécule A possède une conformation fermée tandis que la molécule B a une conformation ouverte. Cependant, aucune attribution de métal dans le site actif n'est possible, probablement dû à la moins bonne résolution atomique de la structure. La structure décrite sous le code 1L8L¹⁹ possède un inhibiteur dans chaque site actif, à savoir AP3. La présence de ce ligand met en évidence des interactions avec les acides aminés pour le stabiliser, comme cela est observable à la Figure 2.6. Pour le cas de la molécule A, la stabilisation s'effectue à l'aide de l'Asp20, l'Asp22, la Gly110, la Lys158, l'Asp183 et de l'Asp179. Dans la molécule B, on retrouve l'Asp20, l'Asp22, la Gly110, la Lys158 ainsi que le Glu29, la Thr182 et l'Arg202.

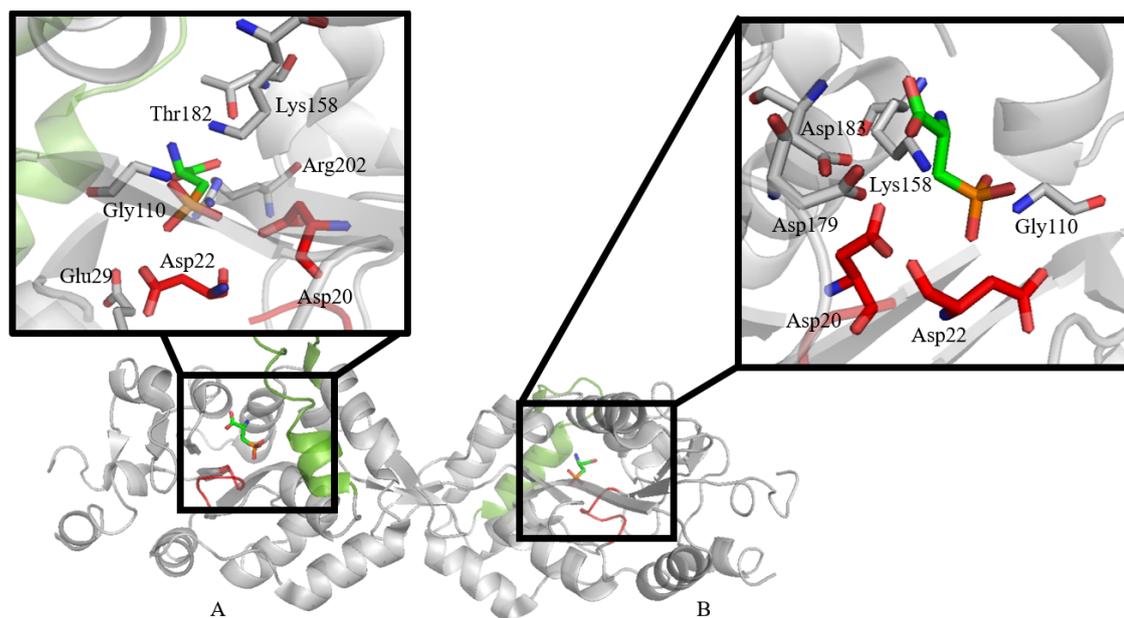


Figure 2.6 – Structure de HPSP découverte par Kim *et al.* contenant de l'AP3 dans chaque molécule (vert clair) en tant que ligands, la boucle dynamique (en bordeaux et orange) et le motif DXDX(T/V) (en rouge) avec une résolution atomique de 2,51 Å. Zoom sur les sites actifs des deux molécules. Détail des acides aminés impliqués dans la liaison au ligand.⁵⁰

Dans le cas de la seconde structure (1L8O) (Figure 2.7),¹⁹ un phosphate et une sérine sont observés dans chaque molécule de l'unité asymétrique. Dans les site actif des deux monomères, la L-sérine interagit avec la Thr182, l'Arg202 et l'Ala181. Le phosphate interagit avec l'Asp20, l'Asp22, la Ser109, la Lys158 et l'Arg202. L'orientation des deux ligands est similaire à celle de AP3 et permet des hypothèses mécanistiques appuyées par la mutagenèse dirigée. Ces analyses ont souligné l'importance de certains acides aminés (Asp20, Asp22, Glu29, Arg65, Asn133, Ser109, Lys158, Glu178, Asp179, Asp183, Arg202). Kim *et al.* attribue un rôle catalytique au Glu29 et à l'Arg65, essentiels à la liaison du substrat, et appuie le rôle de nucléophile de l'Asp20 grâce à son orientation. Les ligands de Kim *et al.* ont des orientations différentes des ligands observées dans MJSP pour la phospho-L-sérine, sur laquelle se base la hypothèse mécanistique de Peeraer *et al.* Un récapitulatif des acides aminés dont l'implication est supposée dans le mécanisme de HPSP est repris au Tableau 2.1, en fonction des différents apports de la littérature. Malgré l'apport des structures de HPSP pour comprendre le mécanisme de celle-ci, aucune de ces structures n'a été obtenue sous forme active, à savoir avec du magnésium. De plus, le rôle de nombreux acides aminés dans le mécanisme reste flou, l'orientation différente des ligands entre MJSP et les structures de Kim *et al.* étant étonnante. Une étude approfondie de l'enzyme permettrait de mieux comprendre son mécanisme.

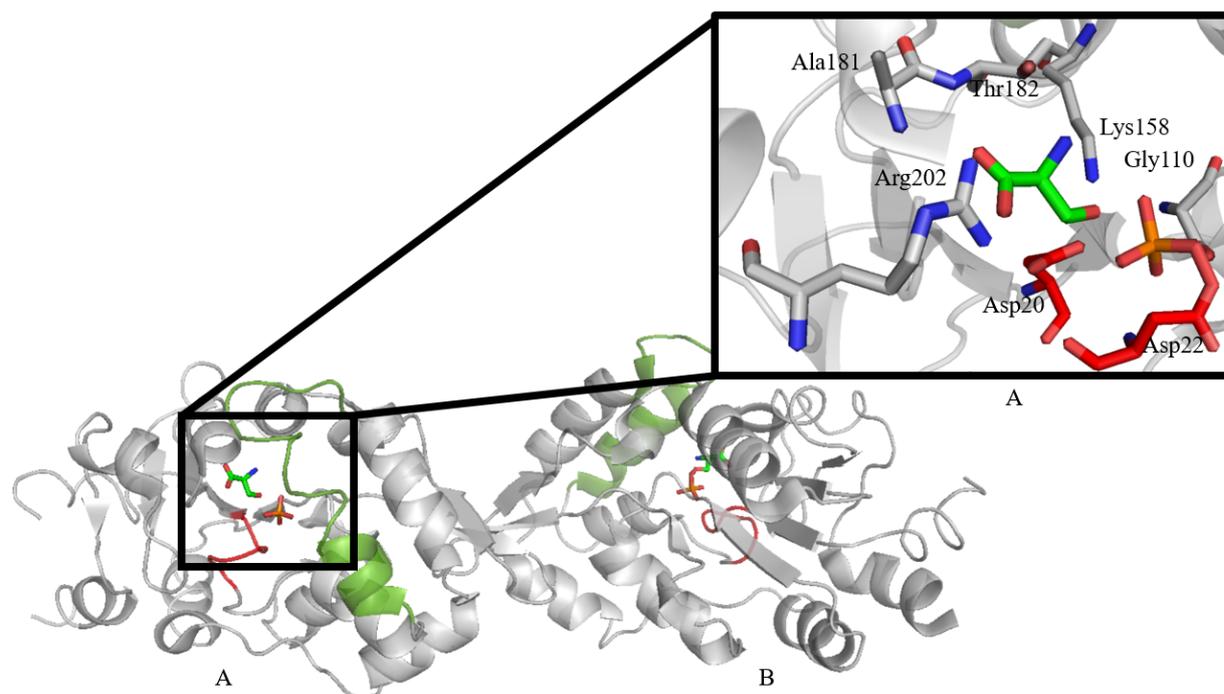


Figure 2.7 – Structure de HPSP découverte par Kim *et al.* contenant un phosphate (orange) et une sérine (en bordeaux) en tant que ligands, la boucle dynamique (vert foncé) et le motif DXDX(T/V) (en rouge) dans chaque molécule avec une résolution atomique de 2,80 Å. Zoom sur le site actif de la molécule A. Détail des acides aminés impliqués dans la liaison au ligand.⁵⁰

Tableau 2.1 – Récapitulatif des acides aminés importants dans le mécanisme de HPSP selon la littérature. L'effet observé lors de la mutagenèse est noté par les signes + (légère augmentation d'activité (entre 100 et 120 %)), – (perte légère d'activité (entre 0 et 50%)) et – – (perte forte d'activité (supérieure à 50%)). PLS = phospho-L-sérine^{46,50}

| Acides aminés | Hypothèse de Peeraer <i>et al.</i> | Hypothèse de Kim <i>et al.</i> | Mutagenèse |
|---------------|---|--|----------------------|
| Asp20 | Nucléophilie | Nucléophilie | D20E – –, D20N – – |
| Asp22 | Protonation | | D22E –, D22N – – |
| Glu29 | Stabilisation L-sérine | Stabilisation PLS | E29D – –, E29Q – – |
| Arg65 | Stabilisation L-sérine | / | R65K – –, R65A – – |
| Ser109 | / | Stabilisation PO ₄ ³⁻ | S109T +, S109A – – |
| Gly110 | Stabilisation PO ₄ ³⁻ | Stabilisation PLS | / |
| Asn133 | / | / | N133D +, N133A – – |
| Lys158 | Stabilisation PO ₄ ³⁻ | Stabilisation PLS/PO ₄ ³⁻ | K158R – –, K158A – – |
| Gly178 | / | / | G178A – |
| Asp179 | / | Stabilisation PLS | D179E –, D179N – – |
| Gly180 | / | / | G180A – – |
| Ala181 | / | Stabilisation L-sérine | / |
| Thr182 | / | Stabilisation L-sérine/PLS | T182S – –, T182V – |
| Asp183 | / | Stabilisation PLS | D183E –, D183N – – |
| Arg202 | / | Stabilisation L-sérine/PO ₄ ³⁻ | R202K – –, R202A – – |

Chapitre 3

Objectifs et stratégies

La haute fréquence d'apparition du cancer du côlon et la mortalité qui en découle nécessitent la mise au point de nouveaux traitements afin de permettre une meilleure qualité de vie des patients. La chimiorésistance des cellules cancéreuses colorectales à la chimiothérapie au 5-FU est aussi un enjeu crucial. De manière à inhiber, partiellement ou totalement, cette résistance, comprendre le mécanisme de l'enzyme impliquée dans la chimiorésistance est plus que nécessaire.

L'objectif de ce mémoire consiste en l'étude mécanistique de la phosphosérine phosphatase humaine (HPSP), une enzyme dont l'importance dans la chimiorésistance au 5-FU a pu être mise en évidence précédemment. Le but est de mieux comprendre le fonctionnement de cette enzyme avant de pouvoir s'intéresser au problème de chimiorésistance qu'elle provoque, mais aussi de concevoir de nouveaux inhibiteurs sur base des conclusions mécanistiques. Pour ce faire, une stratégie décrite à la Figure 3.1 est mise en place. Dans un premier temps, une étude expérimentale de l'enzyme doit être menée. Pour ce faire, une forme recombinante de HPSP (HPSP-His₆) est surexprimée à partir de souches *E. coli* modifiées. Par la présence de la queue polyhistidine, la purification a lieu par chromatographie d'affinité de type IMAC. Par la suite, la détermination des conditions de cristallisation idéales de HPSP est réalisée. Les conditions de co-cristallisation avec le substrat ou un analogue non hydrolysable sont également identifiées, permettant l'obtention de cristaux protéiques avec des ligands. Ces cristaux ont diffracté au synchrotron. Les structures ainsi obtenues permettent de mettre en évidence différentes hypothèses de mécanisme pour HPSP.

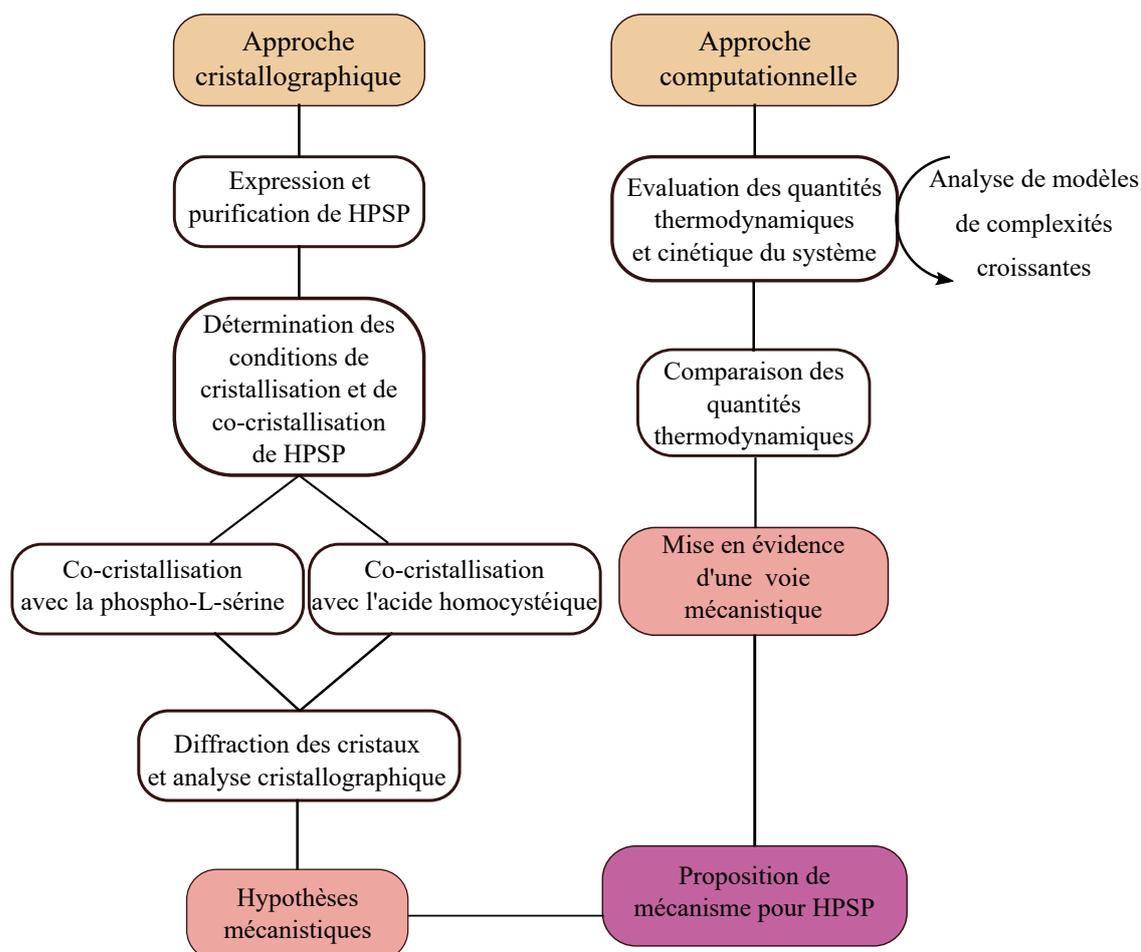


Figure 3.1 – Stratégies et objectifs suivis durant le mémoire.

Dans un second temps, une étude computationnelle est réalisée dans le but de confirmer ou d’infirmer les mécanismes supposés par les structures cristallographiques. Cette étude a pour but de déterminer les quantités thermodynamiques et cinétiques du système lors des différentes réactions impliquées lors de la réaction d’hydrolyse afin de mettre en évidence le chemin mécanistique le plus probable. Afin de souligner l’importance du système enzymatique, la simulation est réalisée à partir d’un système simple composé uniquement des réactifs et des produits, dans l’eau. Différents types de modèles sont ensuite examinés afin de voir l’effet de l’ajout d’acides aminés simulant l’environnement protéique et l’influence de ces acides aminés proches sur les différentes étapes de réaction. Les quantités thermodynamiques peuvent ainsi être rationalisées en fonction de la complexification du système.

Deuxième partie

Approche cristallographique

Chapitre 4

Expression, production et purification

La partie expérimentale de ce mémoire consiste en la résolution et l'affinement de la structure d'HPSP au moyen de la cristallisation, une des techniques les plus employées pour résoudre des structures protéiques à l'échelle atomique.^{60,61} Cependant, l'obtention d'une structure cristallographique est parfois laborieuse et des conditions optimales de cristallisation doivent être investiguées avant d'arriver à l'obtention de la structure souhaitée. Dans ce présent chapitre, la première étape de ce processus, à savoir la préparation de l'échantillon protéique au moyen de la production et la purification de la macromolécule, est détaillée. Ceci est réalisé au moyen d'un exemple représentatif de purification et est d'une importance capitale puisque la cristallisation ne peut avoir lieu que si la protéine possède une très haute pureté et une concentration assez élevée.^{60,61}

4.1 Expression et production de HPSP

4.1.1 Culture bactérienne et induction

À partir de souches *E. coli* thermocompétentes BL21 (DE3) possédant un plasmide codant pour HPSP fournie par le Dr Étienne De Plaen, la culture bactérienne est réalisée dans 400 mL de milieu. Le plasmide utilisé est composé d'un gène codant pour la protéine d'intérêt munie d'une queue polyhistidine en extrémité N-terminale, d'un promoteur T7, d'un gène lac, d'un gène de résistance pour un antibiotique, à savoir la Kanamycine™, ainsi qu'une origine de transcription. Ce système de transformation est connu sous le nom de système de type pET28 b. Il s'agit d'un système très efficace pour permettre l'expression d'une protéine.⁶²

Lors de la préculture, les bactéries sont placées dans un milieu (10 mL) de culture composé de Lysogeny Broth (LB) liquide et de la Kanamycine™ à une concentration de 50 mg mL⁻¹. La présence de Kanamycine™ dans ce milieu induit une pression de sélection sur les bactéries : seules les bactéries ayant incorporé le plasmide possèdent une résistance à cet antibiotique et peuvent se développer. Une fois la croissance cellulaire amorcée, celle-ci continue dans un milieu de culture plus grand (400 mL), de proportions identiques en antibiotiques, jusqu'à atteindre une densité optique comprise entre 0,6 et 0,8, correspondant à la phase de croissance exponentielle de la bactérie. Dès cet instant, l'induction de l'expression de la protéine peut débuter. Cela est possible par le système pET du plasmide. Ce système est connu pour permettre une induction de la protéine par le biais d'une ARN polymérase (T7) provenant de la cellule hôte. Cette ARN polymérase est extrêmement sélective et peut détourner le métabolisme de *E. coli* afin de produire en très grande quantité l'enzyme d'intérêt.

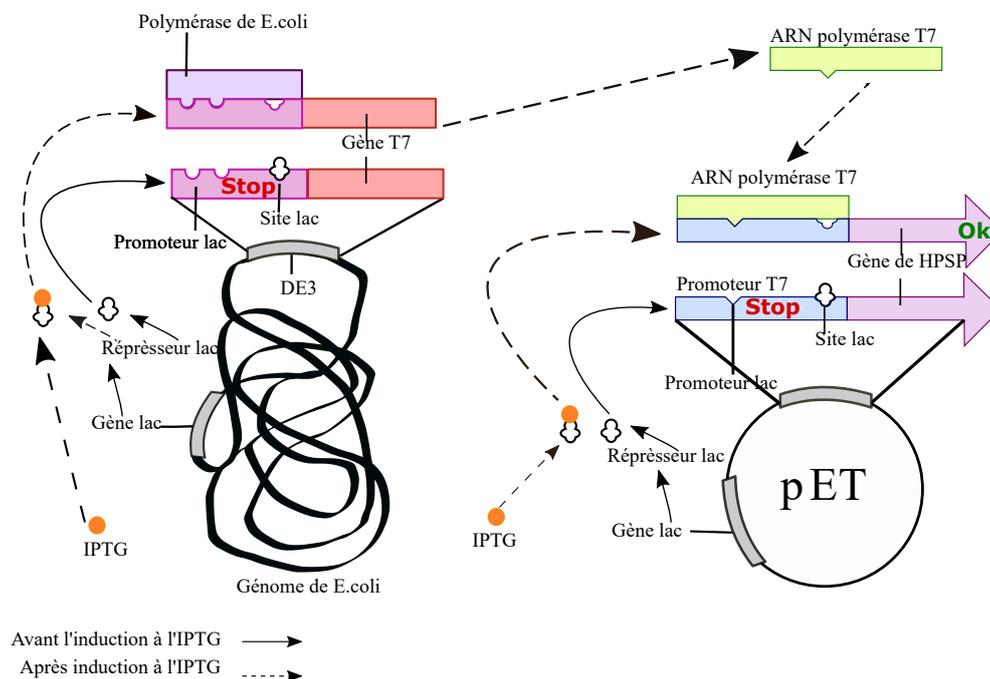


Figure 4.1 – Représentation du principe de l'induction en présence d'IPTG et d'un promoteur T7 dans une cellule de *E. coli*. Adapté de Novagen.⁶²

Le principe d'une induction est montré à la Figure 4.1. Dans la bactérie, elle est rendue possible par ajout d'isopropyl- β -D-galactopyranoside (IPTG). Cette molécule permet l'inactivation

des points de contrôle instaurés dans la bactérie et le plasmide. Ces points de contrôle ont pour but d'empêcher l'expression de HPSP lorsque la croissance cellulaire n'a pas atteint son maximum, évitant une toxicité pour les bactéries et une destruction de celles-ci. Les points de contrôle, situés au niveau du promoteur lac contenu dans le lysogène DE3 et dans le plasmide pET au niveau du promoteur T7, lient habituellement les répresseurs lac ce qui empêche l'expression de l'ARN polymérase T7 et, en conséquence l'expression de HPSP.⁶² Lorsque l'induction est enclenchée par ajout d'IPTG, cette molécule interagit avec les répresseurs lac. Cette liaison empêche la liaison future de ces répresseurs avec leur site de liaison, et rend possible l'expression de la polymérase T7. Cette polymérase T7 se lie au promoteur T7 et induit l'expression de HPSP. En quelques heures à peine, la machinerie basale de *E. coli* produit l'enzyme en grande proportion.

4.1.2 Lyse par sonification

Dès que la protéine est induite, celle-ci peut être extraite de la bactérie. Pour ce faire, une lyse par sonification est employée. Les parois membranaires de *E. coli* sont alors détruites à l'aide d'ultrasons. Les bactéries sont solubilisées dans un tampon de lyse contenant un cocktail d'inhibiteurs de protéases sans EDTA. Une détérioration des différentes protéines présentes dans ce mélange par des protéases est ainsi évitée. Du fait que l'EDTA n'est pas utilisé, la dénaturation des métallo-enzymes, comme HPSP, est empêchée. Dans ce mélange bactérien est plongée une tige métallique, soumise à une puissance de dix watts. Cette tige vibre et provoque des ultrasons, brisant les membranes bactériennes. Les organites et métabolites sont ainsi libérés dans le tampon.

4.2 Purification

Après la lyse, HPSP se trouve dans le surnageant de lyse. Cependant, ce surnageant contient énormément d'autres protéines contaminantes et de l'ADN provenant de *E. coli*. De ce fait, une purification est nécessaire afin de récupérer HPSP. Par la présence d'une queue polyhistidine en extrémité N-terminale de HPSP, cette purification a lieu par chromatographie d'affinité pour un métal de transition (*Immobilized Metal-Affinity Chromatography* ou IMAC). Ce type de purification est couramment utilisé dans un but ultérieur de cristallisation, car il permet d'obtenir la protéine d'intérêt avec une haute pureté.⁶¹

4.2.1 Principe de la purification par IMAC

La chromatographie d'affinité pour un métal (IMAC) est une technique de purification basée sur l'interaction avec un ligand, à savoir une protéine avec une queue polyhistidine, pour un métal tel que le nickel.^{63,64} Cela est possible grâce à la haute affinité du métal pour ce groupement. La colonne utilisée est composée d'une résine d'agarose sur laquelle des molécules d'acide nitrilotriacétique sont liées covalentement (NTA) et coordonnées à des ions nickel. La structure de cette résine est reprise à la Figure 4.2. Le NTA est capable de se coordonner au métal par quatre liaisons.

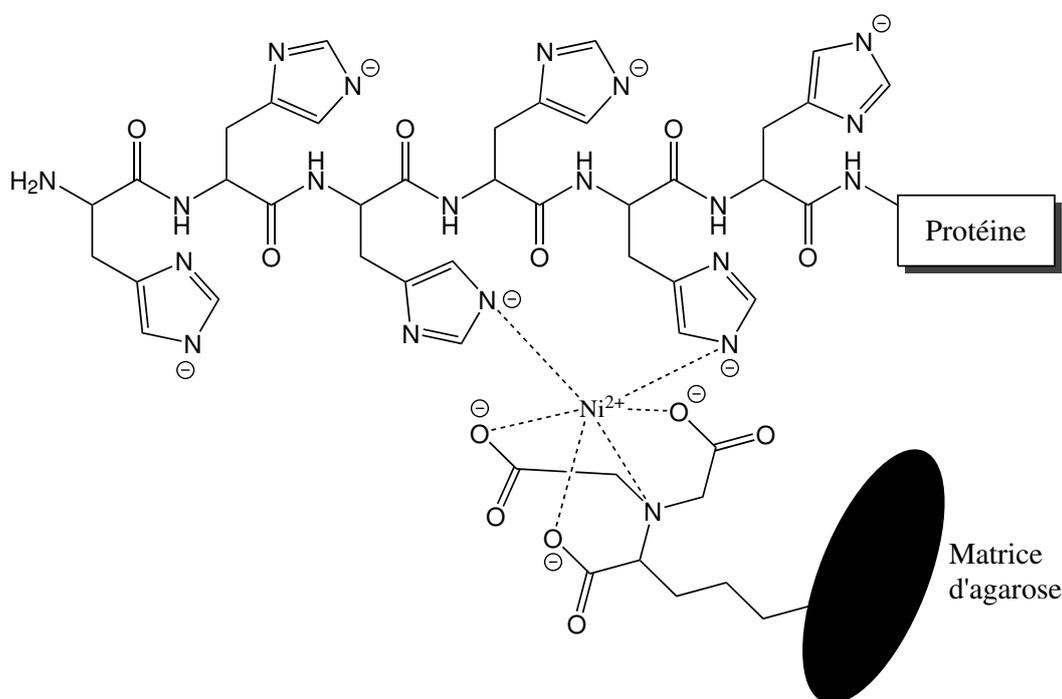


Figure 4.2 – Représentation de la résine d'agarose sur laquelle se trouvent des ions d'acide nitrilotriacétique (NTA) covalentement liés dessus, à pH physiologique. Inspiré de *Block*.⁶⁴

La purification commence par l'équilibration de la colonne au moyen d'un tampon contenant une faible concentration en imidazole, nommé tampon A (Figure 4.3). L'imidazole se lie à la colonne. Une fois cette étape d'équilibration réalisée, le surnageant protéique, obtenu après la lyse, est introduit dans la colonne. L'interaction entre la protéine histaguée, HPSP, et la colonne provoque un déplacement de ligand.⁶³ La colonne est ensuite lavée avec ce tampon A, permettant l'éluion des protéines non liées à la colonne. Cet ensemble de protéines s'appelle le *Flow Through*.

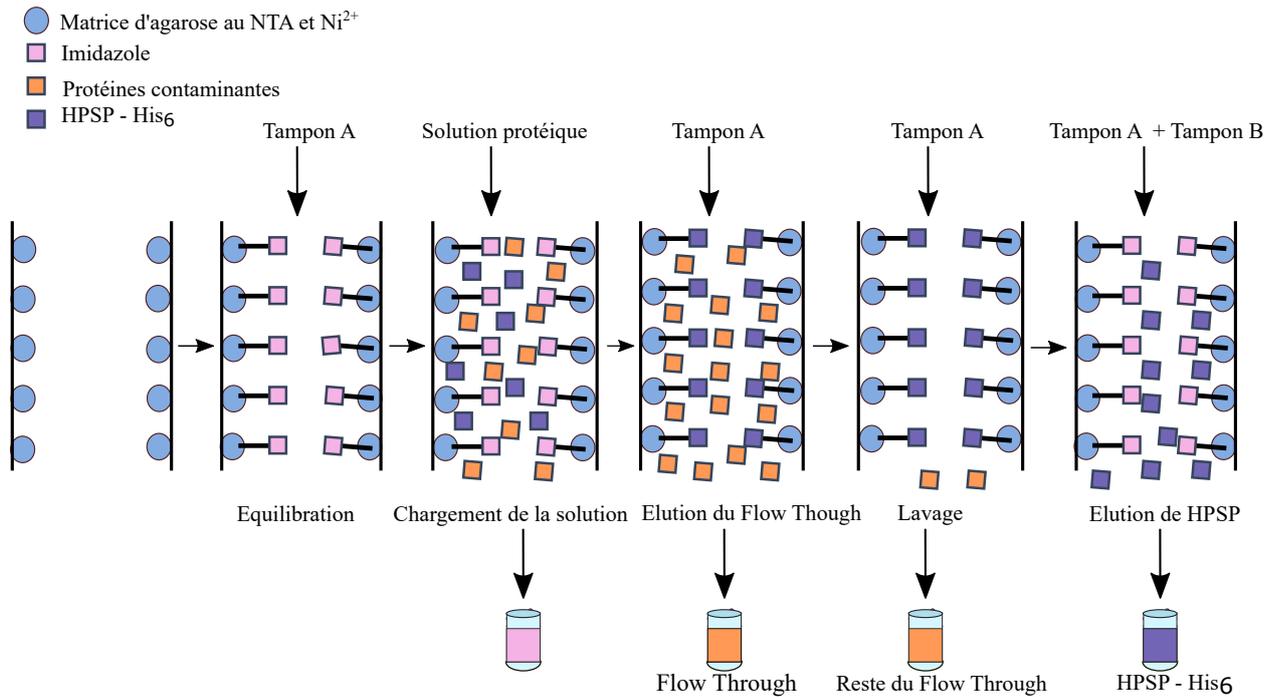


Figure 4.3 – Principe de la purification par IMAC d'une protéine polyhistaguée telle que HPSP.⁶³

Après élution du *Flow Through*, un gradient d'imidazole, réalisé par mélange du tampon A à un second, nommé tampon B, composé d'une plus grande quantité d'imidazole, est introduit afin d'éluer la protéine d'intérêt. L'imidazole de ce second tampon B entre en compétition avec la protéine histaguée au fur à mesure que la proportion de ce tampon augmente. À un moment donné, la compétition est tellement forte que les protéines histaguées se détachent au profit de l'imidazole. La protéine d'intérêt est donc récupérée.

4.2.2 Résultats de la purification par IMAC

Afin de suivre l'élution de la protéine au cours du processus, la colonne est reliée à un spectromètre UV ainsi qu'à un conductimètre. Le spectromètre mesure l'absorbance à 280 nm permettant de suivre l'élution des acides aminés aromatiques et le conductimètre mesure la conductivité en fonction du temps. Le chromatogramme ainsi obtenu au cours d'une purification est présenté à la Figure 4.4.

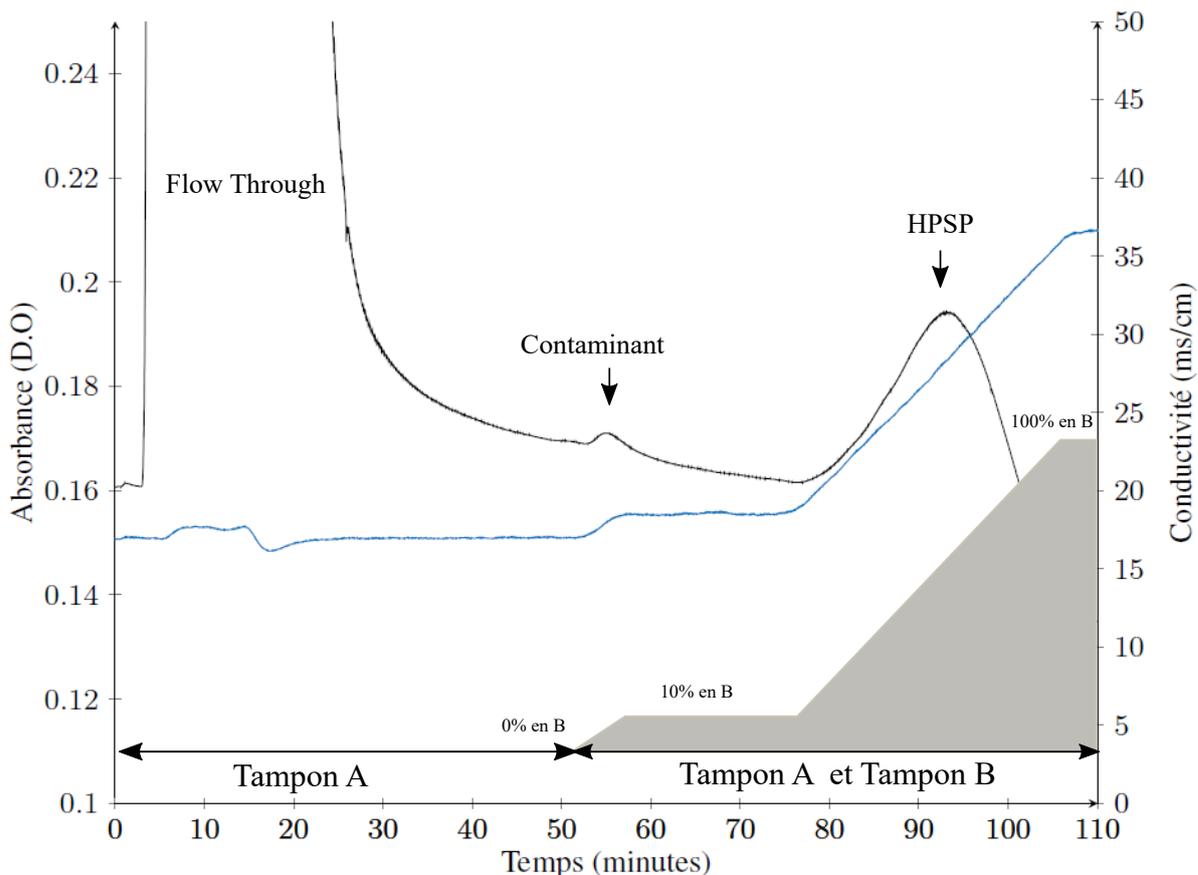


Figure 4.4 – Chromatogramme obtenu après purification par IMAC au nickel. L'absorbance apparaît en noir et la conductivité en bleu.

À partir de la Figure 4.4, on peut observer une élution du *Flow Through*, après 3 minutes et jusque 45 minutes. Ce dernier contient tellement de protéines qu'il y a une saturation de l'absorbance à 280 nm. Après l'élution du *Flow Through*, la colonne est lavée avec le tampon A afin d'assurer l'élimination de protéines contaminantes. À partir de 50 minutes, le mélange des deux tampons est progressivement introduit afin de permettre la formation d'un gradient en imidazole et d'entraîner l'élution de la protéine d'intérêt. Différentes fractions sont récoltées à partir de cet instant, par intervalle d'une minute. On constate deux pics d'élution lors de cette période : un premier, minoritaire, autour de 55 minutes et un second, majoritaire, vers 96 minutes. L'hypothèse est posée que ce premier pic est un contaminant et le second est la protéine d'intérêt. Afin d'apporter des preuves de la nature de ces deux pics, il est nécessaire d'effectuer un suivi de purification. Un premier aliquot, nommé HPSP, comprenant les protéines éluées entre 80 et 100 minutes et un second, nommé protéine contaminante, comprenant les protéines éluées entre 50 et 60 minutes, sont prélevés. Ces deux aliquots sont testés afin de vérifier leur contenu.

4.3 Suivi de purification

Le suivi de purification est réalisé de deux manières différentes, mais complémentaires. Dans un premier temps, la pureté de la protéine est analysée par un gel d'électrophorèse. Par la suite, un test d'activité est employé afin de vérifier si la protéine ainsi obtenue est active.

4.3.1 Suivi de purification par SDS-PAGE

Le suivi de purification par gel d'électrophorèse, aussi communément appelé SDS-PAGE pour *Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis*, est une technique courante permettant la mise en évidence de la pureté des fractions obtenues après purification en se basant sur la masse des différents composés macromoléculaires. Ces différentes fractions sont placées sur un gel d'électrophorèse coloré par la suite à l'aide de bleu de Coomassie ou de nitrate d'argent. Il est ainsi possible d'apprécier l'éventuelle pureté des fractions récoltées après purification.

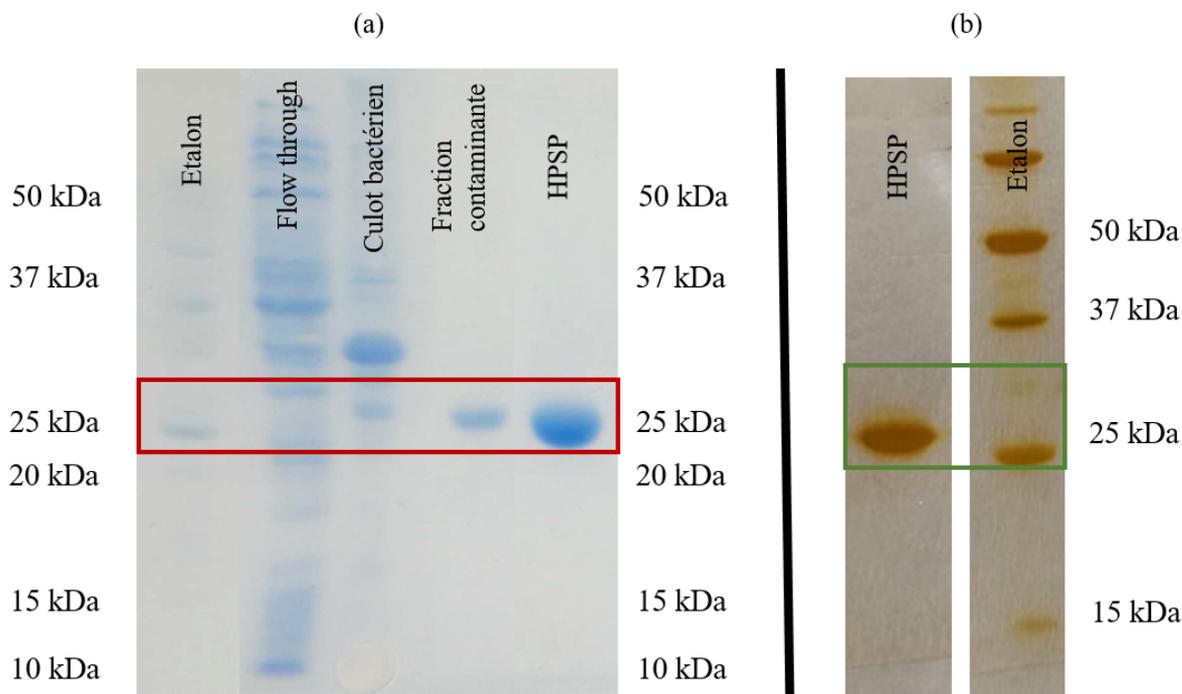


Figure 4.5 – Gel d'électrophorèse SDS-PAGE obtenu après purification. (a) Coloration au bleu de Coomassie. (b) Coloration au nitrate d'argent.

Le gel obtenu après purification, représenté à la Figure 4.5, montre que la fraction nommée HPSP contient une protéine de 25 kDa, correspondant au poids moléculaire attendu. Aucune im-

pureté n'est observée dans cette fraction, ce qui rend possible les essais de cristallogenèse. En ce qui concerne la fraction contaminante, la présence d'une protéine de poids moléculaire similaire est constatée. Il pourrait s'agir d'un autre état de repliement d'HPSP, qui n'éluerait pas au même moment que le reste suite à une affinité plus faible pour la colonne. D'autre part, en ce qui concerne le culot bactérien, une faible présence d'une protéine de 25 kDa est constatée ce qui suggère quelques pertes lors de l'étape de lyse. Cependant, la présence d'une autre protéine de même poids n'est pas à exclure. Le *Flow Through* ne possède pas une quantité importante de protéine de 25 kDa. La majorité de la protéine exprimée est donc récupérée.

4.3.2 Suivi de purification par test d'activité

Le test d'activité utilisé sur HPSP se base sur la dégradation de la phospho-L-sérine, par HPSP, en L-sérine et en phosphate. La présence de phosphate est ensuite dosée par la méthode colorimétrique décrite à la Figure 4.6. Par réaction du phosphate avec une solution formée de vert de malachite et de molybdate d'ammonium, un complexe vert absorbant à 655 nm se forme. L'absorbance de ce complexe est directement proportionnelle à la quantité de phosphate en solution et permet donc de mettre en évidence l'activité des fractions issues de la purification. Ce protocole a été découvert par Itaya *et al.*⁶⁵ et mis au point par un précédent mémoire.⁶⁶ L'avantage de ce test est qu'il permet d'observer rapidement si la protéine est active ou non.

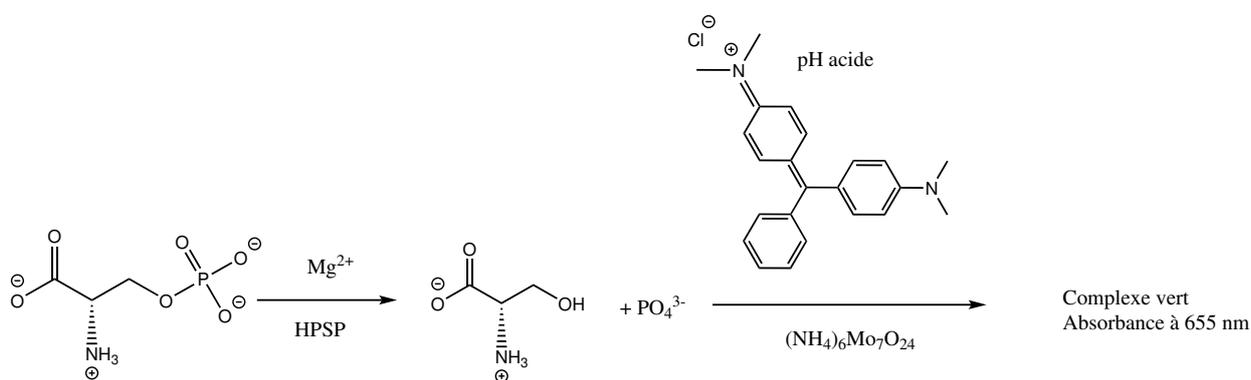


Figure 4.6 – Principe du test d'activité de HPSP.^{65,66}

Les différents résultats des tests, effectués en triplicats sur une quantité d'enzymes identique, sont repris au Tableau 4.1. Le premier résultat montre que la fraction HPSP est bien active. Elle n'a donc pas été dénaturée durant sa production. Le second résultat indique que la fraction contami-

nante possède aussi une légère activité phosphatase. Cette activité confirme l'hypothèse d'un autre état d'HPSP pour cette fraction. Enfin, le *Flow Through* possède lui aussi une activité phosphatase explicable par le fait qu'il existe, dans la machinerie d'*E. coli*, une phosphoserine phosphatase propre à la bactérie libérant du phosphate et donc réagissant au test.

Tableau 4.1 – Résultats issus du test d'activité réalisé sur les fractions récoltées après purification, pour une même quantité d'enzymes. Mesures réalisées en triplicats.

| Fractions | Concentration en PO_4^{3-} (μM) | Écart-type (μM) |
|-----------------------|---|------------------------------|
| <i>Flow Through</i> | 7,962 | 1,406 |
| Fraction contaminante | 5,513 | 0,659 |
| Fraction HPSP | 44,716 | 4,179 |

Au vu des résultats, la fraction nommée HPSP est collectée pour la suite des expériences. La concentration protéique de cette fraction, mesurée par Nanodrop™, montre que dix milligrammes d'enzymes sont récoltés pour 400 mL de culture.

4.4 Conclusions intermédiaires

Au cours de ce mémoire, plusieurs expressions et purifications ont été réalisées. Le processus d'expression et de purification représentatif développé dans ce chapitre montre que la protéine est produite avec succès à l'aide d'un processus d'induction à l'IPTG. La purification par chromatographie d'affinité de type IMAC permet de récolter HPSP sous forme pure et en quantité suffisante, à savoir dix milligrammes pour 400 mL de culture. Le suivi de purification met en évidence que la fraction ainsi récoltée a été obtenue sans contaminant tandis qu'un test d'activité souligne le fait que cette fraction est bien active. Au vu de ces résultats, les essais de cristallogénèse sont débutés.

Chapitre 5

Cristallisation de HPSP

Suite à la purification de HPSP, des essais de cristallisation peuvent être débutés. Il s'agit d'une étape primordiale afin de formuler une hypothèse mécanistique plausible pour l'enzyme. Elle permet l'obtention de cristaux protéiques, ce qui, ultérieurement, rend possible la diffraction par les rayons X des cristaux et la détermination structurale de HPSP, avec ou sans ligands dans son site actif. Toutefois, l'obtention de ces cristaux nécessite la détermination des conditions idéales de cristallisation.

5.1 Principe de la cristallisation

La technique de cristallisation utilisée est celle de la goutte pendante, une des techniques de cristallisation les plus courantes. Cette technique se base sur un processus d'équilibration par diffusion de composés volatils permettant la cristallisation protéique, schématisé à la Figure 5.1, (a).^{60,61} Une combinaison d'une faible quantité de solution protéique concentrée et d'une faible quantité d'une solution contenant différents composés, dont un précipitant, est réalisée. Ce mélange prend la forme d'une goutte sur une lamelle de verre, nommé *cover slide*. Le *cover slide* est ensuite placé et scellé au-dessus d'un réservoir contenant la même solution précipitante. Par effet de dilution, la concentration en solution précipitante est plus faible dans la goutte que celle du réservoir. En conséquence, une diffusion a lieu afin d'augmenter la concentration de la solution précipitante dans la goutte, entraînant une augmentation de la concentration en protéine. Ce processus a lieu jusqu'à l'obtention d'un équilibre des concentrations entre le réservoir et la goutte.

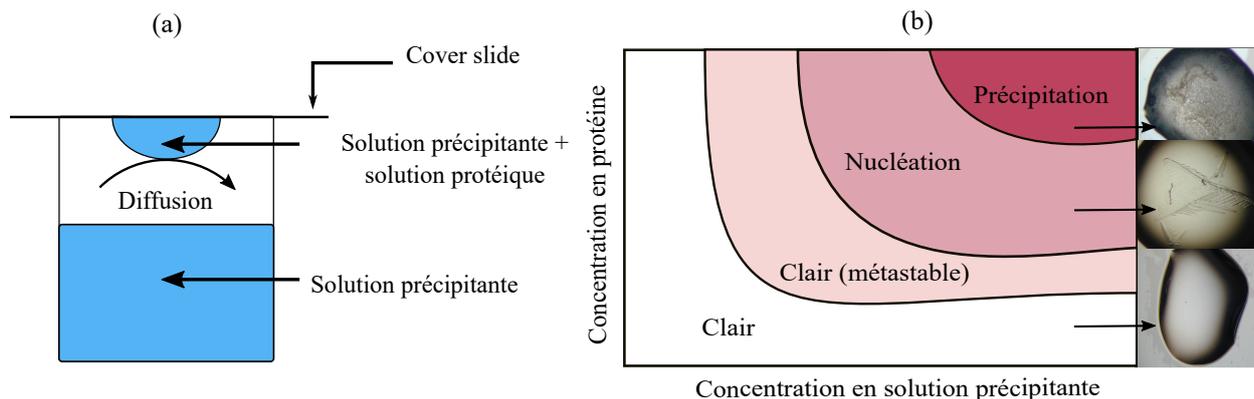


Figure 5.1 – Principe de la cristallisation de protéines (a) Schéma de la méthode de la goutte pendante. (b) Diagramme de phase du processus de cristallisation. Inspiré de Qiagen *et al.*⁶¹

Ce processus de cristallisation peut être expliqué au moyen d'un diagramme de phase présenté à la Figure 5.1, (b). Ce dernier met en évidence le comportement de la goutte en fonction des concentrations des composés respectifs. À basse concentration en protéine et en solution précipitante, la goutte est claire et rien n'y apparaît. Lorsque la sursaturation et la concentration en agent précipitant augmentent, un état métastable s'installe dans la goutte. Il mène ensuite à la phase de nucléation et à la formation de cristaux. La formation des premiers cristaux entraîne une diminution de la concentration en protéine et un retour à l'état métastable. Ainsi de suite, le schéma se répète, permettant ainsi au cristal de grandir. Le choix de la concentration en protéine ainsi que les composés dans le réservoir est primordial. Ceux-ci influencent la vitesse de diffusion. La diffusion jouant un rôle dans la concentration en protéine, elle peut entraîner la précipitation de la protéine si la concentration de cette dernière devient trop grande.⁶¹

5.2 Cristallisation en présence de phospho-L-sérine

Dans l'optique d'introduire des ligands dans le site actif de HPSP, la protéine est cristallisée avec de la phospho-L-sérine. Les conditions de cristallisation utilisées s'inspirent des conditions de Peeraer *et al.*,⁴⁶ conduisant à une structure cristallographique de haute résolution atomique. L'agent précipitant utilisé est du Polyéthylène glycol (PEG) 2000. Cet agent permet la diminution de la solubilité de la protéine.⁶¹ D'autres composés se trouvent dans la solution du réservoir, à savoir un tampon cacodylate et un sel ($MgCl_2$ ou $CaCl_2$). Le tampon a pour but de réguler le pH

lors du phénomène de diffusion. La solution saline, elle, ajuste la force ionique du milieu. Les conditions des différents composés dans le réservoir investiguées sont décrites au Tableau 5.1, les conditions de Peeraer *et al.* figurant à la condition **1**. La concentration de solution protéique s'élève à 6 mg mL^{-1} .

Afin d'isoler les co-cristaux de HPSP avec la phospho-L-sérine, la solution protéique utilisée pour la goutte est composée de la protéine concentrée à 6 mg mL^{-1} , de 5 mM de MgCl_2 afin de rendre la protéine active et de 10 mM de phospho-L-sérine comme substrat de HPSP. Afin de vérifier si la co-cristallisation n'affecte pas l'obtention de cristaux, une cristallisation avec une protéine dite native est réalisée, en absence de phospho-L-sérine et de MgCl_2 dans la solution protéique, sur le même *cover slide* pour servir de goutte de contrôle.

Tableau 5.1 – Conditions du puits pour les essais de cristallisation. La gamme de pH examinée s'étend de 5,5 à 7,0.⁴⁴⁻⁴⁶

| Conditions | Tampon | Sel | PEG |
|------------|----------------------------|---|--------------|
| 1 | Cacodylate $0,1 \text{ M}$ | $0,55 \text{ M} < \text{CaCl}_2 < 0,80 \text{ M}$ | PEG 2000 20% |
| 2 | Cacodylate $0,1 \text{ M}$ | $0,55 \text{ M} < \text{MgCl}_2 < 0,80 \text{ M}$ | PEG 2000 20% |
| 3 | Cacodylate $0,1 \text{ M}$ | $0,05 \text{ M} < \text{CaCl}_2 < 0,30 \text{ M}$ | PEG 2000 20% |

La condition **1**, dérivée de Peeraer *et al.*, est inadéquate. Cette condition mène à l'obtention de peu de cristaux et est peu reproductible. Il en est de même pour la condition **2**. La condition **3** s'avère très intéressante. Elle permet l'obtention de cristaux autant en cristallisation en condition native qu'en co-cristallisation, après seulement une journée. Différentes conditions optimales sont obtenues entre ces deux types de cristallisation. De manière générale, les cristaux obtenus en co-cristallisation sont plus grands que les cristaux obtenus par cristallisation native.

Dans le cas de la cristallisation en condition native, les cristaux apparaissent à partir d'une concentration en CaCl_2 dans le réservoir supérieure à $0,20 \text{ M}$. Aucun cristal n'est observé si la concentration est inférieure. Au-delà, par contre, on observe des cristaux à une gamme de pH allant de 6 à 7. Cette condition de cristallisation est reprise à la Figure 5.2, a. Dans le cas de la co-cristallisation avec la phospho-L-sérine selon la condition **3**, les cristaux n'apparaissent qu'à

basse concentration en sel, inférieure à 0,20 M. Le pH de 6,5 semble optimal pour la formation de cristaux (Figure 5.2, b). Au-delà de cette concentration, les cristaux deviennent progressivement plus petits jusqu'à devenir inexistant. La présence de $MgCl_2$ dans la solution protéique et de $CaCl_2$ dans le réservoir peut s'avérer très intéressante. En effet, vu la faible quantité de $CaCl_2$ dans les réservoirs ayant donné lieu à une cristallisation, il est dès lors possible de retrouver soit du Mg^{2+} , soit du Ca^{2+} dans le site actif.

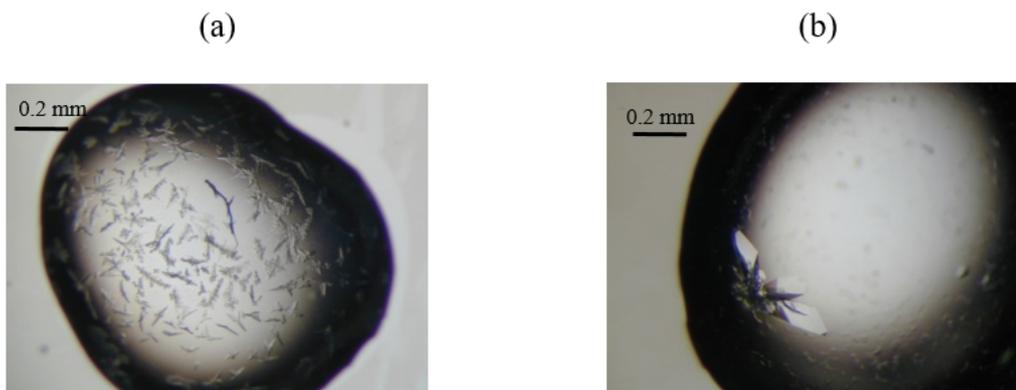


Figure 5.2 – Cristaux obtenus à partir de la condition **3** décrite au Tableau 5.1 (a) par cristallisation standard à pH 6,0 et à une concentration en $CaCl_2$ de 0,30 M. (b) par co-cristallisation, à pH 6,5 et à une concentration en $CaCl_2$ de 0,05 M.

5.3 Cristallisation en présence d'un analogue de substrat

Afin d'approfondir les connaissances mécanistiques sur HPSP, la cristallisation est également réalisée en présence d'un analogue non hydrolysable de phospho-L-sérine, à savoir l'acide homocystéique (HOM). Cette molécule, décrite à la Figure 5.3, possède des similarités structurales avec le substrat de l'enzyme. Les deux différences se trouvent au niveau du soufre remplaçant l'atome de phosphore et de l'absence d'oxygène en position γ , ce qui rend la molécule non hydrolysable.

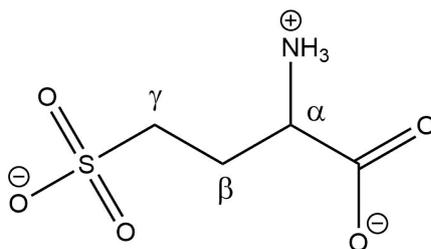


Figure 5.3 – Structure de l'acide homocystéique à pH physiologique.

La co-cristallisation de HPSP avec l'acide homocystéique est réalisée selon les conditions optimales décrites préalablement pour la co-cristallisation avec la phospho-L-sérine. Toutefois, de légères modifications des conditions de cristallisation ont lieu afin que l'acide homocystéique remplace la phospho-L-sérine. Premièrement, ce composé est mis en solution dans le DMSO, ce qui n'était pas le cas de la phospho-L-sérine. Des essais pré-expérimentaux mettent en évidence que la protéine native ne subit aucune influence dans le DMSO lors de sa cristallisation jusqu'à une concentration de 30%. Cependant, afin d'éviter un choc trop brutal entre l'analogue non hydrolysable dilué dans le DMSO et la protéine, il est ajouté directement dans les puits de cristallisation et non dans l'aliquot protéique. Les conditions du réservoir sont reprises au Tableau 5.2. En ce qui concerne l'aliquot protéique, il est composé de la protéine concentrée à 6 mg mL^{-1} ainsi que de 5 mM de MgCl_2 afin de rendre la protéine active.

Tableau 5.2 – Conditions pour les réservoirs de cristallisation lors de la co-cristallisation avec l'acide homocystéique.

| Composant | Concentration |
|---|--|
| Acide homocystéique | 8 mM dissous dans du DMSO 100% (4% de DMSO dans le réservoir) |
| Cacodylate de sodium | $0,1 \text{ M}$, pH 5,5 à 7,0 |
| Chlorure de calcium (CaCl_2) | $0,20 \text{ M} < \text{CaCl}_2 < 0,30 \text{ M}$ |
| PEG 2000 (monométhyl) | 20 % |

Il s'avère que la co-cristallisation conduit à une cristallisation optimale à plus haute concentration en sel que dans le cas de la phospho-L-sérine. L'apparition de cristaux est observée aux alentours de $0,25 \text{ M}$ en CaCl_2 . Le pH est optimal entre 6,0 et 6,5. Les cristaux obtenus sont observables à la Figure 5.4. Ces conditions de cristallisation sont plus proches des conditions de cristallisation de la protéine native. La présence d'acide homocystéique en solution dans le DMSO n'affecte pas le processus de formation de cristaux. Ceux-ci cristallisent en un jour et ont une taille légèrement plus grande.

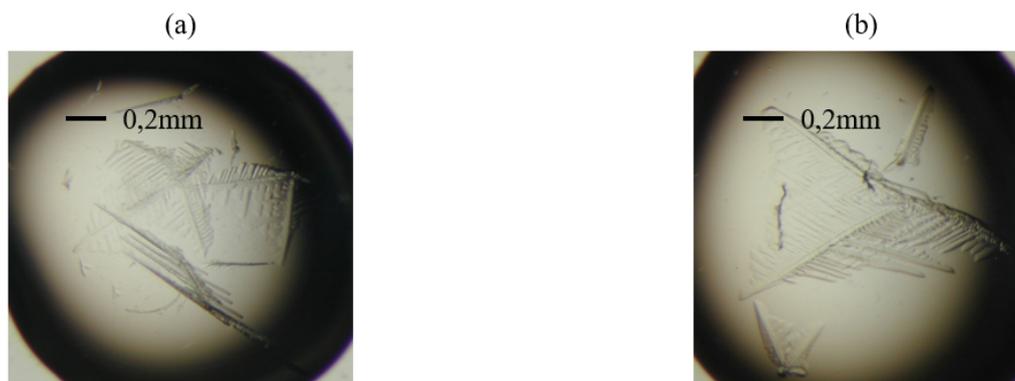


Figure 5.4 – Cristaux obtenus par co-cristallisation de HPSP avec l'acide homocystéique 8 mM dissous dans du DMSO 100% (4% dans le puit) dans du cacodylate de sodium 0,1 M et du PEG 2000 20%(a) pour une concentration en CaCl_2 de 0,25 M, à pH 6,0 (b) pour une concentration en CaCl_2 de 0,30 M, à pH 6,5.

5.4 Conclusions intermédiaires

Les conditions de cristallisation ainsi que les conditions de co-cristallisation optimales sont déterminées dans les différents cas. La protéine cristallise sous forme native dans le cacodylate 0,1 M, à un pH compris entre 6 et 7 et pour une concentration en CaCl_2 supérieure à 0,20 M en présence de 20% de PEG 2000. Lorsqu'une co-cristallisation avec de la phospho-L-sérine est réalisée, elle se produit pour des conditions en CaCl_2 inférieure à 0,2 M pour une même concentration de PEG et de cacodylate à pH 6,5. Enfin, si une cristallisation en présence d'acide homocystéique est envisagée, les conditions de co-cristallisations optimales se situent aux alentours d'une concentration en CaCl_2 de 0,25 M et un pH de 6,5, la concentration optimale en PEG 2000 et en cacodylate restant inchangée.

Chapitre 6

Étude structurale de HPSP

L'obtention d'une structure tridimensionnelle protéique à partir de cristaux nécessite la diffraction de rayons X (RX) par ces derniers. La source de RX est un rayonnement synchrotron. Ce type de rayonnement permet une meilleure modulation des longueurs d'onde, une analyse plus rapide, car les RX produits sont plus intenses, ainsi qu'une meilleure focalisation des faisceaux. La collecte est réalisée au synchrotron SOLEIL, situé à Paris. Pour faciliter la compréhension, les notions liées à l'affinement des structures cristallographiques sont expliquées ci-dessous. Le principe même de la diffraction est repris de manière plus détaillée dans la partie V.

6.1 Principe de l'affinement de structure protéique

Lors d'une expérience de diffraction, un cliché de diffraction, tel que celui présenté à la Figure 6.1, est obtenu. À partir de ce cliché, des intensités diffractées sont extraites. Elles fournissent les facteurs de structure observés, F_o . Le modèle affiné permet quand à lui le calcul des facteurs de structure calculés, F_c . Ces données permettent l'élaboration de cartes densités, outils du cristallographe pour affiner la structure atomique d'une macromolécule. Deux types de cartes, traitables par l'expérimentateur, sont ainsi isolables afin d'affiner le modèle : la carte de densité référant aux facteurs $F_o - F_c$ et la carte de densité référant aux facteurs $2F_o - F_c$. (voir partie V) L'affinement consiste à affiner les coordonnées x, y et z ainsi que le facteur d'agitation thermique^A des atomes.

A. Facteur rendant compte de la vibration des différents atomes et de la cohérence de la position des atomes par rapport aux cartes de densités. Il est aussi appelé facteur B.⁶⁷

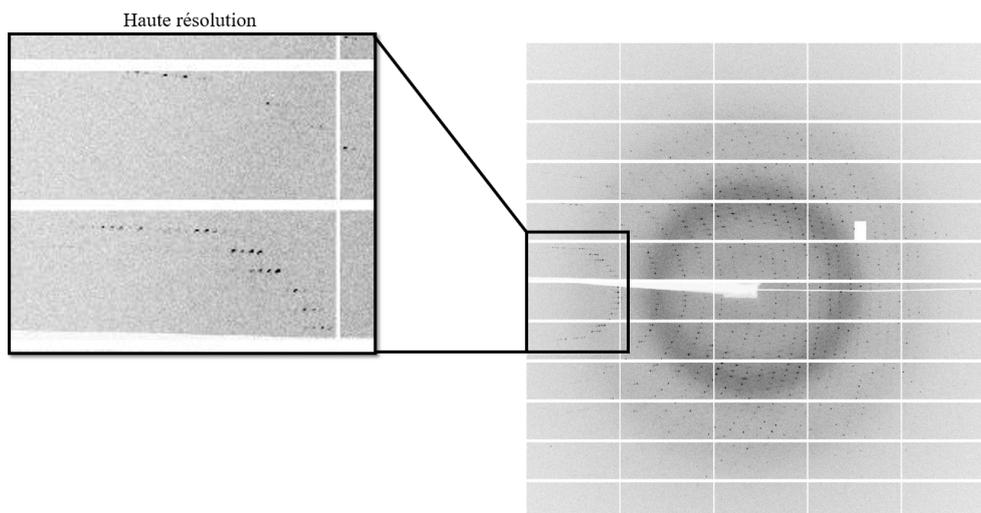


Figure 6.1 – Cliché de diffraction d'un cristal décrit dans le présent chapitre. Résolution de 1,57 Å.

Ce processus utilise une alternance d'optimisations suivie de corrections manuelles cherchant à positionner au mieux les atomes dans les cartes de densité. La structure est validée lorsqu'elle est jugée concordante avec les données expérimentales, notamment à l'aide de trois indicateurs :

Le *R-factor* indique l'écart entre les facteurs de structure calculés et les facteurs de structure observés, selon l'équation 6.1. Ce facteur mesure l'erreur commise dans le modèle et est directement influencé par la résolution atomique de la structure. On distingue deux types de *R-factors*, calculés sur des jeux de données différents, le R_{work} et le R_{free} . Pour des structures de qualité, les *R-factors* attendus sont de l'ordre de 20 % avec une différence comprise entre les deux *R-factors* entre 2% et 7%.⁶⁸

$$R - factor = \frac{\sum |F_o - F_c|}{\sum F_o} \quad (6.1)$$

Les RMSD, pour *Root-mean-square deviation*, renseignent le respect du modèle d'un point de vue chimique, pour les distances de liaisons ($RMS(bonds)$) et pour les angles ($RMS(angles)$). Le $RMS(bonds)$ montre un bon affinement de structure s'il possède une valeur inférieure à 0,02 Å. Au-delà de 0,03 Å, cela montre des erreurs dans le modèle proposé.⁶⁸

Les cartes de Ramachandran et de rotamères renseignent des angles de torsions les plus fréquemment observés dans les protéines. La différence entre ces deux notions s'explique par le fait que les cartes de Ramachandran s'intéressent aux angles de torsion du squelette carboné de la chaîne principale tandis que les rotamères s'intéressent, eux, aux chaînes latérales. Il est souhai-

table d'avoir un pourcentage d'angle de Ramachandran hors des cartes proche de 0,2%. Pour les rotamères, 1% des chaînes latérales hors des conditions idéales est tolérable.

La qualité d'une structure est aussi influencée par sa résolution atomique. Elle se définit comme étant la capacité à aller au-delà de grandes valeurs d'angles de diffraction lors de la collection des données sans rencontrer de phénomène d'interférence (Figure 6.1). Elle dicte la morphologie des cartes de densité. Plus le cliché de diffraction montre des spots éloignés du centre, plus le cristal diffracte à une haute résolution atomique et au mieux les cartes de densités sont définies. Si les cartes de densités sont bien définies, l'emplacement des atomes dans la macromolécule est plus précis. Une structure est dite de haute résolution atomique entre 1,2 Å et 2,0 Å, de moyenne résolution atomique entre 2,0 Å et 3,0 Å et de faible résolution atomique à partir de 3,0 Å.⁶⁸

6.2 Structure de HPSP en présence de phospho-L-sérine

Une collecte des intensités de diffraction est effectuée sur les cristaux obtenus en présence de phospho-L-sérine, par co-cristallisation, à un pH de 6,5 et à une concentration en CaCl_2 de 0,15 M observable à la Figure 6.2. Les cartes de densités sont obtenues après remplacement moléculaire^B et affinement à l'aide du programme Phenix,⁶⁹ interprétée via le programme Coot⁷⁰ tandis que l'attribution des éléments de structures secondaires a été réalisée avec le programme DSSP.⁷¹

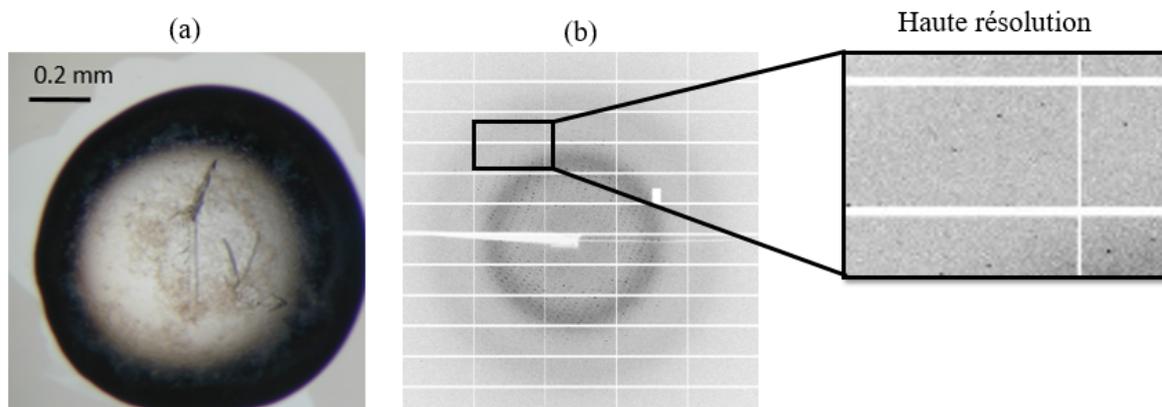


Figure 6.2 – (a) Cristaux obtenus par co-cristallisation avec la phospho-L-sérine, à pH 6,5 et à une concentration en CaCl_2 de 0,15 M dans un tampon cacodylate 0,1 M et une solution de PEG 2000 20%. (b) Cliché de diffraction du cristal. Résolution de 1,93 Å.

B. À partir de la structure INNL de Peeraer *et al.*

La résolution atomique s'élève à 1,93 Å. La structure ainsi obtenue après affinement est déposée dans la PDB sous le code 6HYJ. Elle est présentée à la Figure 6.3 et possède deux molécules dans son unité asymétrique, nommées A et B. Dans chacune d'entre elles, un ion métallique est retrouvé dans le site actif. La boucle dynamique, décrite dans la littérature et également localisée entre la Val40 à la Pro57, possède une structure repliée dans la molécule A et totalement dépliée dans la molécule B. La boucle de B étant moins définie, l'affinement a été réalisé seulement pour le squelette carboné de la chaîne principale. La présence de ligands dans la poche catalytique est constatée. Il s'agit du réactif de l'enzyme, la phospho-L-sérine, dans la molécule A et du produit de l'enzyme, la L-sérine, dans la molécule B. Dans les paragraphes ci-dessous, la détermination du type de métal ainsi que l'étude des ligands sont développées. L'étude de la boucle dynamique composant HPSP fera l'objet d'une étude détaillée.

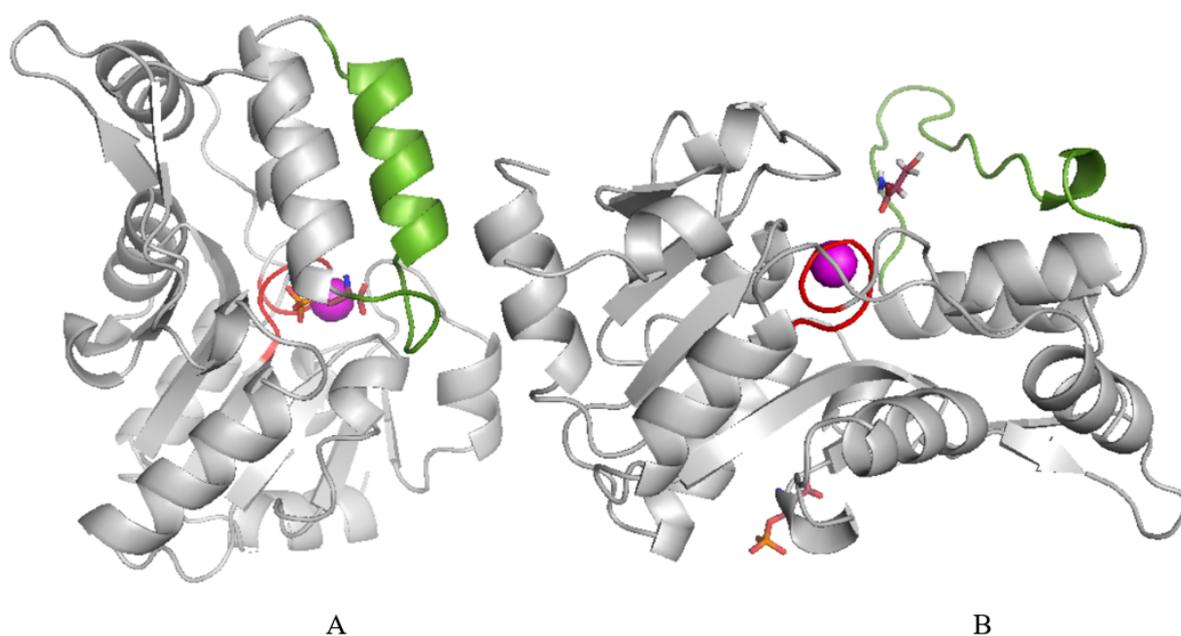


Figure 6.3 – Structure de la protéine issue du cristal obtenu par co-cristallisation à pH 6,5 et à une concentration en CaCl_2 de 0,15 M. Structure des deux monomères A et B. La boucle dynamique apparaît en vert, le métal en mauve, le motif catalytique en rouge et la phospho-L-sérine en bordeaux.

Les données relatives à la diffraction sont reprises au Tableau 6.1. Une différence acceptable entre les *R-factors* de l'ordre de 6,36% est remarquée. Il en est de même pour les valeurs de RMSD. Le facteur d'agitation thermique moyen est de 49 Å². En ce qui concerne les cartes de Ramachandran, deux acides aminés possèdent des angles de torsions hors des angles idéaux : la Val44 et l'Ala51, tous deux provenant de la molécule B. Ces acides aminés composent la boucle dynamique dépliée et possèdent donc des angles de torsion inhabituels dus au mouvement de cette dernière. Le pourcentage de rotamères hors des cartes est acceptable. Ces rotamères hors des cartes sont associés à la Phe11, l'Ile174, la Gln204 et la Leu220, tous compris dans la molécule B. La position de ces acides aminés semble correcte malgré des angles de torsion inhabituels.

Tableau 6.1 – Données relatives à la diffraction conduisant à la structure présentée à la Figure 6.3. Les données entre parenthèses correspondent à la plus grande sphère de résolution.

| Données de diffraction | | Affinement | |
|--|---------------------------|---|-------|
| Groupe d'espace | <i>C</i> 222 ₁ | <i>R</i> _{work} final (%) | 19,98 |
| Sous-unité | 2 | <i>R</i> _{free} final (%) | 26,34 |
| Paramètre de maille (Å et °) | a = 48,67, b = 128,97 | RMS(bonds) (Å) | 0,008 |
| | c = 155,71, α, β, γ = 90 | RMS(angles) (Å) | 1.19 |
| Résolution (Å) | 34,23 - 1,993 | Facteur B (macromolécule) (Å ²) | 48,89 |
| | (1,998 - 1,929) | Nombre de molécules d'eau | 185 |
| Réflexions unique | 36957 (3228) | Nombre de ligands | 3 |
| Complétude (%) | 98,66 (87,57) | Rotamères défavorisés (%) | 1,09 |
| Signal sur bruit (<i>I</i> / <i>σ</i>) | 11,27 (0,96) | Ramachandran défavorisés (%) | 0,45 |
| Multiplicité | 2,0 (1,9) | Ramachandran favorisés (%) | 94,77 |

La structure obtenue pour chaque monomère (Figure 6.4) est composée de treize hélices α ainsi que huit plans β et est similaire à celle retrouvée dans la littérature (Figure 6.5). Le motif catalytique, en rouge, est constitué de quatre boucles. Le motif correspondant au motif coiffe des HAD, en orange, est fait de quatre hélices α au lieu de trois dans la structure décrite dans la littérature. Cette boucle supplémentaire, en vert, est la boucle dynamique de HPSP.⁵⁶

L'étude plus approfondie de la structure obtenue nécessite de porter son intérêt sur le cation métallique qui compose le site actif. La cristallographie ne donne pas d'indication directe sur la nature de ce métal. Toutefois, cette information est relativement importante et peut être déterminée.

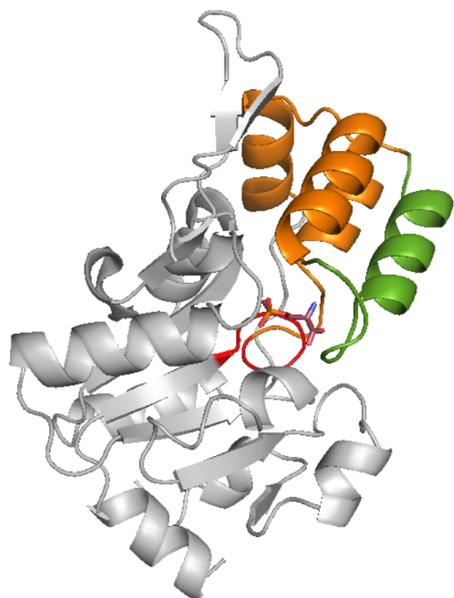


Figure 6.4 – Structure de la molécule A du dimère de HPSP. La boucle dynamique est en vert tandis que le motif de coiffe est en orange. Le motif catalytique apparaît en rouge.

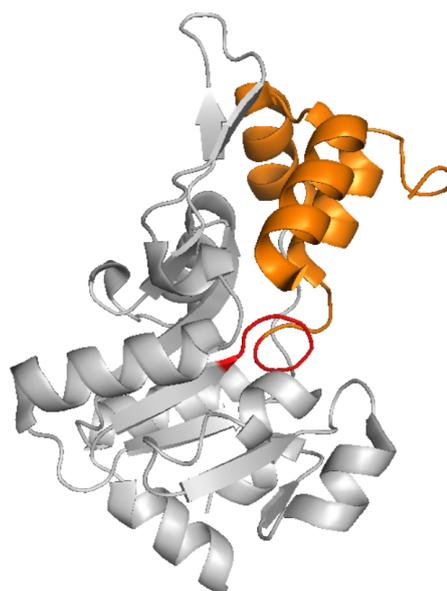


Figure 6.5 – Structure d'une molécule du dimère de HPSP présenté dans la PDB obtenue par Peeraer *et al.* (1NNL).^{19,46} Le motif de coiffe est en orange tandis que le motif catalytique apparaît en rouge.

Pour ce faire, une analyse selon trois indicateurs est utilisée, à savoir la géométrie des cartes de densités autour du métal, le facteur d'agitation thermique des atomes environnants et du métal et les distances de liaisons entre le métal et ces ligands. À partir des cartes de densité (Figure 6.6 et 6.7), dans le cas des deux molécules A et B, une coordination du métal de type octaédrique est constatée, avec six ligands. Les ligands impliqués dans cette sphère de coordination sont trois molécules d'eau et trois acides aminés : l'Asp20, l'Asp22 et l'Asp179, soit les ligands décrits dans la littérature.⁴⁴⁻⁴⁶ Cette coordination octaédrique est typique de la présence de magnésium, mais elle n'exclut pas la présence du calcium puisque ce dernier peut présenter une sphère de coordination à six, sept ou huit ligands.⁴⁶ Il est donc nécessaire d'analyser d'autres indicateurs.

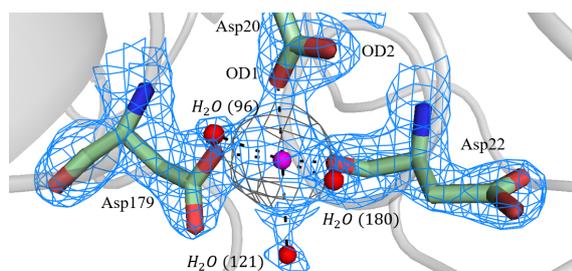


Figure 6.6 – Carte de densité du site actif de la molécule A.

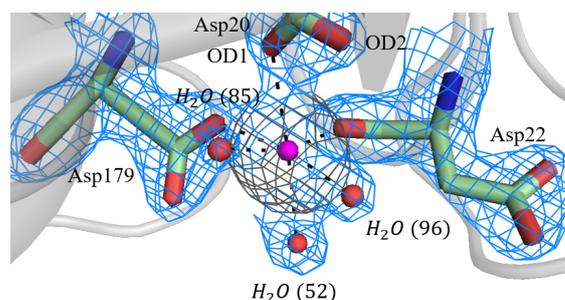


Figure 6.7 – Carte de densité du site actif de la molécule B.

Les facteurs d'agitation thermique appartiennent aux indicateurs couramment utilisés pour créditer la présence d'un type de métal ou non. En effet, ceux-ci témoignent de l'agitation des atomes et de la cohérence avec la carte de densité. Pour qu'un atome soit considéré comme cohérent avec la densité qui l'entoure, il est nécessaire que son facteur d'agitation thermique soit faible et cohérent avec l'environnement. Dans le cas présent, une variation notable de ces facteurs est constatée entre la présence d'un ion magnésium ou calcium dans les molécules. Si du calcium est placé dans le site actif des deux monomères, un facteur d'agitation thermique de $36,11 \text{ \AA}^2$ est obtenu pour la molécule A et de $38,07 \text{ \AA}^2$ pour la molécule B. À l'inverse, si du magnésium est dans le site actif, un facteur d'agitation plus faible, de $23,74 \text{ \AA}^2$, est observé pour la molécule A et de $27,09 \text{ \AA}^2$ pour la molécule B. Cependant, les facteurs d'agitation obtenus pour le magnésium ne sont pas cohérents avec les facteurs d'agitation obtenus dans le cas des acides aminés environnants. En effet, la valeur de ceux-ci s'approche des valeurs de facteurs d'agitation thermique obtenues pour le calcium. Enfin, les distances de liaisons entre les ligands et le métal sont étudiées. Celles-ci (Tableau 6.2 et Figure 6.8) suggèrent la présence de calcium. En conclusion, sur base des trois indicateurs, un ion calcium est présent dans les deux sites actifs.

| Ligands | A (Å) | B (Å) |
|---------|-------|-------|
| Eau 1 | 2,43 | 2,38 |
| Eau 2 | 2,42 | 2,20 |
| Eau 3 | 2,24 | 2,60 |
| Asp 20 | 2,25 | 2,38 |
| Asp 22 | 2,31 | 2,24 |
| Asp 179 | 2,18 | 2,31 |

Tableau 6.2 – Distances entre l'ion métallique et les ligands pour la molécule A et B.^{46,72}

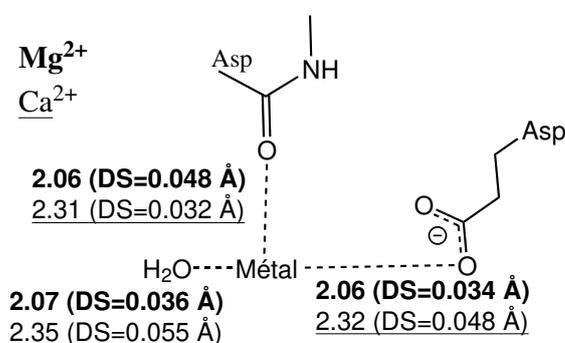


Figure 6.8 – Distances M-O provenant de la CSD pour les aspartates et l'eau : moyenne et déviation standard (DS) En gras : Mg²⁺ souligné : Ca²⁺.^{46,72,73}

La structure protéique obtenue après affinement met en évidence la présence de ligands dans l'unité asymétrique. Ces ligands varient en fonction de la molécule. En plus d'un métal dans le site actif, on retrouve une molécule de phospho-L-sérine (PLS) dans de la molécule A. Dans le cas de la molécule B, une molécule de phospho-L-sérine est aussi observable, mais à la surface de la macromolécule tandis qu'une molécule de L-sérine se loge dans le site actif. L'orientation

de ces ligands s'explique par une série d'interactions stabilisantes au sein même de la protéine par l'intermédiaire d'acides aminés. En ce qui concerne le site actif de la molécule A, la Figure 6.9 souligne la présence de phospho-L-sérine selon un facteur d'occupation de 46%. Celle-ci est en interaction avec des acides aminés mentionnés dans la littérature comme étant essentiels pour l'activité de HPSP. Effectivement, on observe une proximité entre l'Asp20 et l'Asp22 et le substrat, ce qui est intéressant compte tenu du rôle supposé de nucléophile et de base de ces deux acides aminés. De plus, le substrat est stabilisé dans cette position par la présence de l'Arg65, la Lys158 et de la Thr182. Un relai de ponts hydrogène, partant de l'Arg65, stabilise l'amine de la phospho-L-sérine. Le récapitulatif des différentes interactions est repris au Tableau 6.3.

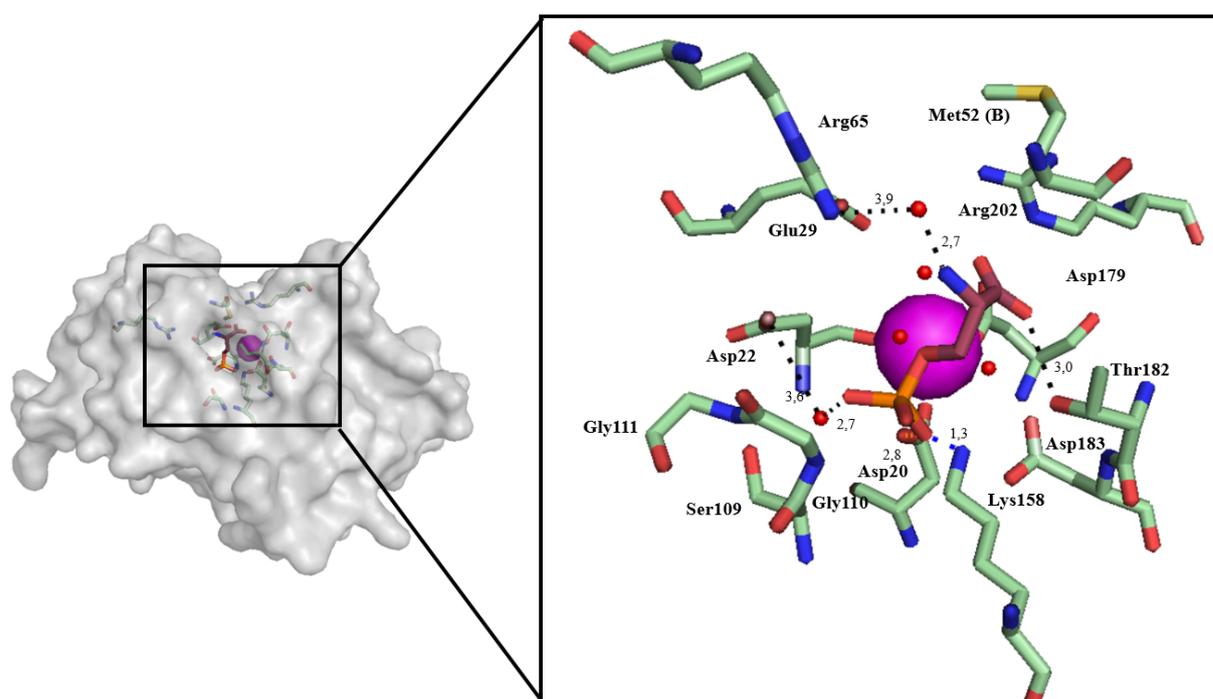


Figure 6.9 – Structure de la molécule A. Vision en volume. La phospho-L-sérine apparaît en bordeaux et le métal en rose. Les ponts hydrogènes apparaissent en noir, les interactions ioniques en bleu et les distances montrant une possible réactivité en orange.

Tableau 6.3 – Interactions entre les acides aminés et le substrat contenu dans la molécule A, présenté à la Figure 6.9. Les atomes concernés par la mesure des distances sont soulignés.

| Ligands | Acide aminé | Distance (Å) | Type d'interactions |
|--------------------|-----------------------------|--------------|---------------------|
| PLS (<u>OP</u>) | Asp 20 (<u>COO</u>) | 2,8 | / |
| PLS (<u>OP</u>) | Eau - Asp 22 (<u>COO</u>) | 2,7 - 3,4 | Pont hydrogène |
| PLS (<u>NH3</u>) | Eau - Arg 65 (<u>NH2</u>) | 2,7 - 3,9 | Pont hydrogène |
| PLS (<u>OP</u>) | Lys 158 (<u>NH3</u>) | 1,3 | Ionique |
| PLS (<u>COO</u>) | Thr 182 (<u>OH</u>) | 3,2 | Pont hydrogène |

La Met52 adopte deux conformations en fonction de la présence de la phospho-L-sérine, justifiable par l'encombrement stérique (Figure 6.10). Quand le site actif est vide, la Met52 adopte la position A, observée dans 57% des cas. Toutefois, lorsque le substrat est dans le site actif, la Met52 est en position B, évitant un encombrement stérique défavorable (43% des cas). Cette méthionine étant présente dans la boucle dynamique, il pourrait s'agir d'un contrôle de fermeture de celle-ci lors de l'arrivée du substrat, à l'image d'un loquet.

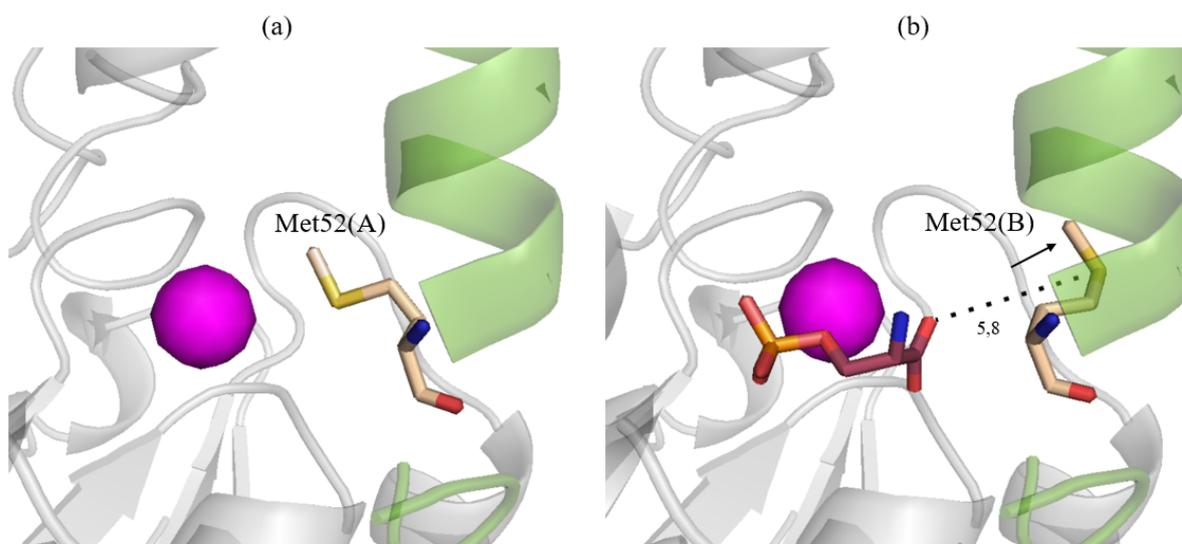


Figure 6.10 – Positionnement de la Met52 en fonction de l'occupation du site actif (a) lorsque le site actif est vide (57% des cas) (b) lorsque la phospho-L-sérine occupe le site actif (43% des cas). L'ion métallique est en rose et l'hélice dynamique est en vert. La phospho-L-sérine est en bordeaux.

Dans la molécule B, une molécule de phospho-L-sérine n'est pas observable dans le site actif, mais à l'extérieur de la protéine (Figure 6.11). Cette phospho-L-sérine est stabilisée à l'aide principalement de deux acides aminés, à savoir la Phe131 et la Lys168. L'interaction avec la phénylalanine est de type π -cation tandis que celle avec la lysine est de type ionique. Le site actif de la molécule B, présentée à la Figure 6.12, contient, lui, une molécule de L-sérine. Elle semble être en train de sortir de la protéine. La stabilisation de ce ligand est garantie par la présence de l'Arg202 ainsi que des acides aminés d'une molécule voisine dans la maille cristalline. Le Glu29 repousse le ligand vers la sortie. Ces interactions sont reprises au Tableau 6.4.

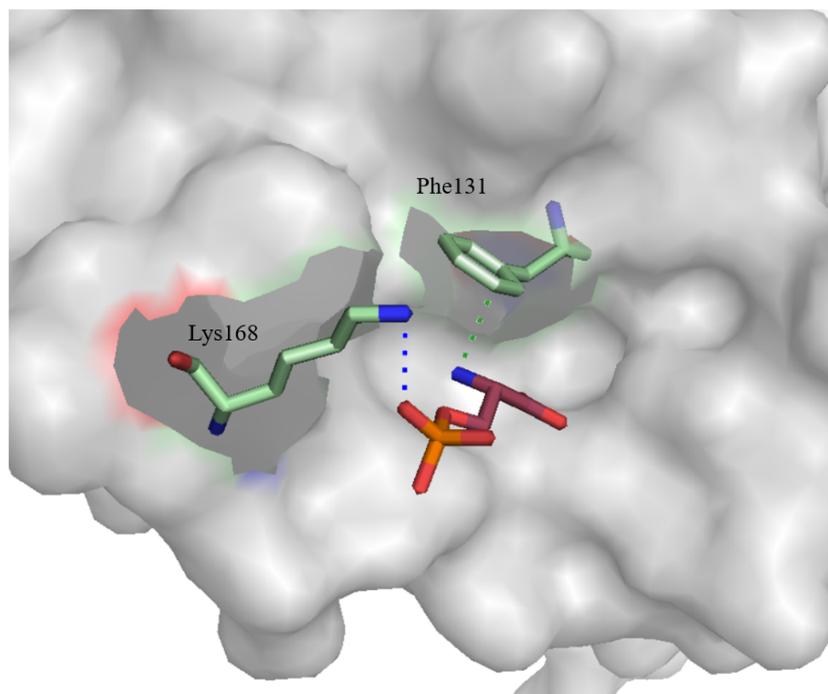


Figure 6.11 – Stabilisation de la phospho-L-sérine sur la surface de la molécule B. Vision en volume. Les interactions π -cation apparaissent en vert et les interactions ioniques en bleu.

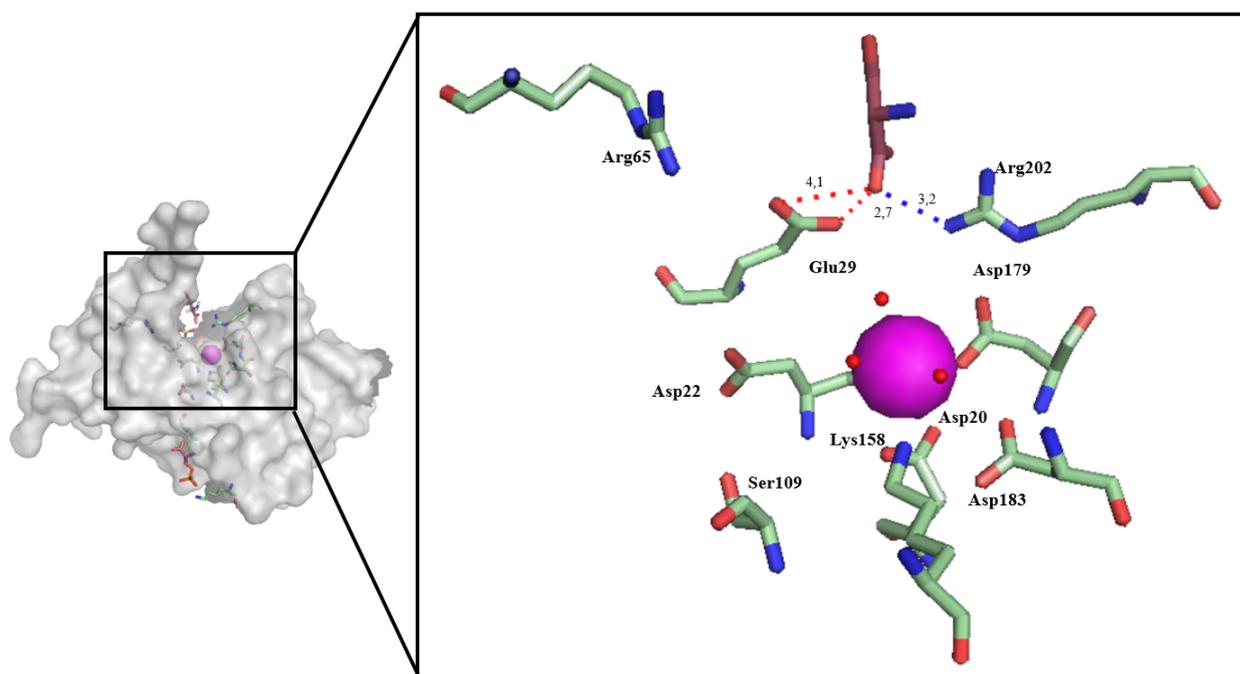


Figure 6.12 – Structure de la molécule B. Vision en volume. La L-sérine apparaît en bordeaux et le métal en rose. Les interactions ioniques apparaissent en bleu et les distances entre le ligand et le Glu29 apparaissent en rouge.

Tableau 6.4 – Interactions entre les acides aminés et les ligands contenu dans la molécule B, présenté à la Figure 6.12. Les atomes concernés par la mesure des distances sont soulignés.

| Ligands | Acide aminé | Distance (Å) | Type d'interactions |
|--------------------|------------------------|--------------|---------------------|
| PLS (<u>N</u> H3) | Phe 131 (<u>A</u> R.) | 3,0 | π -cation |
| PLS (<u>O</u> P) | Lys 168 (<u>N</u> H3) | 3,5 | Ionique |
| SER (<u>C</u> OO) | Glu 29 (<u>C</u> OO) | 2,7 | / |
| SER (<u>C</u> OO) | Glu 29 (<u>C</u> OO) | 4,1 | / |
| SER (<u>C</u> OO) | Arg 202 (<u>N</u> H2) | 3,2 | Ionique |

La position des ligands dans les deux molécules souligne l'importance de certains acides aminés dans le mécanisme de HPSP. En effet, le rôle de l'Asp20 et de l'Asp22 est cohérent avec le positionnement de ces derniers par rapport au substrat de l'enzyme. L'Asp20 est judicieusement orienté afin d'attaquer le phosphate tandis que l'Asp22 est proche du substrat. Une interaction entre ces deux derniers après hydrolyse est possible afin de permettre la protonation de la L-sérine. L'importance de l'Arg65, la Lys158 et de la Thr182 est soulignée afin de stabiliser le substrat dans la poche enzymatique. Ces trois résidus permettent l'orientation du substrat favorisant la réaction. Enfin, le Glu29 et l'Arg202 peuvent avoir un rôle de liaison du produit lors de la sortie du site actif.

6.3 Structure de HPSP en présence de L-sérine et de phosphate

Le second cristal analysé au synchrotron est représenté à la Figure 6.13. Ce cristal est obtenu par co-cristallisation en présence de phospho-L-sérine, à pH 6,5 et à une concentration en CaCl_2 de 0,05 M. Le cristal ainsi obtenu diffracte à une résolution atomique de 1,57 Å.

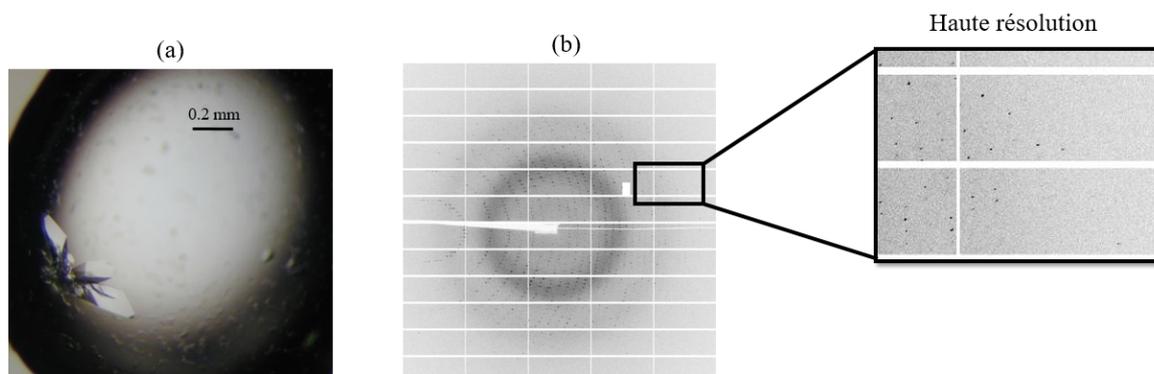


Figure 6.13 – (a) Cristaux obtenus par co-cristallisation avec la phospho-L-sérine, à pH 6,5 et à une concentration en CaCl_2 de 0,05 M dans un tampon cacodylate 0,1 M et de PEG 2000 20%. (b) Cliché de diffraction du cristal. Résolution de 1,57 Å.

Après affinement des données de diffraction, l'unité asymétrique possède également deux molécules, nommées C et D. De manière identique que précédemment, un ion métallique est retrouvé dans chaque site actif ainsi que la boucle dynamique. Celle-ci a pu être affinée dans son entièreté pour les deux molécules et possède le même dynamisme d'ouverture que précédemment (fermé dans C et ouvert dans D). Des ligands sont observables dans les sites actifs des deux molécules. Toutefois, il ne s'agit plus de phospho-L-sérine mais de L-sérine et d'un phosphate, les produits d'hydrolyse.

La structure, déposée dans la PDB sous le code 6HYY, est présentée à la Figure 6.14. Les informations relatives à la diffraction sont reprises au Tableau 6.5. On peut constater que la structure possède un R_{work} (19,45%) et un R_{free} (22,66%) relativement bas, ce qui montre une très bonne cohérence du modèle avec les valeurs expérimentales. Les valeurs basses de RMSD appuient cette constatation. De manière générale, la structure respecte les cartes de Ramachandran sauf dans le cas de deux acides aminés, le Glu224 de la molécule A et la Val40 de la molécule B. Cela s'explique par la présence respective de ces acides aminés en extrémité terminale et dans la boucle dynamique de HPSP. Il en découle une certaine mobilité. Il en est de même pour les rotamères. Ils sont tous cohérents avec la carte de densité sauf l'Arg202 de la molécule B, un acide aminé possédant un caractère dynamique détaillé plus loin.

Tableau 6.5 – Données relatives à la diffraction de la structure présentée à la Figure 6.14. Les données entre parenthèses correspondent à la plus grande sphère de résolution.

| Données de diffraction | | Affinement | |
|---------------------------------|--|---|-------|
| Groupe d'espace | $C222_1$ | R_{work} final (%) | 19,45 |
| Sous-unité | 2 | R_{free} final (%) | 22,56 |
| Paramètre de maille (Å et °) | a = 49,14, b = 129,14 | RMS(bonds) (Å) | 0,007 |
| | c = 155,79, $\alpha, \beta, \gamma = 90$ | RMS(angles) (Å) | 1,20 |
| Résolution (Å) | 32,28 - 1,57 | Facteur B (macromolécule) (Å ²) | 34,35 |
| | (1,622 - 1,566) | Nombre de molécules d'eau | 347 |
| Réflexions unique | 69622 (6498) | Nombre de ligands | 3 |
| Complétude (%) | 99,25 (93,82) | Rotamères défavorisés (%) | 0,27 |
| Signal sur bruit (I/σ) | 22,30 (2,07) | Ramachandran défavorisés (%) | 0,46 |
| Multiplicité | 2,0 (2,0) | Ramachandran favorisés (%) | 96,29 |

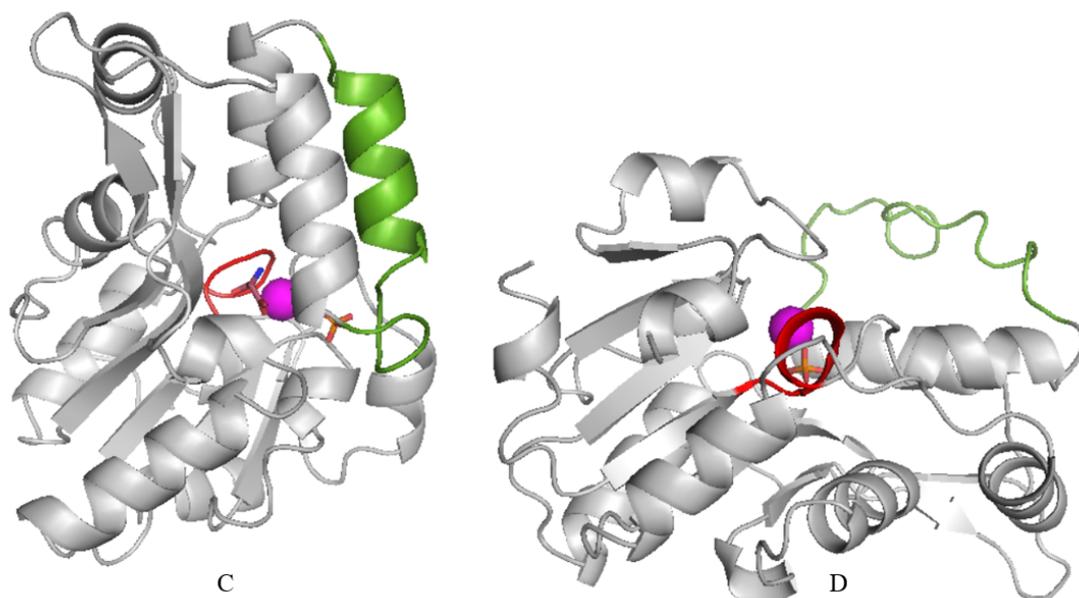


Figure 6.14 – Structure de HPSP obtenue à pH 6,5 et à une concentration en CaCl_2 de 0,05 M. Structure des monomères C et D. La boucle dynamique apparaît en vert, le phosphate en orange, le métal en mauve, le motif catalytique en rouge et la L-serine en bordeaux.

Afin de déterminer le métal dans le site actif, la même méthodologie que dans le cas de la structure précédente est appliquée. Les deux cartes de densités (Figures 6.15 et 6.16) ont mis en évidence deux sites actifs foncièrement différents. Dans le cas de la molécule C, il s'agit bien des six ligands habituels pour l'ion métallique de HPSP. Cependant, pour la molécule D, l'Asp20 ne coordonne plus le métal avec un seul oxygène du carboxylate mais avec les deux. Deux molécules d'eau sont visibles, une des trois molécules habituellement observée dans la structure ayant été remplacée par une liaison au phosphate, soit une géométrie à sept ligands. La carte $F_O - F_C$ indique un manque d'électrons dans la molécule D en présence de Mg^{2+} . Ces différences de géométrie suggèrent la présence de Mg^{2+} dans la molécule C et de Ca^{2+} dans la molécule D. Les facteurs d'agitations thermiques (facteurs B) et les longueurs de liaisons confirment cette hypothèse.

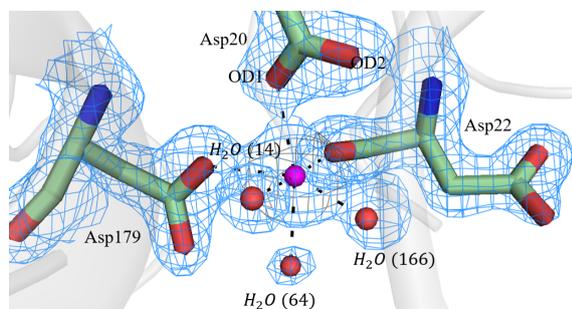


Figure 6.15 – Carte de densités du site actif de la molécule C en présence de Mg^{2+} . Facteur B de $17,25 \text{ \AA}^2$ pour le Mg^{2+} .

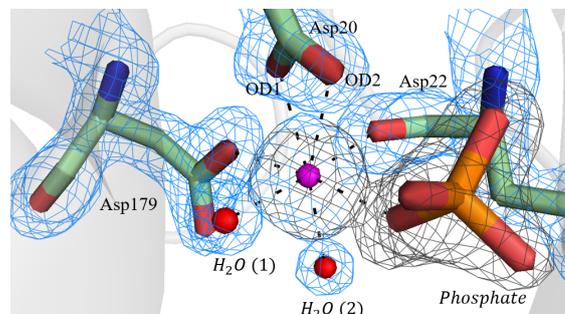


Figure 6.16 – Carte de densités du site actif de la molécule D en présence de Ca^{2+} . Facteurs B de $23,76 \text{ \AA}^2$ pour le Ca^{2+} .

La seconde information observable sur cette structure est la présence de ligands. Les sites actifs de la molécule C et de la molécule D contiennent des ligands différents. Dans la molécule C, on retrouve une molécule de L-sérine ainsi qu'un ion phosphate alors que dans la molécule D seulement un ion phosphate est observable. En ce qui concerne la molécule C, observable à la Figure 6.17, la fonction alcool de la L-sérine est orientée dans la poche catalytique vers l'Asp22. Cela pourrait souligner le rôle de l'Asp22 lors de la protonation de la L-sérine. Une stabilisation par la Ser109 et la Gly110 via un pont hydrogène est possible. L'interaction entre la Lys158 et le ligand est de type ionique et stabilise aussi la molécule de L-sérine. Le Glu29 et l'Arg65 jouent aussi un rôle stabilisant indirect à l'aide de différentes molécules d'eau, à savoir deux molécules d'eau pour l'Arg65 et une pour le Glu29. Le phosphate, lui, est stabilisé par l'Arg202. Deux conformations pour l'Arg202 sont observables en fonction de l'état de liaison du phosphate. La position de l'Arg202 liant le phosphate (Arg202 (B)) possède un facteur d'occupation de 40%. Le phosphate interagit avec la Thr182, la Lys158 et une molécule d'eau. Les ligands n'interagissent pas avec la Met52. Celle-ci possède une position équivalente à la position de la Met52(B) précédemment mise en évidence. La protéine possède une boucle dynamique fermée, confirmant le rôle de loquet de l'acide aminé. Les deux ligands sont aussi proches de l'Asp20, dont le rôle nucléophile a déjà été mentionné.

Tableau 6.6 – Interactions entre les acides aminés et les ligands présentés à la Figure 6.17. Les atomes concernés par la mesure des distances sont soulignés.

| Ligands | Acide aminé | Distance (Å) | Type d'interactions |
|--------------------|----------------------------|--------------|---------------------|
| SER (<u>COO</u>) | Asp 20 (<u>COO</u>) | 3,2 | / |
| SER (<u>NH3</u>) | Asp 22 (<u>COO</u>) | 2,5 | Pont hydrogène |
| SER (<u>OH</u>) | Asp 22 (<u>NH</u>) | 3,0 | Pont hydrogène |
| SER (<u>OH</u>) | Ser109 (<u>OH</u>) | 3,0 | Pont hydrogène |
| SER (<u>COO</u>) | Gly110 (<u>NH</u>) | 3,0 | Pont hydrogène |
| SER (<u>COO</u>) | Lys158 (<u>NH3</u>) | 2,4 | Ionique |
| SER (<u>NH3</u>) | Eau - Glu 29 (<u>O</u>) | 2,6 - 2,1 | Pont hydrogène |
| SER (<u>NH3</u>) | Eau - Arg65 (<u>O</u>) | 3,8 - 4,7 | Pont hydrogène |
| PO4 (<u>O</u>) | Eau - SER (<u>COO</u>) | 3,6 - 2,4 | Pont hydrogène |
| PO4 (<u>O</u>) | Thr182 (<u>OH</u>) | 2,7 | Pont hydrogène |
| PO4 (<u>O</u>) | Arg202 (A) (<u>NH2</u>) | 4,1 | Ionique |
| PO4 (<u>O</u>) | Arg 202 (B) (<u>NH2</u>) | 2,4 | Ionique |

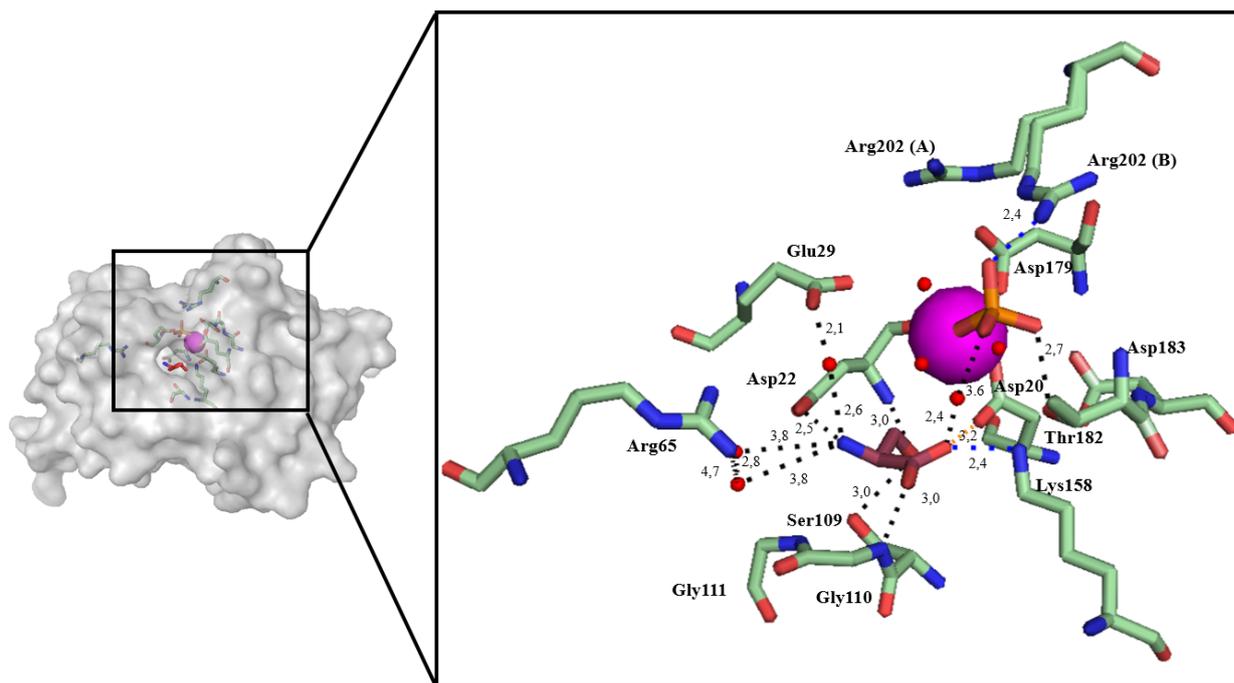


Figure 6.17 – Structure du site actif de la molécule C. Vision en volume. La L-sérine apparaît en bordeaux, le phosphate en orange et le métal en rose. Les interactions de type pont hydrogène sont en noir et les interactions ioniques en bleu. Les distances pouvant montrer une possible orientation due à une réactivité sont en orange.

D'un point de vue mécanistique, la molécule C est très intéressante. Elle confirme l'importance de divers acides aminés démontrée dans la littérature par mutagenèse (Asp20, Asp22, Asp179, Glu29, Arg65, Ser109, Lys158, Asp183, Arg202). Néanmoins, le rôle de certains acides aminés semble s'expliquer différemment que les hypothèses posées dans la littérature. En cohérence avec la littérature, cette structure confirmerait le rôle de l'Asp20 comme nucléophile. L'orientation de l'Asp22 indique également le rôle de ce dernier comme intervenant dans la protonation de la L-sérine. L'importance de la Ser109, du Glu29, de l'Arg202 et de la Lys158 est mise en évidence selon un rôle de stabilisation et d'orientation des ligands. Toutefois, la Lys158 pourrait jouer également un rôle lors de l'étape de régénération. L'Asp20 serait, dans ce cas, régénéré par cet acide aminé et non par l'Asp183. De plus, l'Asp183 est moins proche du phosphate pour permettre une régénération. Enfin, l'Asp179 ne montre pas d'influence dans la stabilisation des ligands.

La molécule D, quant à elle (Figure 6.18), possède un phosphate proche de l'Asp20 et de l'Asp22. La Ser109 et la Lys158 interviennent une nouvelle fois dans la stabilisation du ligand. Le phosphate interagit avec l'ion calcium. L'Arg65, Asp183 et l'Arg202 sont relativement loin du phosphate, aucune interaction n'est possible. La position de ces acides aminés souligne le rôle de la Lys158 dans l'hydrolyse de l'Asp20 lorsque celui-ci est lié au phosphate de par sa proximité. La proximité avec l'Asp22 peut s'expliquer par un pont hydrogène si cet acide aminé est protoné.

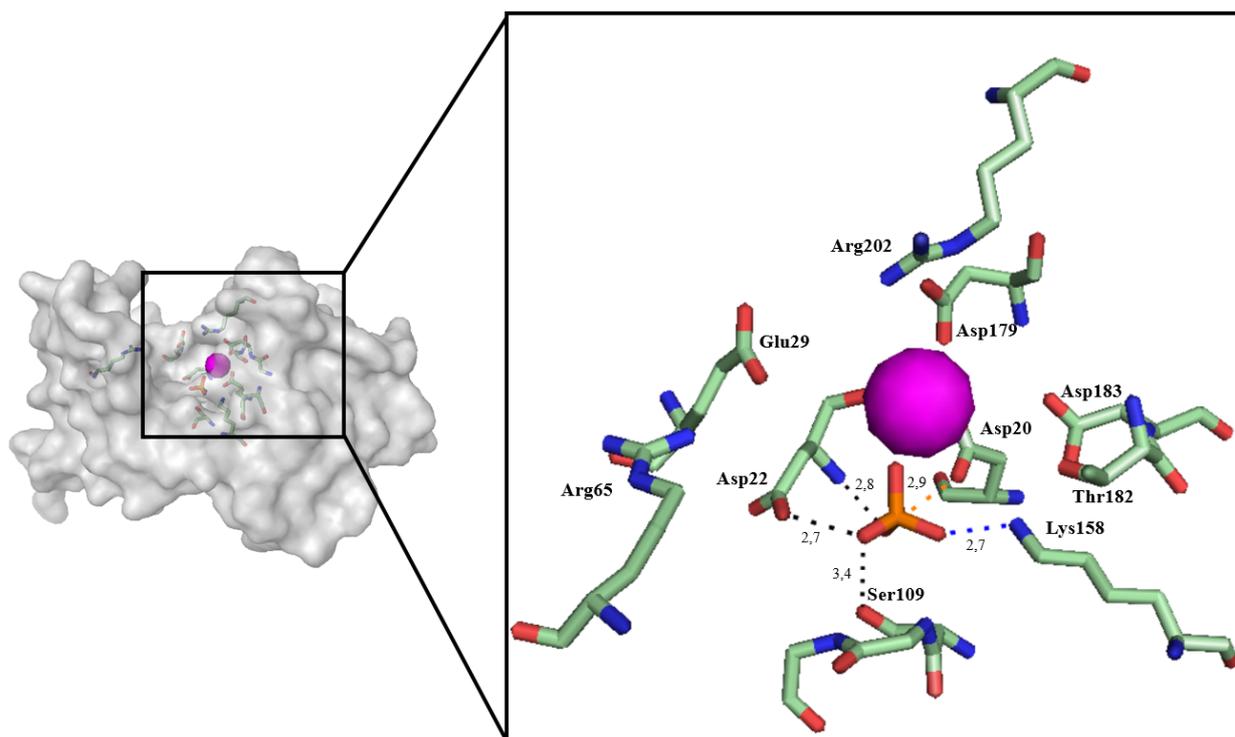


Figure 6.18 – Structure du site actif de la molécule D. Vision en volume. Le phosphate apparaît en orange et le métal en rose. Les interactions de type pont hydrogène sont en noir et les interactions ioniques en bleu. Les distances pouvant montrer une possible orientation due à une réactivité sont en orange.

Tableau 6.7 – Interaction entre les acides aminés et les ligands présentés à la Figure 6.18. Les atomes concernés par la mesure des distances sont soulignés.

| Ligands | Acide aminé | Distance (Å) | Type d'interactions |
|------------------|------------------------|--------------|---------------------|
| PO4 (<u>O</u>) | Asp 20 (<u>COO</u>) | 2,9 | / |
| PO4 (<u>O</u>) | Asp 22 (<u>COO</u>) | 2,7 | Pont hydrogène |
| PO4 (<u>O</u>) | Asp 22 (<u>NH</u>) | 2,8 | Pont hydrogène |
| PO4 (<u>O</u>) | Ser 109 (<u>OH</u>) | 3,4 | Pont hydrogène |
| PO4 (<u>O</u>) | Lys 158 (<u>NH3</u>) | 2,7 | Ionique |

6.4 Structure de HPSP en présence d'un analogue de substrat

Le dernier cristal analysé au synchrotron est présenté à la Figure 6.19. Il a été obtenu par co-cristallisation avec l'acide homocystéique, à un pH de 6,5 et à une concentration en CaCl_2 de 0,30 M. Le cristal ainsi obtenu a diffracté à une résolution atomique de 1,99 Å.

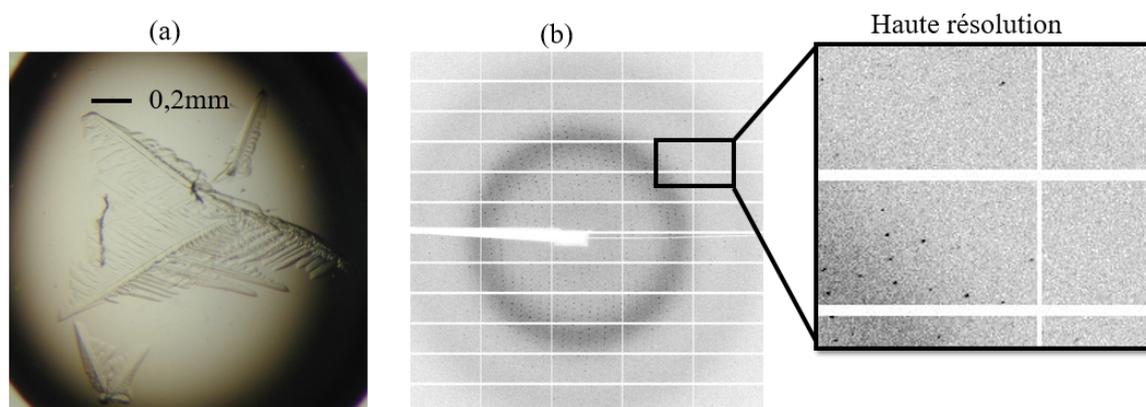


Figure 6.19 – (a) Cristaux obtenus par co-cristallisation avec l'acide homocystéique, à pH 6,5 et à une concentration en CaCl_2 de 0,30 M dans un tampon cacodylate 0,1 M, avec 8 mM d'acide homocystéique dilué dans le DMSO pour un pourcentage de 4% dans le puit et du PEG 2000 20%. (b) Cliché de diffraction correspondant. Résolution de 1,99 Å.

La structure obtenue, présentée à la Figure 6.20 et déposée dans la PDB selon le code 6Q6J, est fortement similaire aux structures précédentes. Elle possède deux molécules dans l'unité asymétrique (E et F) ainsi qu'un calcium dans chaque site actif. La boucle dynamique a également pu être définie dans son entièreté excepté la chaîne latérale de l'Arg50 de la molécule B. Dans le cas de la molécule E, la boucle est fermée tandis que dans le cas de la molécule F, elle est ouverte. Les données de diffractions sont présentées dans le Tableau 6.8. Les données de *R-factors* (19,82% et 24,76%) sont relativement cohérentes avec une résolution atomique de 1,99 Å. Un acide aminé se situe à l'extérieur des cartes de Ramachandran. Il s'agit de l'Ala55 de la molécule B, comprise dans la boucle dynamique. Au niveau des rotamères, un seul acide aminé est considéré comme non idéal. Il s'agit de la Leu220 de la molécule B. Cette structure se démarque dans les ligands qu'elle contient. Aucun ligand n'est observé dans le site actif de la molécule E. À l'inverse, une molécule d'acide homocystéique se trouve dans le site actif de la molécule F. À l'extérieur de la protéine, un ion calcium supplémentaire est observé proche de la molécule E et un ion chlorure près de la molécule F.

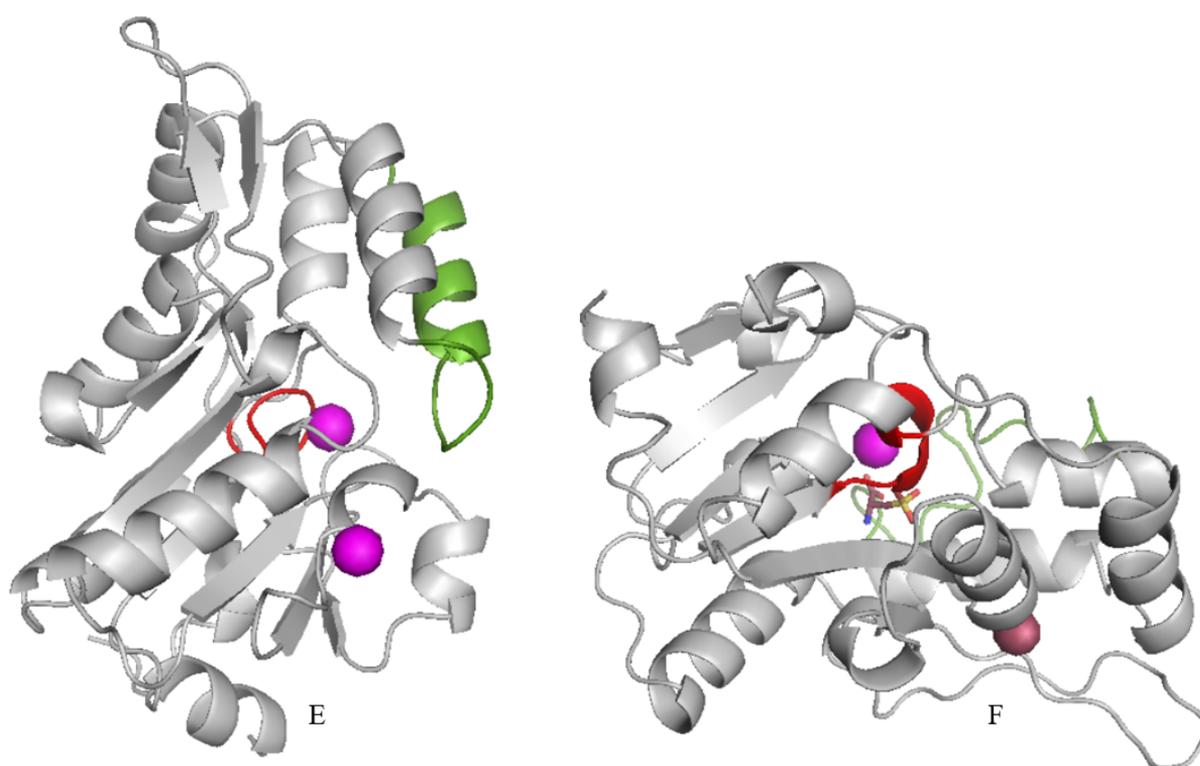


Figure 6.20 – Structure de HPSP obtenue pour des conditions de cristallisation avec l'acide homocystéique à pH 6,5 et à une concentration en CaCl_2 de 0,30 M. La boucle dynamique apparaît en vert, les ions calcium en mauve, le motif catalytique en rouge et l'acide homocystéique et l'ion chlorure en bordeaux.

Tableau 6.8 – Données relatives à la diffraction de la structure présentée à la Figure 6.20. Les données entre parenthèses correspondent à la plus grande sphère de résolution.

| Données de diffraction | | Affinement | |
|---------------------------------|--|---|-------|
| Groupe d'espace | $C222_1$ | R_{work} final (%) | 19,82 |
| Sous-unité | 2 | R_{free} final (%) | 24,76 |
| Paramètre de maille (Å et °) | a = 49,23, b = 130,49 | RMS(bonds) (Å) | 0,009 |
| | c = 157,76, $\alpha, \beta, \gamma = 90$ | RMS(angles) (Å) | 1,20 |
| Résolution (Å) | 34,65 - 1,985 | Facteur B (macromolécule) (Å ²) | 46,02 |
| | (2,056 - 1,985) | Nombre de molécules d'eau | 195 |
| Réflexions unique | 35482 (3380) | Nombre de ligands | 4 |
| Complétude (%) | 99,55 (96,24) | Rotamères défavorisés (%) | 0,28 |
| Signal sur bruit (I/σ) | 17,40 (2,29) | Ramachandran défavorisés (%) | 0,23 |
| Multiplicité | 2,0 (2,0) | Ramachandran favorisés (%) | 96,07 |

D'un point de vue mécanistique, la position de l'acide homocystéique dans la protéine est intéressante. Comme démontré à la Figure 6.21, ce ligand est en interaction avec de nombreux acides aminés. La partie sulfonate est à proximité de l'Asp22. Par cette proximité et le caractère acide fort de l'acide homocystéique, un hydrogène doit être porté par l'Asp22 afin d'expliquer une interaction favorable. Cet hydrogène permet la formation d'un pont hydrogène et confirme le rôle de protonation de l'Asp22. Un motif de type *oxyanion hole* est observé entre un oxygène du sulfonate, la Ser109, la Gly110 et la Gly111.^C La formation de cet intermédiaire améliore l'attaque nucléophile puisque l'oxygène se comporte comme un sulfonate ($S-O^-$) et non plus comme une sulfone ($S=O$).⁷⁴⁻⁷⁶ L'attaque nucléophile est ainsi favorisée par l'Asp20 afin de combler le manque de liaison au niveau du soufre. Ce mécanisme de stabilisation connu dans des cas de protéases et d'estérases n'avait jamais été mis en évidence dans le cas de HPSP.⁷⁴⁻⁷⁷

Un pont hydrogène avec Thr182 est observable pour stabiliser le carboxylate ainsi qu'une stabilisation indirecte avec l'Asp183 avec des molécules d'eau. La fonction ammonium est stabilisée par une succession de pont hydrogène par le biais de molécules d'eau. L'Asp20 est préférentiellement orienté pour attaquer le soufre. L'hélice dynamique ouverte dans cette structure s'explique par la présence d'acide homocystéique et non de phospho-L-sérine dans le site actif.

Tableau 6.9 – Interaction entre les acides aminés et le ligand présentés à la Figure 6.21. Les atomes concernés par la mesure des distances sont soulignés.

| Ligands | Acide aminé | Distance (Å) | Type d'interactions |
|---------------------|----------------------------|--------------|----------------------|
| HOM (<u>SO</u> 3) | Asp20 (<u>CO</u> O) | 3,6 | / |
| HOM (<u>S</u> OOO) | Asp22 (<u>CO</u> OH) | 2,3 | Pont hydrogène |
| HOM (S <u>OO</u> O) | Asp22 (<u>N</u> H) | 2,5 | Pont hydrogène |
| HOM (S <u>OO</u> O) | Ser109 (<u>O</u> H) | 1,9 | <i>Oxyanion hole</i> |
| HOM (S <u>OO</u> O) | Gly110 (<u>N</u> H) | 2,6 | <i>Oxyanion hole</i> |
| HOM (S <u>OO</u> O) | Gly111 (<u>N</u> H) | 2,3 | <i>Oxyanion hole</i> |
| HOM (<u>N</u> H3) | Eau - Lys158 (<u>N</u> H) | 3,6 - 3,6 | Pont hydrogène |
| HOM (<u>CO</u> O) | Thr182 (<u>O</u> H) | 3,4 | Pont hydrogène |
| HOM (<u>CO</u> O) | Eau - Thr182 (<u>N</u> H) | 2,3 - 3,4 | Pont hydrogène |

C. Un *oxyanion hole* est un trou où va venir se positionner un anion qui va être stabilisé par plusieurs ponts hydrogènes.

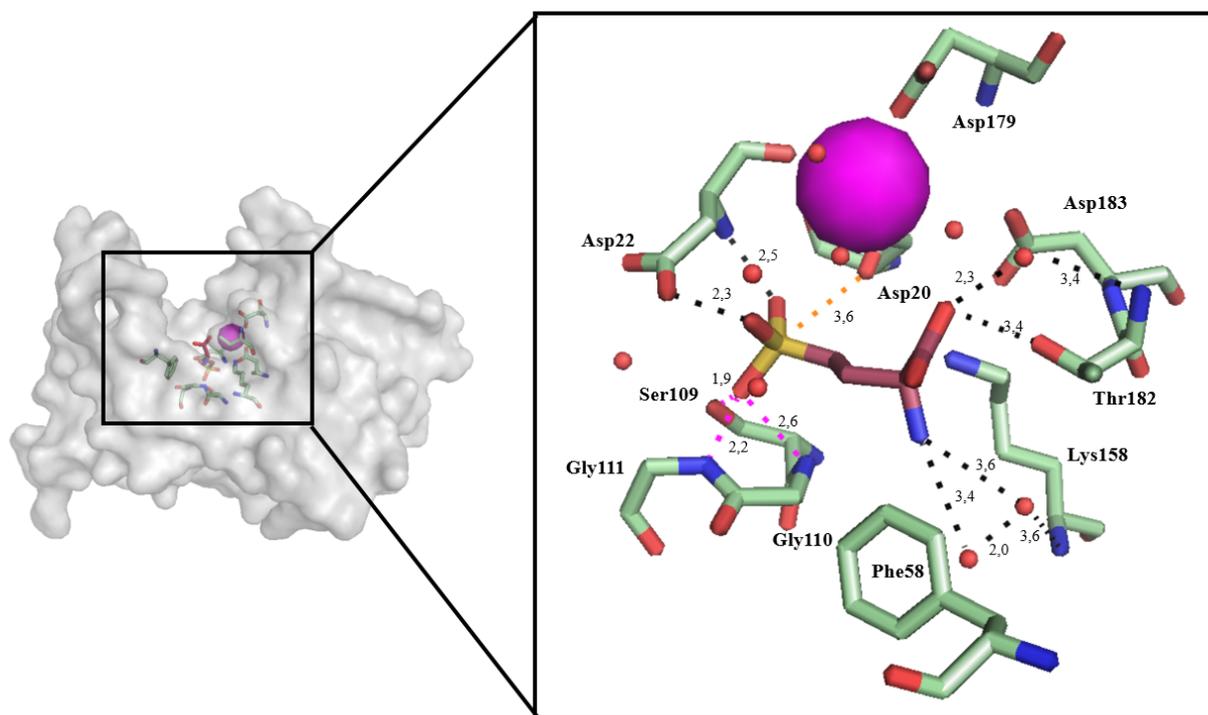


Figure 6.21 – Structure du site actif de la molécule F. Vision en volume. Les interactions ioniques sont en bleu, les interactions de type *Oxyanion hole* en rose et les ponts hydrogènes en noir. Les distances montrant une proximité pouvant être expliquées par une réactivité apparaissent en orange.

L'acide homocystéique se trouve à une position différente que la phospho-L-sérine obtenue dans la première structure. Cette dernière n'était pas stabilisée par la Ser109 et les deux glycines 110 et 111. La différence entre les deux positions est montrée à la Figure 6.22. Un déplacement vers l'intérieur du site actif est identifiable par rapport à la phospho-L-sérine. Cette différence de position peut s'expliquer du fait que l'acide homocystéique éjecte une molécule d'eau pour prendre sa place et s'approche des aspartates catalytiques. Ce rapprochement permet la stabilisation au niveau de l'*oxyanion hole* et facilite l'attaque nucléophile. En présence de phospho-L-sérine, une approche du substrat similaire peut être envisagée au niveau du mécanisme.

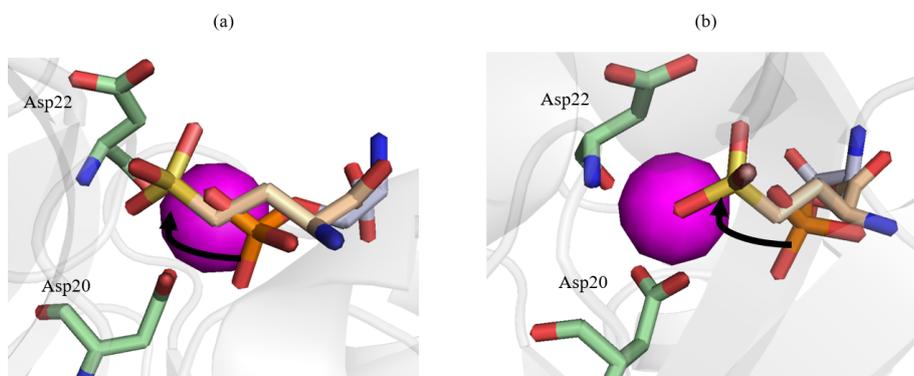


Figure 6.22 – Positionnement de l'acide homocystéique (mauve) dans le site actif de la molécule F. Comparaison à la phospho-L-sérine (beige) dans la molécule A (a) Vue de face (b) Vue de profil.

6.5 Étude structurale de la boucle dynamique de HPSP

Une autre information importante des structures obtenues est la localisation de la boucle dynamique. Cette boucle, composée des résidus compris entre la Val40 et la Pro57, adopte plusieurs conformations en fonction de l'état d'occupation du site actif par des ligands. Deux conformations peuvent être observées, à savoir une conformation ouverte et une conformation fermée observables à la Figure 6.23. Celles-ci sont développées ci-dessous pour les molécules A, B, C et D. Les molécules avec l'analogue non hydrolysable ont été exclues afin d'assurer une conformation représentative d'une protéine active.

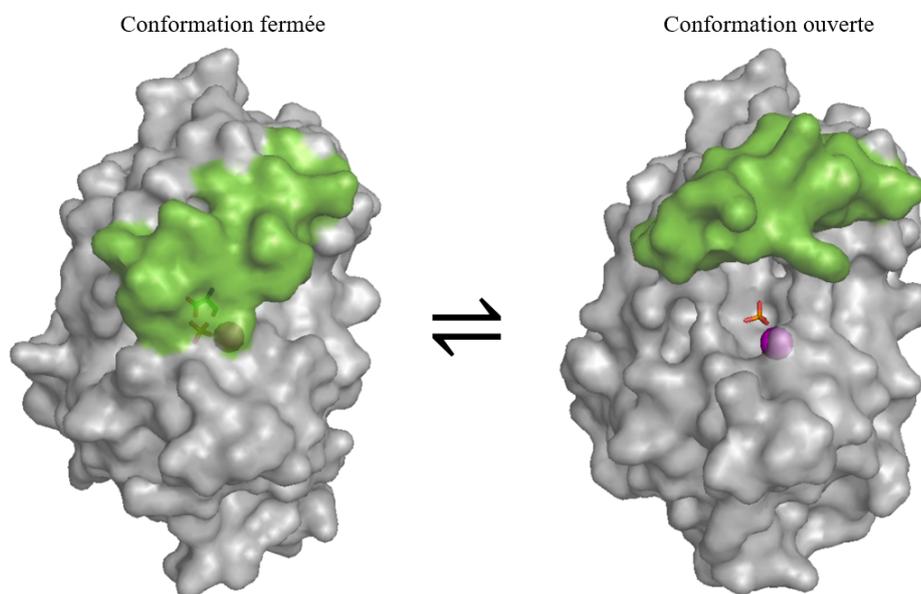


Figure 6.23 – Changement de conformation entre la forme fermée et la forme ouverte de HPSP, visualisé à l'aide des molécules C et D.

La conformation fermée, observée dans le cas de la molécule A et de la molécule C, est reprise à la Figure 6.24, a. Cette conformation possède la particularité d'adopter une forme d'hélice α , semblable à celle observée dans la PDB.⁵⁰ La structure de cette hélice est fortement conservée entre les différentes structures, malgré de légères différences au niveau de la conformation de la boucle terminant l'hélice, mise en évidence à la Figure 6.24 par un rond noir.

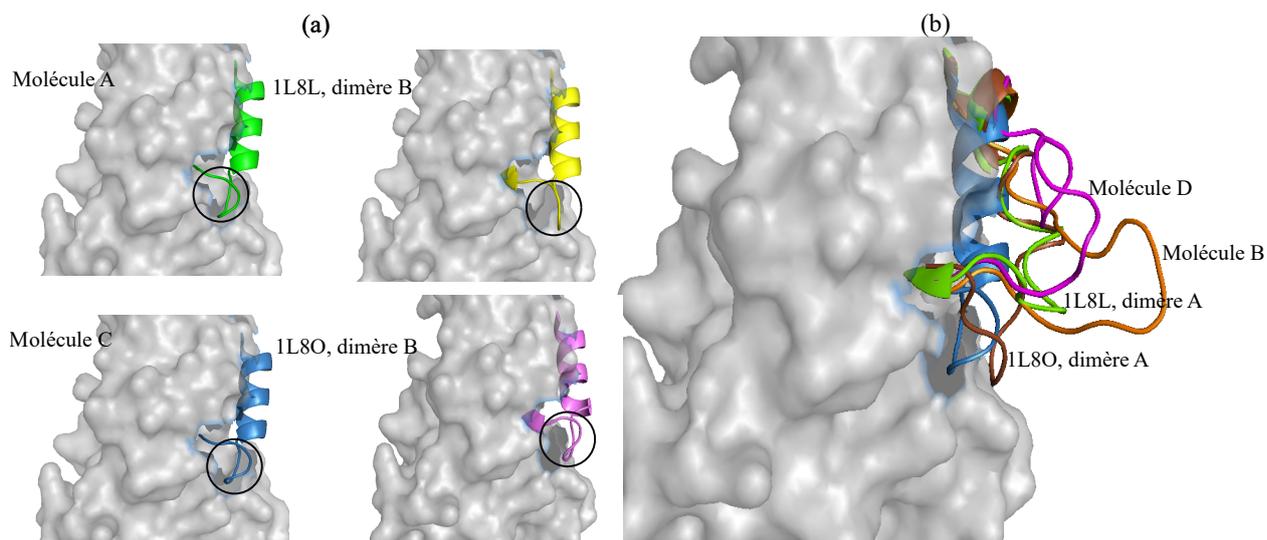


Figure 6.24 – Comparaison de la conformation de la boucle dynamique dans l'enzyme en fonction des molécules. Comparaison des structures observables dans la littérature (1L8L et 1L8O⁵⁰). (a) Conformation fermée (b) Conformation ouverte.

La conformation ouverte, reprise pour les molécules B et D à la Figure 6.24, b, correspond à une configuration dépliée de la boucle dynamique, admettant beaucoup de degrés de liberté. De ce fait, plusieurs conformations différentes sont obtenues. La Figure 6.24 souligne les différentes structures observées pour cette boucle dynamique, que ce soit celle des molécules B et D ou encore de HPSP dans la PDB.⁵⁰ La juxtaposition de ces conformations donne accès au caractère dynamique de cette boucle. En partant de la conformation en hélice α de la boucle, la conformation présentée dans la structure du monomère A reprise dans la PDB par Kim *et al.* de la structure 1L8L souligne une légère conformation en hélice qui tend à être perdue progressivement dans le monomère A de la structure 1L8O jusqu'à perdre totalement cette conformation dans la molécule B et dans la D. Lorsque la conformation est ouverte, comme c'est le cas dans la molécule B, la phospho-L-sérine peut entrer dans le site actif de HPSP. Une fois le substrat dans la poche catalytique, la boucle dynamique change de conformation et adopte la conformation fermée. Au moment de cette fermeture, la Met52 change de position, à l'image d'un loquet.

6.6 Hypothèse mécanistique

L'apport de ces structures cristallographique est important afin de comprendre au mieux le mécanisme mis en œuvre par l'enzyme. Le rôle d'une multitude d'acides aminés a pu être mis en évidence par le biais de la cristallographie. Suite à ces informations, un mécanisme peut donc être

suggéré. Ce dernier est élaboré sur base d'une vision globale de la position des ligands dans le site actif, disponible à la Figure 6.25 (en trois dimensions) et à la Figure 6.26 pour une vision des interactions (à deux dimensions).

Au point de départ, l'enzyme est sous sa forme ouverte, avec une structure de boucle dynamique dépliée. Cela laisse le site actif accessible afin que la phospho-L-sérine puisse entrer à l'intérieur du site actif (Figure 6.25 et 6.26). Une fois la phospho-L-sérine entrée, une interaction de type pont hydrogène avec l'Asp22, une interaction de type ionique avec la Lys158 et une interaction de type pont hydrogène avec la Thr182 orientent convenablement le substrat (Figure 6.25 et Figure 6.26, I). La boucle dynamique se replie et la Met52 change de conformation pour fermer l'hélice. La phospho-L-sérine avance dès lors légèrement pour prendre une position comparable à l'acide homocystéique afin de se rapprocher des aspartates catalytiques. Cela provoque le départ d'une molécule d'eau. S'il s'agit de phospho-L-sérine dans le site actif, la boucle dynamique reste en position fermée. Le substrat est stabilisé par la Ser109, la Gly110 et la Gly111 au sein d'un *oxy-anion hole* et par la Thr182 par pont hydrogène (Figure 6.25 et Figure 6.26, II). Une fois le substrat positionné, la proximité entre la phospho-L-sérine et l'Asp20 permet l'attaque par cet acide aminé sur le phosphate afin de permettre l'hydrolyse du substrat. Une molécule de L-sérine déprotonée et d'un Asp20 phosphorylé se trouve alors dans le site actif.

Du fait de la liaison du phosphate à l'Asp20, un espace se crée entre l'Asp22 et la L-sérine déprotonée. La L-sérine s'approche de cet acide aminé par des interactions avec la Ser109 et la Gly110, permettant sa protonation (Figure 6.25 et Figure 6.26, III). Une fois cette L-sérine protonée, elle sort du site actif. Pour ce faire, elle est d'abord stabilisée à l'aide du Glu29, de l'Arg65, de la Ser109 et de la Lys158 (Figure 6.25 et Figure 6.26, III). Ensuite, la boucle dynamique va se déplier de façon à ouvrir la poche catalytique. Le Glu29 change légèrement de configuration éjectant la L-sérine de l'enzyme à l'aide d'interactions supplémentaires avec l'Arg202 jouant le rôle de pivot (Figure 6.25 et Figure 6.26, IV).

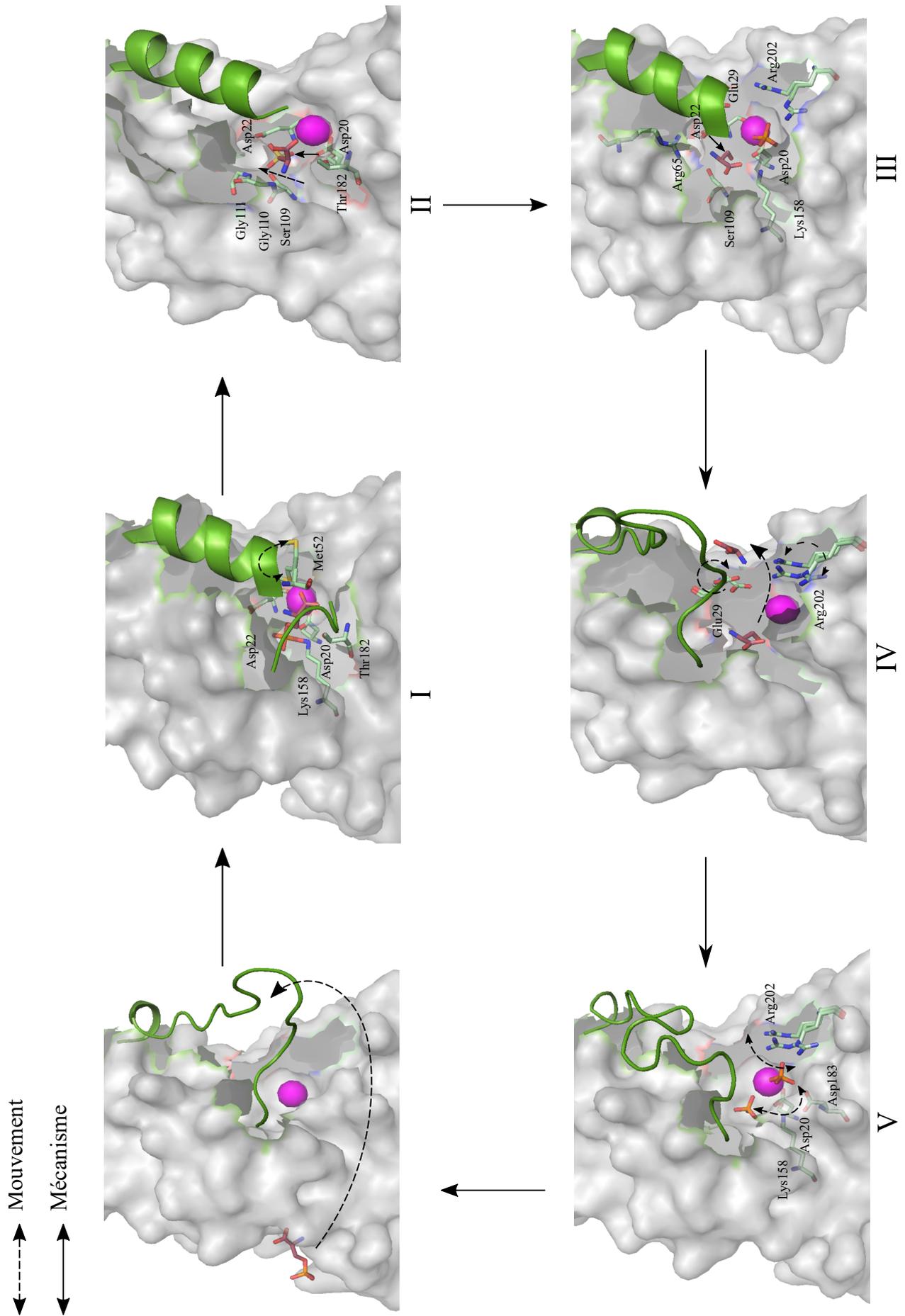


Figure 6.25 – Représentation des différentes conformations des ligands dans la protéine obtenues à partir des structures résolues par diffraction. Visions en 3D.

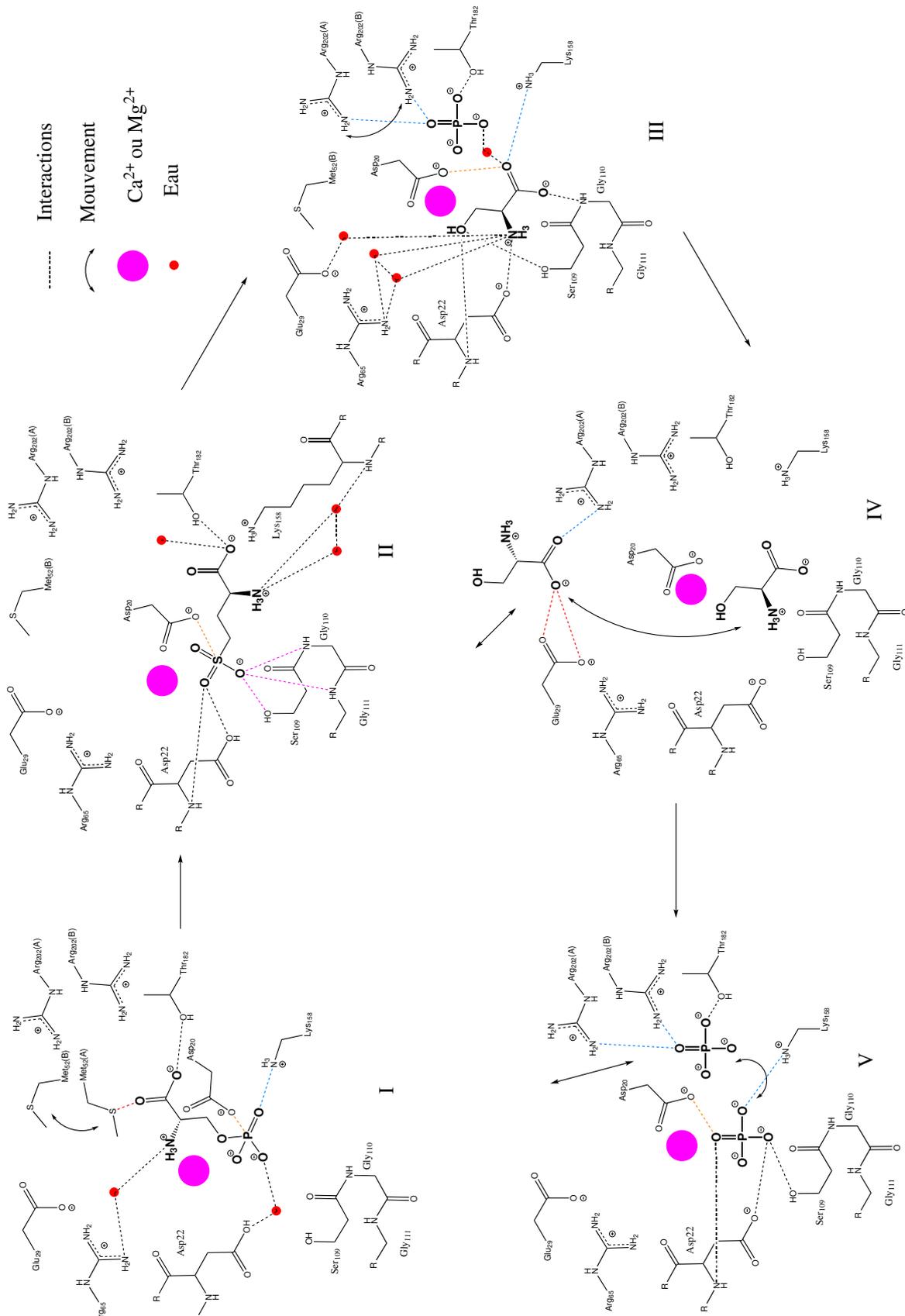


Figure 6.26 – Représentation des différentes conformations des ligands dans la protéine obtenues à partir des structures résolues par diffraction. Vision en 2D. Les interactions ioniques sont en bleu, les interactions au sein de l'*Oxyanion hole* en rose et les ponts hydrogènes en noir. Les distances défavorables sont en rouge tandis que les distances pouvant montrer une possibilité de réactivité sont en orange.

D'autre part, l'Asp20 est régénéré afin de libérer le phosphate (Figure 6.25 et Figure 6.26, V). Les structures cristallographiques obtenues laissent suggérer deux types de mécanismes en fonction de la position du phosphate lors de cette régénération. Cela peut être réalisé soit au moyen de l'Asp183, comme le décrit la littérature, ou de la Lys158, comme laisse suggérer les structures cristallographiques. Dans le premier cas, l'Asp183 active une molécule d'eau, en la déprotonant. Cette molécule d'eau attaque le phosphate porté par l'Asp20 par une attaque nucléophile. Dans le second cas, la Lys158 joue le rôle d'activateur d'eau pouvant attaquer de manière similaire l'Asp20. Après régénération, le phosphate se coordonne à l'Arg202 jouant le rôle de pivot pour le faire sortir de l'enzyme. L'implication des différents acides aminés, en comparaison aux informations fournies par la littérature,^{46,50} est reprise au Tableau 6.10.

Tableau 6.10 – Récapitulatif des acides aminés importants dans le mécanisme de HPSP selon la littérature. Apport de ce travail dans l'hypothèse mécanistique. PLS = phospho-L-sérine.^{46,50}

| AA | Peeraer <i>et al.</i> | Kim <i>et al.</i> | Ce travail |
|--------|---|--|-------------------------------|
| Asp20 | Nucléophilie | Nucléophilie | Nucléophile |
| Asp22 | Protonation | Stabilisation PLS | Protonation/Stabilisation |
| Glu29 | Stabilisation L-sérine | Stabilisation PLS | Stabilisation/sortie L-sérine |
| Met52 | / | / | Loquet |
| Arg65 | Stabilisation L-sérine | / | Stabilisation L-sérine |
| Ser109 | / | Stabilisation PO ₄ ³⁻ | Stabilisation PLS/L-sérine |
| Gly110 | Stabilisation PO ₄ ³⁻ | Stabilisation PLS | Stabilisation PLS/L-sérine |
| Gly111 | / | / | Stabilisation PLS |
| Lys158 | Stabilisation PO ₄ ³⁻ | Stabilisation PLS/PO ₄ ³⁻ | Stabilisation/ Régénération |
| Asp179 | / | Stabilisation PLS | Structural |
| Ala181 | / | Stabilisation L-sérine | / |
| Thr182 | / | Stabilisation PLS/L-sérine | Stabilisation PLS |
| Asp183 | / | Stabilisation du PLS | Structural/Régénération |
| Arg202 | / | Stabilisation L-sérine/PO ₄ ³⁻ | Stabilisation produits/Pivot |

6.7 Conclusions intermédiaires

Les cristaux obtenus ont été diffractés au synchrotron. Trois structures protéiques sont obtenues. La première structure possède deux molécules dans son unité asymétrique, nommées respectivement A et B. La molécule A possède un calcium dans son site actif ainsi qu'une phospho-L-sérine. La molécule B possède dans son site actif une L-sérine et également un calcium comme

métal. Une phospho-L-sérine est aussi observable à l'extérieur de l'enzyme. La seconde structure est composée d'un dimère, la molécule C et la molécule D. Le site actif de la molécule C comprend un magnésium, une L-sérine et un phosphate tandis que la molécule D possède un ion calcium et un phosphate. Dernièrement, une troisième structure avec un dimère (molécule E et F) composé d'un calcium dans chaque site actif est obtenue. La molécule F possède un acide homocystéique dans son site actif. Ces six molécules possèdent également une hélice dynamique. Les molécules A, C et E se trouvent en conformation fermée, sous forme d'hélice α . À l'inverse, les molécules B, D et F se trouvent en conformation ouverte, totalement dépliée. Ces différentes conformations permettent aux acides aminés composant la boucle dynamique d'interagir avec les ligands et d'enclencher la réaction. La Met52 joue le rôle de loquet pour fermer l'hélice.

Une hypothèse mécanistique est envisagée. Au départ de la réaction, la phospho-L-sérine est à l'extérieur de l'enzyme et la boucle dynamique est ouverte. Le substrat entre à l'intérieur de l'enzyme, provoquant le changement de conformation de la boucle dynamique. L'Asp22, la Lys158 et le Thr182 stabilisent le substrat. La Met52 change de conformation et ferme l'enzyme. Le substrat bouge légèrement et s'approche des aspartates catalytiques. La stabilisation a dès lors lieu grâce à la Ser109, les Gly110 et 111 et la Thr182. Un motif *oxyanion hole* est observé pour la stabilisation par la Ser109, la Gly110 et la Gly111. L'Asp20 attaque la phospho-L-sérine et l'hydrolyse a lieu, générant un phosphate lié à l'Asp20 et une L-sérine déprotonée dans la poche catalytique. Par la liaison du phosphate à l'Asp20, la L-sérine déprotonée se positionne à proximité de l'Asp22. Ce dernier protone la L-sérine. L'hydrolyse du phosphate par une molécule d'eau activée est réalisée au moyen de la Lys158 ou de l'Asp183. Par changement de conformations de l'hélice dynamique, les deux produits sortent de l'enzyme au moyen de l'Arg202 et, pour la L-sérine, d'une interaction défavorable du Glu29.

Troisième partie

Approche computationnelle

Chapitre 7

Étude computationnelle du mécanisme

Afin d'appuyer l'hypothèse mécanistique précédemment décrite, des analyses computationnelles sont effectuées. Celles-ci ont pour but la détermination des quantités thermodynamiques et cinétiques impliquées dans les réactions menant à l'hydrolyse de la phospho-L-sérine par utilisation de la DFT, et plus particulièrement au moyen de la fonctionnelle M06, et de la base 6-311G+(d,p). Cette détermination se fait dans l'approximation de modèles simplifiés dans l'optique de mettre en évidence l'influence des acides aminés autour des réactifs et produits. Dans ce présent chapitre, les trois réactions mentionnées dans le mécanisme (attaque de l'Asp20 sur le phosphate, protonation de la L-sérine générée et hydrolyse du phosphate de l'Asp20 phosphorylé, aussi appelée régénération) sont étudiées. Les différents états de transition de ces réactions, pour des modèles de complexités croissantes, sont déterminés et optimisés. À partir de ces états de transition, les complexes intermédiaires sont ensuite obtenus pour les réactifs et les produits. Ces complexes modélisent le rapprochement des différents réactifs ou produits dans l'espace et la stabilisation de ces derniers avant réaction. Les notions théoriques abordées lors de cette analyse computationnelle sont reprises dans la partie V.

7.1 Réaction d'attaque nucléophile par l'Asp20

Les structures cristallographiques suggèrent que la première étape du mécanisme, à savoir l'attaque nucléophile sur le phosphate, s'effectue à l'aide de l'Asp20. Afin de légitimer cette hypothèse, la réaction, décrite à la Figure 7.1, est analysée.

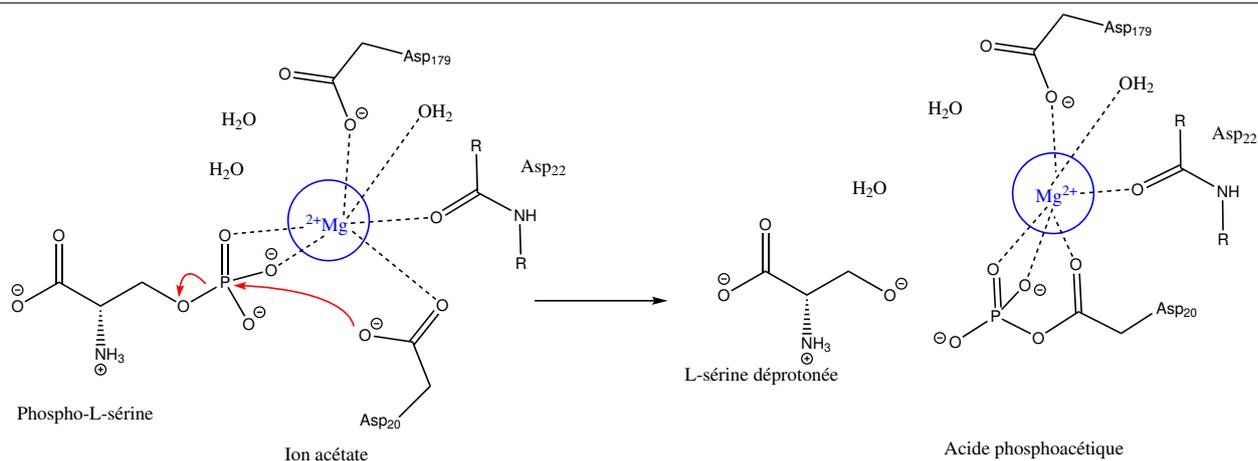


Figure 7.1 – Première étape de réaction du système enzymatique. Hydrolyse de la phospho-L-sérine (PLS^{2-}) en L-sérine déprotonée, dans l'eau, à pH physiologique. Mécanisme impliquant une attaque par l'Asp20. Des ions acétates remplacent les aspartates durant la simulation.

7.1.1 Aspects structuraux

Lorsque la réaction est modélisée au moyen d'un ion acétate (représentant l'Asp20) et d'une phospho-L-sérine pour donner une L-sérine déprotonée et un ion phosphoacétate, les analyses ont souligné l'incapacité d'isoler un état de transition, le système étant trop négatif. Effectivement, suite à l'excès de charges négatives des réactifs et des produits, il est impossible d'optimiser un état de transition cohérent et représentatif de la réalité avec seulement ces quatre composés. Les répulsions étant trop grandes, les espèces se placent de manière à éloigner au maximum les charges négatives. Il en résulte une mauvaise orientation des parties réactionnelles, comme montrée à la Figure 7.2 et à la Figure 7.3. Face à ce problème, d'autres molécules ont été ajoutées au système afin de permettre une orientation cohérente des réactifs et des produits avec les données structurales.

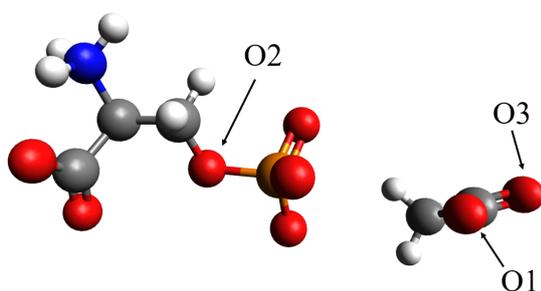


Figure 7.2 – Structure optimisée du complexe formé entre la phospho-L-sérine et l'ion acétate modélisant l'Asp20, calculée avec la méthode M06/6-311+G(d,p)/IEF-PCM.

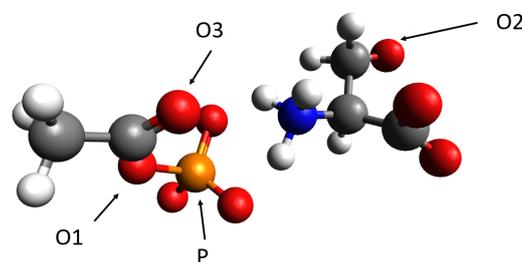


Figure 7.3 – Structure optimisée du complexe formé entre la L-sérine et l'ion phosphoacétate modélisant l'Asp20 phosphorylé, calculée avec la méthode M06/6-311+G(d,p)/IEF-PCM.

Des modèles plus complexes sont donc envisagés : trois modèles sont ainsi conçus et trois états de transition ont pu être isolés. Ceux-ci sont présentés à la Figure 7.4. Le premier modèle (modèle **a**) ajoute un magnésium au système afin d'écranter les charges négatives. L'ajout de cet ion permet d'isoler l'état de transition **1a** possédant une fréquence imaginaire de $141,0i \text{ cm}^{-1}$. Deuxièmement, un modèle comprenant un magnésium et un second aspartate est étudié (modèle **b**). Cet aspartate est crucial dans le site actif, car, comme il a pu être démontré dans la partie cristallographique, il possède un rôle structural dans l'enzyme en complexant l'ion Mg^{2+} . L'état de transition **1b** mis en évidence pour ce modèle possède une fréquence imaginaire de $150,6i \text{ cm}^{-1}$. Enfin, un modèle avec une sphère de coordination complète pour le magnésium est construit par ajout de trois molécules d'eau ainsi qu'un N-méthylacétamide, afin de modéliser la chaîne principale de l'Asp22 (modèle **c**). Il s'agit d'un modèle simplifié de la situation protéique, sans la chaîne latérale de l'Asp22. Ce modèle de départ est repris à la Figure 7.5 (**1c1**). Toutefois, une étude préliminaire du système a mis en évidence la nécessité de retirer deux molécules d'eau pour que la phospho-L-sérine puisse approcher l'acétate modélisant l'Asp20 et enclencher la réaction. Le coût de ce retrait d'un point de vue thermodynamique est détaillé au point suivant. Un état de transition avec une seule molécule d'eau est donc obtenu (**1c2**) et présenté à la Figure 7.4. Cet état de transition possède une coordination à six ligands, soit un oxygène de l'acétate modélisant l'Asp20 (O_3), les deux oxygènes de l'acétate modélisant l'Asp179 (O_4 et O_5), le carbonyle de l'Asp22 (O_6) et deux atomes d'oxygène du phosphate. La molécule d'eau fait un pont hydrogène avec l'oxygène (O_2) de la L-sérine déprotonée. Cette double coordination pour l'Asp179 et la position de la molécule d'eau ne respectent pas les données cristallographiques. Toutefois, la simplification du système et l'absence de la chaîne latérale de l'Asp22 peut en être la cause. La fréquence imaginaire s'élève à $217,6i \text{ cm}^{-1}$. Les trois fréquences imaginaires de ces trois états de transition mettent en évidence l'élongation des liaisons O_1-P et O_2-P , montrant une rupture de la liaison O_2-P et une formation de la liaison O_1-P . Le Tableau 7.1 permet une comparaison de ces distances en fonction du modèle. En ce qui concerne les liaisons O_1-P , une similarité est observée dans les deux premiers modèles (modèle **a** et modèle **b**). Cette liaison est légèrement affaiblie dans le cas de la sphère complète, permettant une hydrolyse plus aisée de l'Asp20. En ce qui concerne la liaison O_2-P , les liaisons obtenues pour les deux premiers modèles sont identiques tandis que la liaison dans le troisième modèle est légèrement plus forte.

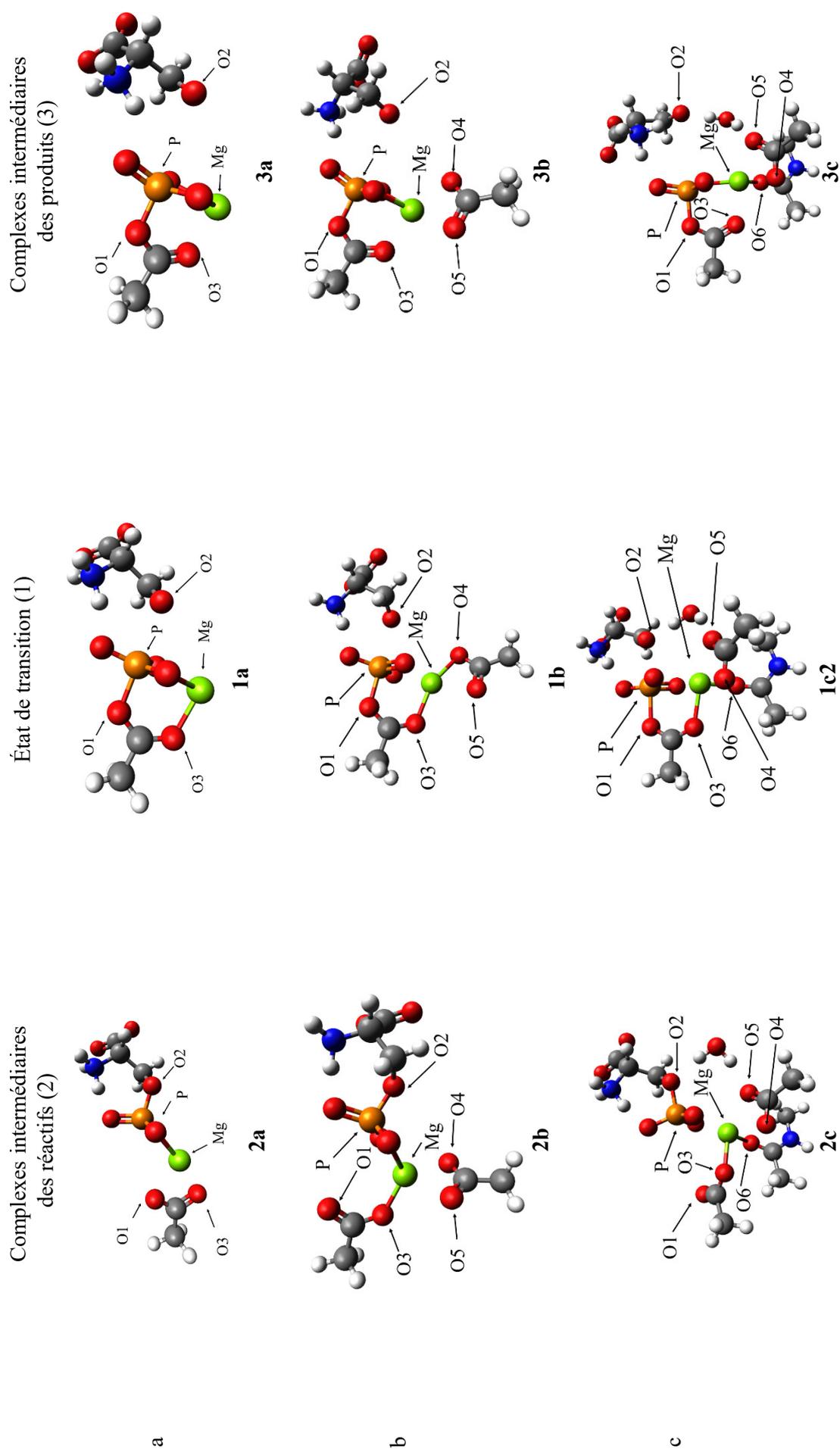


Figure 7.4 – Structures optimisées des complexes intermédiaires et des états de transition formés entre la L-sérine déprotonée et l'ion phosphoacétate modélisant l'Asp20 phosphorylé en fonction de différents modèles étudiés (**a**, **b** ou **c**), calculées avec la méthode M06/6-311+G(d,p)/IEF-PCM.

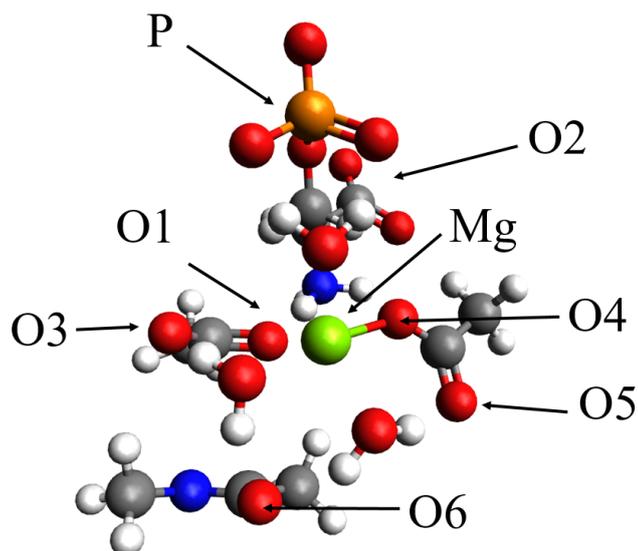


Figure 7.5 – Structure optimisée du complexe formé entre la phospho-L-sérine, l'ion acétate modélisant l'Asp20 en présence de Mg^{2+} entouré d'une sphère de coordination complète (trois molécules d'eau, un ion acétate modélisant l'Asp179 et de N-méthylacétamide modélisant l'Asp22 en plus de l'Asp20) (**1c1**), calculée avec la méthode M06/6-311+G(d,p)/IEF-PCM.

Grâce à ces états de transition, les complexes intermédiaires, montrant respectivement le rapprochement des réactifs ou des produits dans l'espace, sont identifiés dans le cas des trois modèles. Ces complexes, pour les réactifs, sont repris à la Figure 7.4 (**2a**, **2b** et **2c**). De manière cohérente, dans les trois modèles, le Tableau 7.1 souligne une forte élongation pour les liaisons O_1-P et O_3-P par rapport aux états de transition correspondants, puisque le phosphate est dans ce cas lié à la phospho-L-sérine comme le montre la liaison O_2-P légèrement plus forte que dans le cas d'une phospho-L-sérine seule. Cette élongation de O_1-P est d'autant plus forte dans le modèle de la sphère complète. Ce phénomène s'explique par l'encombrement stérique. Les complexes obtenus pour les produits sont repris à la Figure 7.4 (**3a**, **3b** et **3c**) pour les trois modèles. Dans le cas du produit, selon le Tableau 7.1, lorsqu'il y a formation du complexe intermédiaire, une élongation de la liaison O_2-P montre un affaiblissement de la liaison au profit des liaisons O_1-P et O_3-P , indiquant le transfert du phosphate à l'acétate. Les distances de liaisons O_1-P et O_3-P sont semblables dans le cas présent. La distance O_2-P est plus grande dans le troisième modèle, soulignant un éloignement de la L-sérine déprotonée par rapport à l'Asp20 après réaction.

D'un point de vue général, comme le souligne le Tableau 7.1, les liaisons C–O₁ et C–O₃ varient peu entre les différents modèles. Cependant, une variation de longueur de liaison est remarquée entre les réactifs, les produits et les états de transition. En ce qui concerne la liaison C–O₁, la variation des distances entre les composés indique un affaiblissement de la liaison au cours de la réaction, dû à une liaison du phosphate et donc à un appauvrissement électronique. À l'inverse, la liaison C–O₃ est renforcée au cours de la réaction pour atteindre une valeur proche du phosphoacétate isolé.

Tableau 7.1 – Distances interatomiques optimisées (Å) pour les réactifs de départ, les produits, les états de transition et les complexes intermédiaires correspondants. PLS²⁻ est la phospho-L-sérine, PAc²⁻ est le phosphoacétate et Ac⁻ l'acétate. Le C est le carbone portant O₁ et O₃.

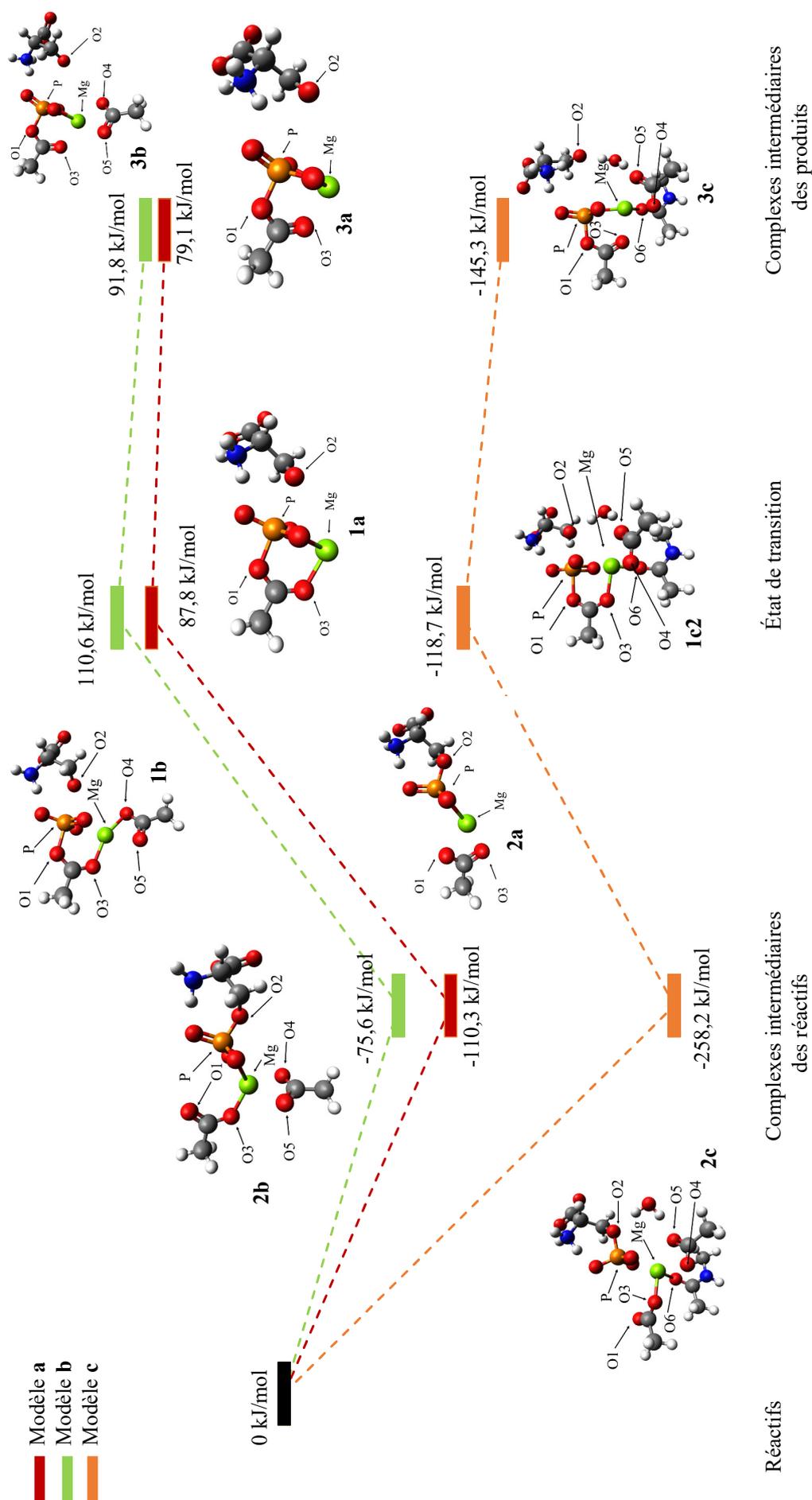
| Type | Molécules | O ₁ –P | O ₂ –P | O ₃ –P | C–O ₁ | C–O ₃ |
|--------------------|-------------------|-------------------|-------------------|-------------------|------------------|------------------|
| Molécules isolées | Ac ⁻ | / | / | / | 1,254 | 1,254 |
| | PLS ²⁻ | / | 1,701 | / | / | / |
| | PAc ²⁻ | 1,797 | / | 3,916 | 1,311 | 1,222 |
| Réactifs | 2a | 3,806 | 1,648 | 4,573 | 1,260 | 1,268 |
| | 2b | 3,941 | 1,657 | 4,045 | 1,234 | 1,280 |
| | 2c | 4,367 | 1,673 | 3,712 | 1,239 | 1,272 |
| État de transition | 1a | 1,906 | 2,556 | 3,053 | 1,280 | 1,247 |
| | 1b | 1,919 | 2,556 | 3,064 | 1,290 | 1,246 |
| | 1c2 | 2,012 | 2,459 | 3,119 | 1,272 | 1,249 |
| Produits | 3a | 1,766 | 3,585 | 2,937 | 1,306 | 1,233 |
| | 3b | 1,762 | 3,658 | 2,926 | 1,306 | 1,232 |
| | 3c | 1,783 | 4,418 | 3,018 | 1,304 | 1,225 |

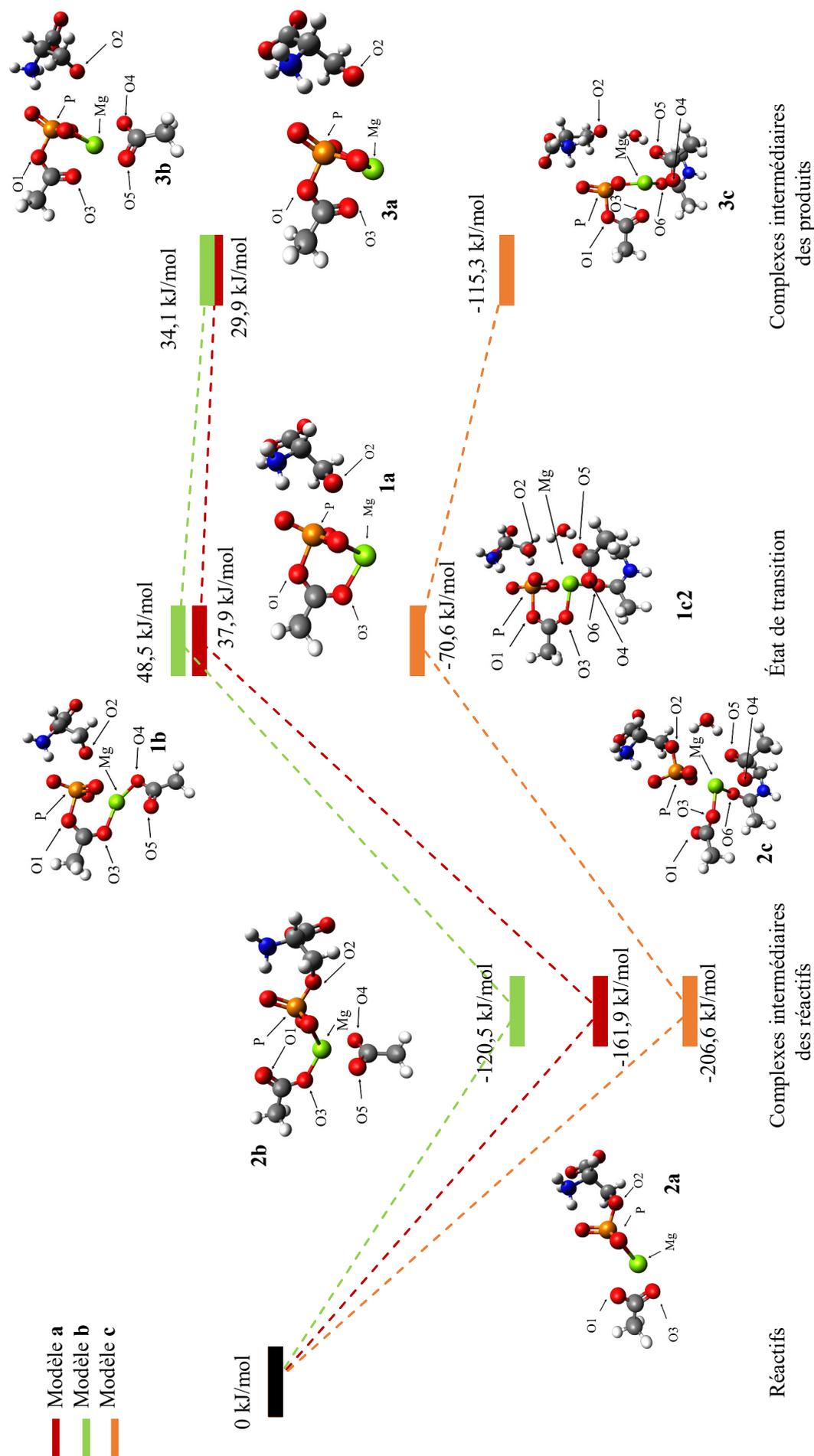
7.1.2 Aspects thermodynamiques et cinétiques

À l'aide de ces structures, les quantités thermodynamiques sont évaluées. Les ΔG_r^0 de la réaction d'attaque nucléophile ainsi obtenus à partir de ces quantités sont résumés à la Figure 7.6 ainsi que les ΔH_r^0 à la Figure 7.7. Suite à la réaction d'attaque nucléophile, les réactions de protonation et d'hydrolyse prennent le relais. Il n'y a donc pas de sens à considérer les produits à l'infini l'un de l'autre, les composés étant physiquement proches dans la poche catalytique. Les produits considérés lors de l'analyse thermodynamique sont donc les complexes intermédiaires. Au premier abord, une constatation peut être réalisée sur l'évolution de l'énergie libre et l'évolution de l'enthalpie. En effet, celles-ci évoluent de manière similaire, soulignant une influence constante de l'entropie.

Dans tous les modèles, la réaction a lieu en une étape, montrant un caractère de SN_2 . Le premier modèle contenant du Mg^{2+} (modèle **a**, courbe rouge) montre une énergie de réaction (ΔG_r^0) valant $79,1 \text{ kJ mol}^{-1}$. Lorsqu'on ajoute l'acétate modélisant l'Asp179 au système (modèle **b**, courbe verte), la réaction devient légèrement plus défavorable que dans le cas précédent. Toutefois, l'allure du diagramme reste cohérente avec le modèle précédent. Cette augmentation d'énergie de réaction s'explique par le fait que compléter la sphère de coordination du magnésium de manière partielle n'est pas optimal. En effet, ce dernier se retrouve avec une coordination à quatre ligands (O3, O4, O5 et un oxygène du phosphate), ce qui est déstabilisant pour cet atome qui préfère une coordination avec six ligands. D'autre part, l'élaboration d'un modèle avec une sphère de coordination complète à six ligands pour le magnésium montre la nécessité de retirer deux molécules d'eau à ce modèle pour que la phospho-L-sérine puisse approcher l'acétate. Le diagramme de la Figure 7.8 montre que retirer ces molécules d'eau est thermodynamiquement favorable si le point de départ est la sphère de coordination comprenant trois molécules d'eau ($-258,6 \text{ kJ mol}^{-1}$). L'approche de la phospho-L-sérine entraîne le départ de la seconde molécule d'eau. Deux chemins mécanistiques sont envisageables. L'obtention d'une sphère de coordination complète pour le Mg^{2+} (modèle **c**, courbe orange) a un effet stabilisant non négligeable sur la réaction d'attaque nucléophile qui suit. En effet, de cette manière, la réaction devient spontanée avec un ΔG_r^0 de $-145,3 \text{ kJ mol}^{-1}$ et un pK de $-25,9$. Cette analyse montre que l'ajout d'acides aminés dans le système tend à rendre possible cette réaction dans un système encombré ressemblant à l'environnement protéique. L'encombrement stérique mêlé à des interactions supplémentaires dans le système biologique ont donc un réel effet au niveau de la thermodynamique de la première étape du mécanisme.

D'un point de vue cinétique, l'énergie d'activation ΔG_r^\ddagger est plus grande dans le cas du modèle avec le magnésium ($198,1 \text{ kJ mol}^{-1}$) que dans le modèle avec le magnésium et l'aspartate ($186,2 \text{ kJ mol}^{-1}$). Cette baisse d'énergie d'activation peut être expliquée par une meilleure orientation réactionnelle. Le modèle avec une sphère de coordination complète possède une énergie d'activation plus basse que les deux premières ($139,5 \text{ kJ mol}^{-1}$). Ce système est donc à la fois favorisé cinétiquement et thermodynamiquement. En conclusion, l'attaque par l'Asp20 est donc envisageable comme première étape mécanistique.

Figure 7.6 – Diagramme d'énergie libre de Gibbs (ΔG_r^0 , kJ/mol) de la réaction d'attaque nucléophile par l'Asp20.

Figure 7.7 – Diagramme d'enthalpie (ΔH_r^0 , kJ/mol) de la réaction d'attaque nucléophile par l'Asp20.

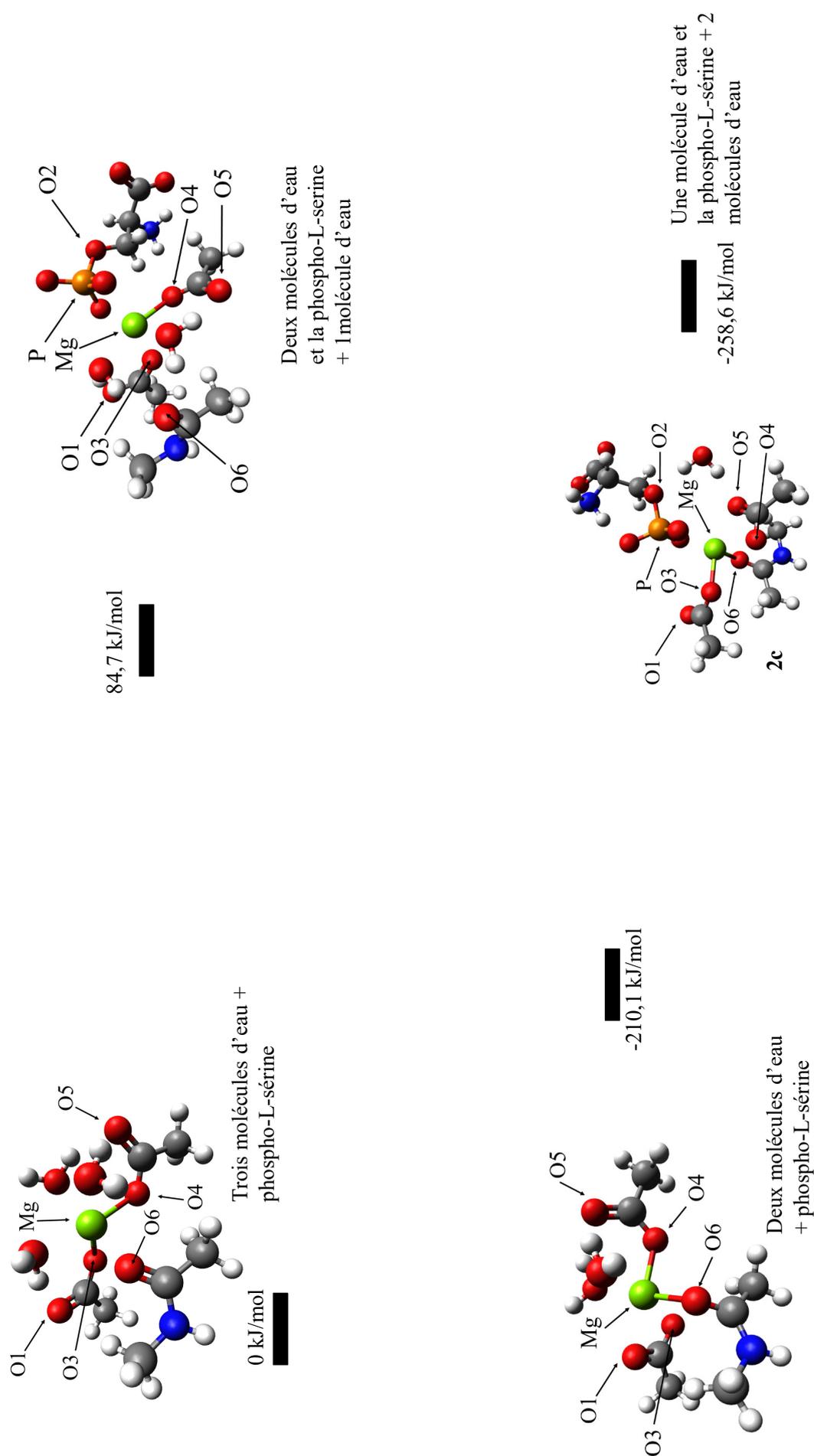


Figure 7.8 – Diagramme d'énergie libre de Gibbs (ΔG_r^0 , kJ/mol) pour retirer des molécules d'eau de la sphère de coordination. Les espèces ajoutées hors de la sphère sont à l'infini (symbolisé avec le +).

7.2 Protonation de la L-sérine par l'Asp22

La protonation de la L-sérine est réalisée au moyen de l'Asp22. Afin de prouver la faisabilité de cette réaction, une détermination des quantités thermodynamiques est entreprise pour trois modèles différents : un contenant un acétate modélisant l'Asp22 et la L-sérine déprotonée, un avec un ion Mg^{2+} supplémentaire et un dernier avec une coordination complète pour cet ion Mg^{2+} . Afin de simplifier les simulations, dans ces modèles, il est supposé que l'Asp20 a déjà été régénéré. Aucun phosphate n'est donc présent dans la poche catalytique. La coordination à six ligands est réalisée par trois molécules d'eau, le carbonyle de l'Asp22, deux oxygènes chacun associé à un acétate modélisant l'Asp20 et l'Asp179. Cette hypothèse s'appuie sur le fait que l'Asp20 influence la réaction que de manière structurale dans le cas présent.

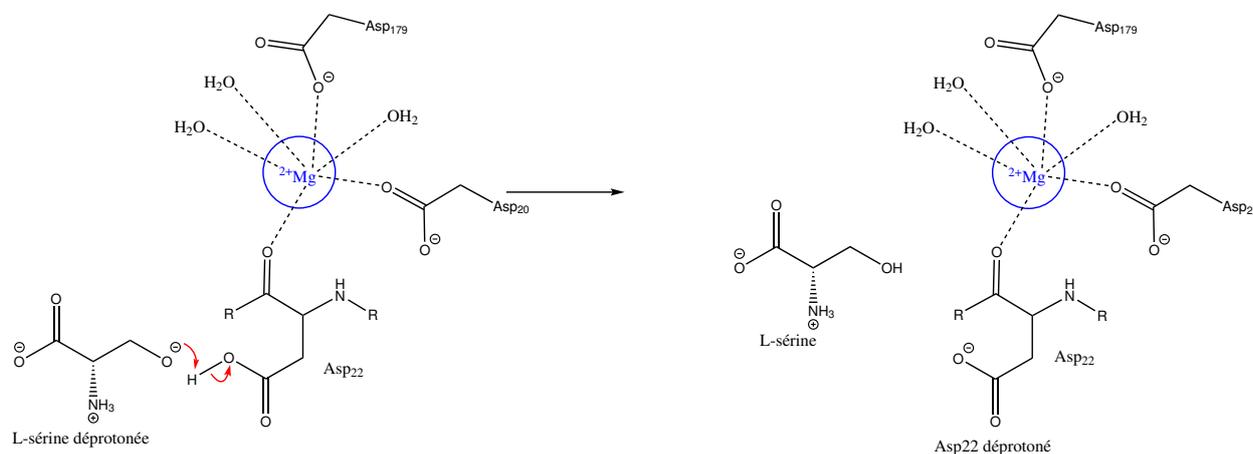


Figure 7.9 – Deuxième étape de réaction du système enzymatique. Réaction de protonation de la L-sérine par l'Asp22.

La réaction de protonation ayant lieu dans son intégralité dans la poche catalytique, les complexes intermédiaires sont respectivement considérés comme étant les réactifs et les produits de la réaction. Les trois modèles étudiés sont repris, par ordre de complexité croissante, aux Figures 7.10, 7.11 et 7.12 dans le cas des complexes intermédiaires des produits. Les quantités thermodynamiques correspondantes sont données dans le Tableau 7.2. L'analyse a mis en évidence que la réaction de protonation ne passait pas par un état de transition à proprement dit. Qu'importe le modèle, le Tableau 7.2 montre que la réaction de protonation est spontanée. Une légère différence d'énergie libre est observée entre les deux premiers modèles, explicable par une déstabilisation légère due à la présence d'un magnésium tricoordiné. Toutefois, bien que l'énergie de réaction soit relativement semblable entre ces deux modèles, elle est nettement moins favorable que dans

le cas du troisième modèle, où une différence de 30 kJ mol^{-1} est observée. Cette différence peut venir d'une meilleure stabilisation de la base correspondant à l'Asp22 à l'aide de la sphère de coordination. Cette stabilisation renforce le caractère acide de cet acide aminé.

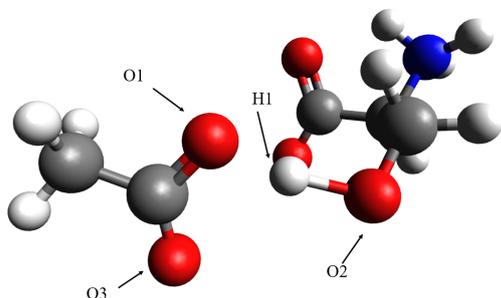


Figure 7.10 – Structure du complexe intermédiaire des produits résultant de la protonation de la L-sérine par l'Asp22, modélisé par l'acétate, calculée avec la méthode M06/6-311+G(d,p)/IEF-PCM. (Ac^-/Ser)

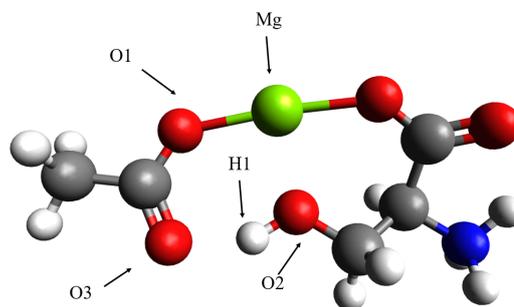


Figure 7.11 – Structure du complexe intermédiaire des produits résultant de la protonation par l'Asp22, modélisé par l'acétate, de la L-sérine protonée en présence de magnésium, calculé avec la méthode M06/6-311+G(d,p)/IEF-PCM. ($\text{Mg}(\text{Ac}^-/\text{Ser})^+$)

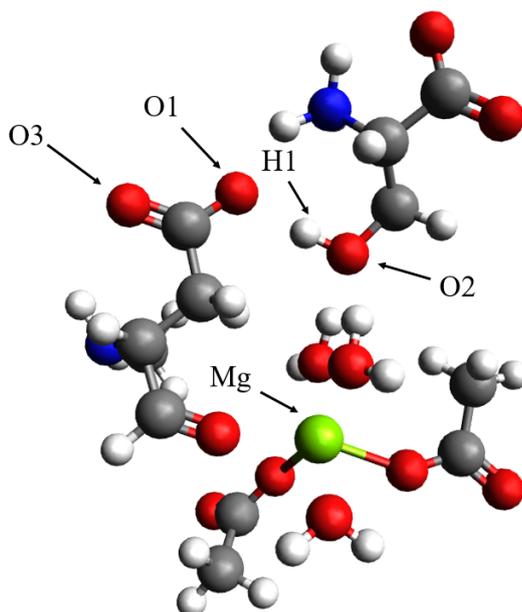


Figure 7.12 – Structure du complexe intermédiaire des produits résultant de la protonation par l'Asp22 déprotoné de la L-sérine protonée, en présence de magnésium entouré d'une sphère de coordination complète de trois molécules d'eau, de deux acétates pour symboliser l'Asp20 et l'Asp179 en plus de l'Asp22, calculée avec la méthode M06/6-311+G(d,p)/IEF-PCM. ($(\text{Mg}/\text{Ac}_2/(\text{H}_2\text{O})_3 (\text{Asp}_{22}^-/\text{Ser}))^-$)

Tableau 7.2 – Quantités thermodynamiques (kJ mol^{-1}) issues de la réaction de protonation de la L-sérine déprotonée par l'Asp22 protoné.

| Modèle | Réactif | Produit | ΔG_r^0 | ΔH_r^0 |
|--------|---|--|----------------|----------------|
| 1 | HAc/Ser ⁻ | Ac ⁻ /Ser | -75,4 | -68,7 |
| 2 | Mg(HAc/Ser ⁻) ⁺ | Mg(Ac ⁻ /Ser) ⁺ | -70,2 | -74,1 |
| 3 | (Mg/Ac ₂ /(H ₂ O) ₃ (Asp ₂₂ /Ser ⁻)) ⁻ | (Mg/Ac ₂ /(H ₂ O) ₃ (Asp ₂₂ ⁻ /Ser)) ⁻ | -103,3 | -89,7 |

7.3 Hydrolyse de l'Asp20

La dernière étape du mécanisme de HPSP, l'étape d'hydrolyse du phosphate, n'est pas clairement définie par la cristallographie. Bien que cette régénération a probablement lieu avec une molécule d'eau activée, deux mécanismes peuvent être envisagés pour l'activation de cette dernière, à savoir une activation soit par la Lys158 ou soit par l'Asp183. Les deux hypothèses sont testées afin de déterminer la plus probable. Dans le cas de l'hydrolyse par la Lys158, cette lysine perdrait temporairement un proton par effet des acides aminés environnant et activerait une molécule d'eau, comme cela est décrit à la Figure 7.13. Cette molécule d'eau réaliserait une attaque nucléophile sur l'Asp20 afin de libérer le phosphate. Après réaction, l'Asp20 donnerait son proton au milieu afin d'adopter le bon état de protonation.

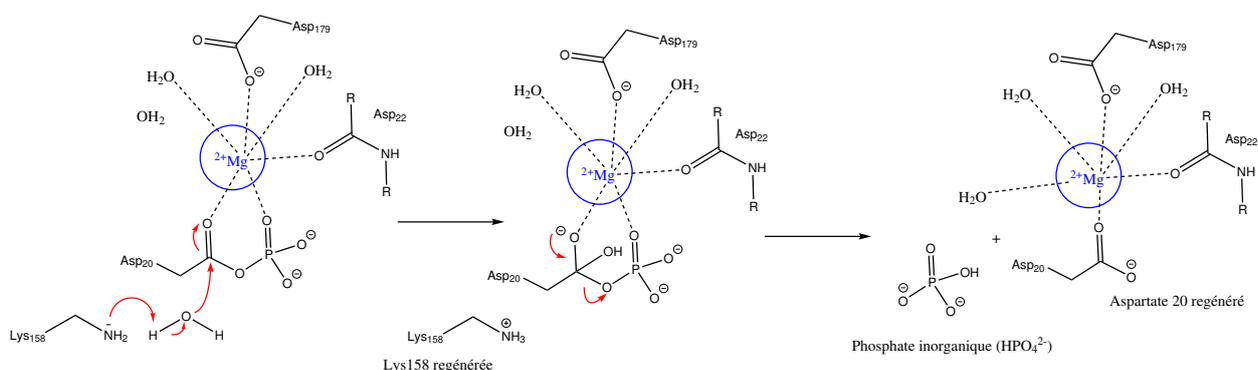


Figure 7.13 – Dernière étape de réaction du système enzymatique. Réaction d'hydrolyse de l'Asp20 phosphorylé engendrée par une molécule d'eau activée par la Lys158 afin de libérer le phosphate inorganique.

Dans le cas d'une réaction d'hydrolyse par l'Asp183, décrite à la Figure 7.14, l'Asp183 activerait la molécule d'eau en se protonant. Cette molécule réagirait sur le carbone de l'Asp20 et permettrait, par substitution nucléophile, la libération du phosphate. Après réaction, l'Asp20 ainsi que l'Asp183 seraient déprotonés en vue d'adopter le bon état de protonation. Dans les deux cas

de figure, la réaction a lieu en deux étapes. Il y a d'abord échange de proton entre l'activateur et l'eau ce qui permet l'activation d'une molécule d'eau. La molécule d'eau déprotonée attaque ensuite le carbonyle de l'Asp20 ce qui mène à la formation d'un intermédiaire, suivi du départ du phosphate en groupe partant. Les étapes d'activation et l'étape de départ du phosphate sont traitées séparément.

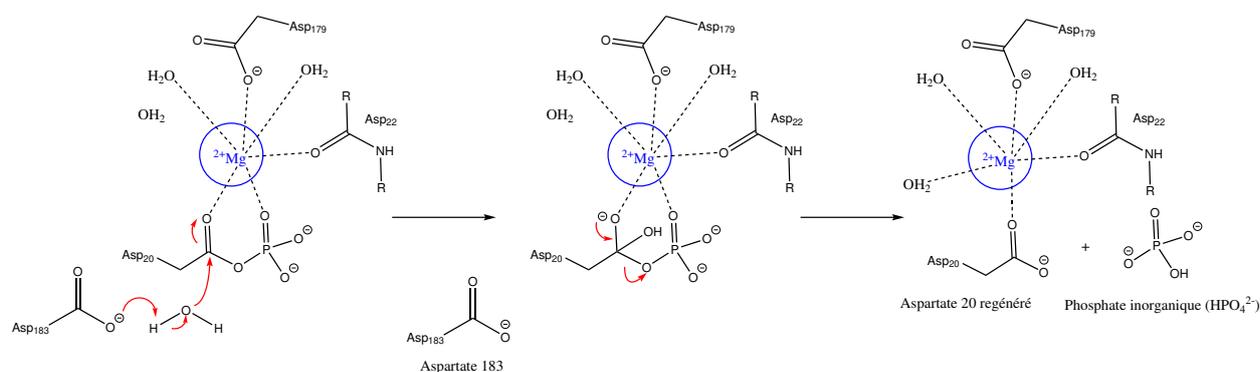
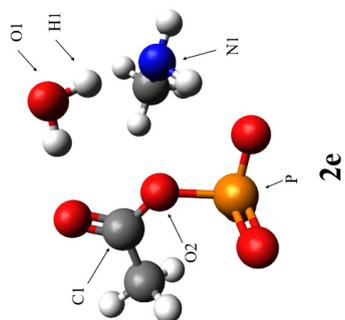


Figure 7.14 – Dernière étape de réaction du système enzymatique. Réaction d'hydrolyse de l'Asp20 phosphorylé engendrée par une molécule d'eau activée par l'Asp183 afin de libérer le phosphate inorganique.

7.3.1 Activation à partir de la Lys158

7.3.1.1 Aspects structuraux

Afin de simuler l'étape d'hydrolyse par la Lys158, une méthylamine ainsi qu'un phosphoacétate sont nécessaires comme réactif de départ. Différents modèles, de complexité croissante, sont élaborés pour traiter la réaction. Un premier état de transition (**1e**) est obtenu à partir d'une forme déprotonée de Lys158 avec de l'eau (modèle **e**). L'état de transition possède une fréquence imaginaire de 390,6i cm⁻¹ et est représenté à la Figure 7.15. À partir de ce modèle simple, une complexification, par ajout de Mg²⁺, est réalisée afin d'en voir l'influence sur le système (modèle **f**). L'état de transition obtenu (**1f**) possède une fréquence imaginaire de 600,3i cm⁻¹. Un dernier modèle est réalisé avec une sphère de coordination complète pour le Mg²⁺ (modèle **g**). Cette sphère comprend deux acétates modélisant l'Asp22 et l'Asp179 ainsi qu'un phosphoacétate représentant l'Asp20, le tout entouré de trois molécules d'eau. Cependant, différentes options devant être envisagées en ce qui concerne le nombre de molécules d'eau et la coordination du phosphate, l'optimisation de ce modèle n'a pas encore été fructueuse.

Complexes intermédiaires
des réactifs (2)

État de transition (1)

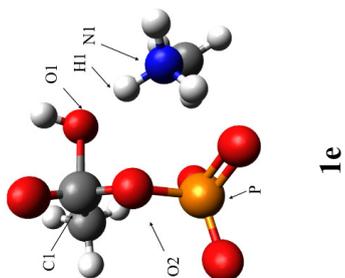
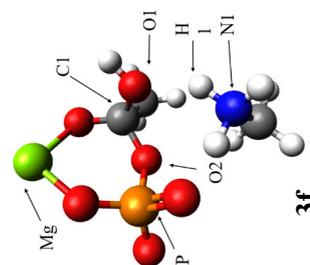
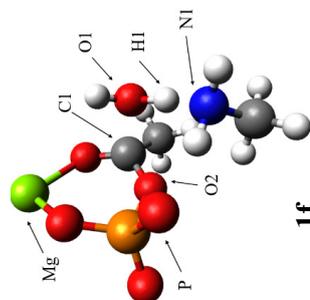
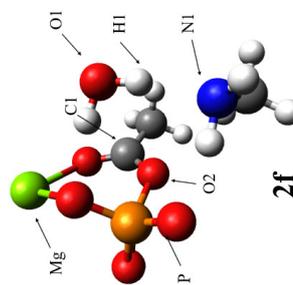
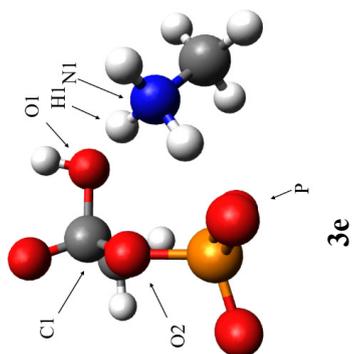
Complexes intermédiaires
des produits (3)

Figure 7.15 – Structures optimisées des complexes intermédiaires et des états de transition formés entre la lysine, simplifiée au moyen d'une méthylamine protonée, et une molécule d'eau déprotonée en fonction de la complexification du modèle, calculée avec la méthode M06/6-311+G(d,p)/IEF-PCM. (modèle **e** et **f**).

Les fréquences imaginaires obtenues soulignent la formation de la liaison N_1-H_1 aux dépens de la rupture de la liaison H_1-O_1 . Il y a donc protonation de l'amine et déprotonation de l'eau. L'ensemble des distances concernant les états de transition sont reprises au Tableau 7.3. Une variation de la distance des liaisons N_1-H_1 et H_1-O_1 est observée en fonction du modèle. La liaison du proton est renforcée par rapport à l'oxygène plutôt qu'à l'amine dans le second modèle. La variation des distances entre les deux états de transition explique la différence de fréquence imaginaire. La liaison C_1-O_2 est affaiblie par rapport au phosphoacétate isolé tandis que la liaison O_2-P est renforcée. Cela indique une régénération simplifiée lors de la deuxième étape d'hydrolyse. À partir de l'état de transition obtenu, les complexes intermédiaires sont optimisés. Les complexes correspondants aux réactifs sont repris à la Figure 7.15 (**2e** et **2f**). L'eau est observée avant activation en présence d'une lysine déprotonée. Les longueurs de liaisons N_1-H_1 et H_1-O_1 sont relativement identiques entre les réactifs des deux premiers modèles. Les distances de liaisons C_1-O_2 et O_2-P ont, elles aussi, une valeur proche de celles de la molécule isolée. Les complexes des produits sont repris dans cette même figure (**3e** et **3f**). Ces complexes soulignent le départ du proton provenant de l'eau vers le phosphoacétate afin de former, par attaque de l'hydroxyle généré sur le carbone, un carbonyle activé sous forme de O^- , et une lysine protonée. À l'aide du Tableau 7.3, le mouvement de l'oxygène O_1 au cours de la réaction est observable. Effectivement, en plus du mouvement de l'hydrogène afin de protoner l'amine, l'oxygène se rapproche du carbone C_1 . La même tendance que dans l'état de transition est observée pour les longueurs C_1-O_2 et O_2-P , soit un affaiblissement de la première et un renforcement de la seconde favorisant la suite de l'hydrolyse.

Tableau 7.3 – Distances interatomiques optimisées (Å) des réactifs de départ, des produits, des états de transition et des complexes intermédiaires correspondants durant la réaction d'activation par la Lys158.

| Type | Molécules | N_1-H_1 | H_1-O_1 | N_1-O_1 | O_1-C_1 | C_1-O_2 | O_2-P |
|--------------------|------------|-----------|-----------|-----------|-----------|-----------|---------|
| Molécule isolée | PAc^{2-} | / | / | / | / | 1,311 | 1,797 |
| Réactifs | 2e | 1,875 | 0,983 | 2,848 | 3,509 | 1,304 | 1,796 |
| | 2f | 1,868 | 0,984 | 2,841 | 3,371 | 1,292 | 1,789 |
| État de transition | 1e | 1,078 | 1,593 | 2,663 | 1,841 | 1,380 | 1,704 |
| | 1f | 1,270 | 1,223 | 2,489 | 1,791 | 1,354 | 1,719 |
| Produits | 3e | 1,046 | 1,777 | 2,810 | 1,521 | 1,445 | 1,673 |
| | 3f | 1,034 | 1,834 | 2,761 | 1,463 | 1,425 | 1,673 |

7.3.1.2 Aspects thermodynamiques et cinétiques

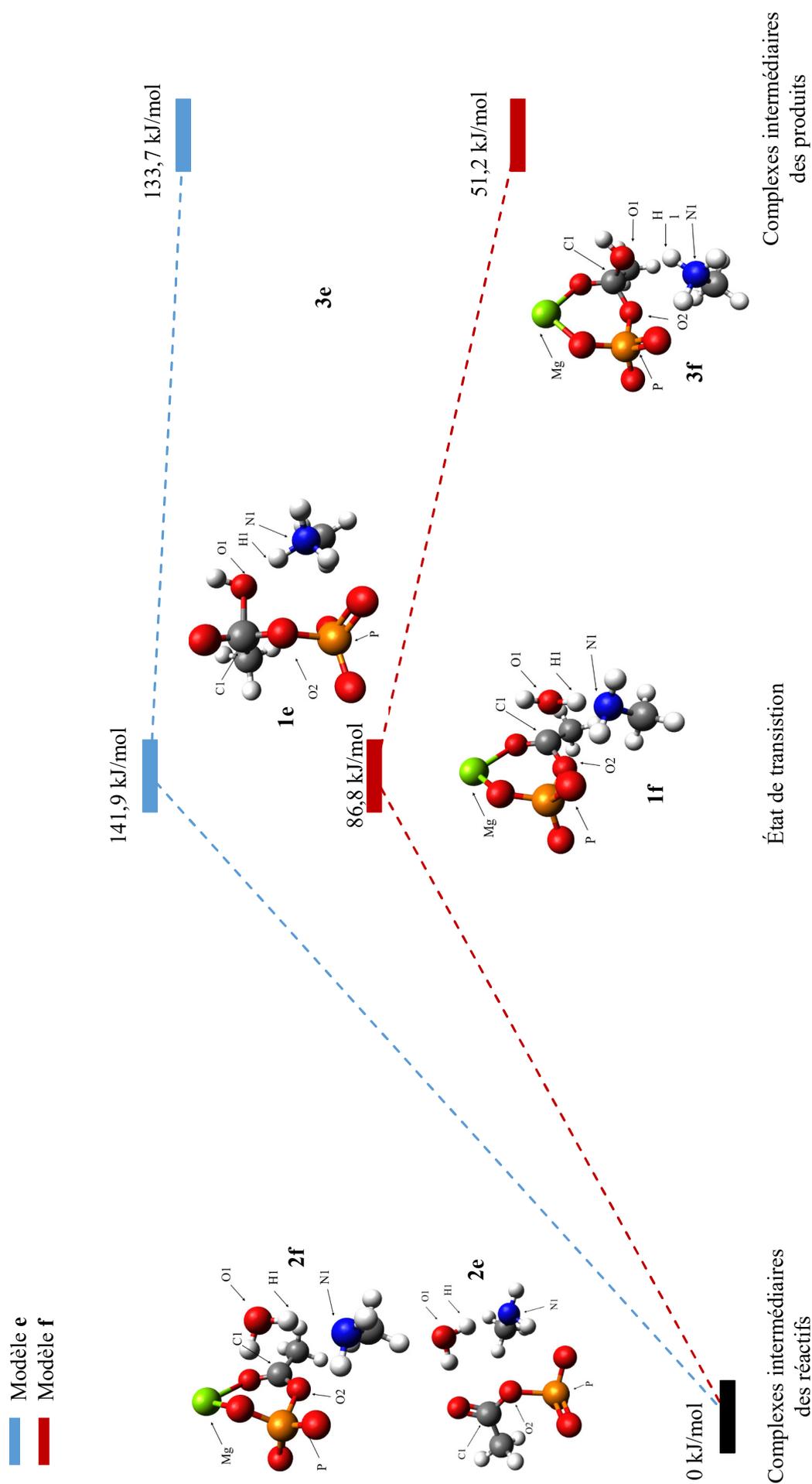
Dans le but de réaliser une étude thermodynamique de l'activation par la Lys158, les produits et les réactifs considérés sont les complexes intermédiaires. En effet, les réactifs et les produits ne sont pas à l'infini l'un de l'autre étant donné que dans la protéine la réaction ne s'arrête pas là : la seconde étape d'hydrolyse a ensuite lieu. Dans le modèle le plus simple (modèle **e**), la réaction d'hydrolyse avec la Lys158 est défavorable avec un ΔG_r^0 de $133,7 \text{ kJ mol}^{-1}$. Cependant, la Figure 7.16 montre une influence sur l'énergie des espèces intermédiaires lors de l'ajout d'un magnésium au système (modèle **f**) diminuant le ΔG_r^0 à une valeur de $51,2 \text{ kJ mol}^{-1}$ et à une constante d'équilibre de 0,98.

D'un point de vue cinétique, la réaction est plus favorisée dans le cas du modèle **f** (ΔG_r^\ddagger de $86,8 \text{ kJ mol}^{-1}$ au lieu de $141,9 \text{ kJ mol}^{-1}$) dans le modèle **e**. La thermodynamique et la cinétique privilégient toutes deux la réaction en présence de Mg^{2+} .

7.3.2 Activation à partir de l'Asp183

7.3.2.1 Aspect structuraux

D'une manière similaire, l'étape d'activation est réalisée à partir de l'Asp183 avec les trois modèles équivalents. Le premier modèle (modèle **e'**), composé d'un acétate modélisant l'Asp183, d'une molécule d'eau et d'un phosphoacétate permet d'isoler un état de transition **1e'**, observable à la Figure 7.17, dont la fréquence imaginaire s'élève à $435,7i \text{ cm}^{-1}$. Un modèle plus complexe (modèle **f'**), contenant un magnésium, est aussi simulé. L'état de transition de celui-ci (**1f'**) possède une fréquence imaginaire est de $548,5i \text{ cm}^{-1}$. Le dernier modèle complète la sphère de coordination du magnésium (modèle **g'**). Toutefois, pour les mêmes raisons qu'énoncées précédemment dans le modèle équivalent pour l'analyse de l'activation par la Lys158, les résultats de ce modèle n'ont pas encore été analysés.

Figure 7.16 – Diagramme d'énergie libre de Gibbs (ΔG_r^0 , kJ/mol) de l'activation d'une molécule d'eau par la Lys 158.

Dans les trois modèles, cette fréquence souligne la rupture de la liaison H_1-O_2 au profit de la liaison H_1-O_1 . L'eau est donc déprotonée. L'ensemble des distances concernant l'état de transition sont reprises au Tableau 7.4. Ce Tableau met en évidence un affaiblissement de la liaison O_1-H_1 entre l'état de transition dès lors de la complexification du modèle au profit d'un renforcement de la liaison H_1-O_2 . De manière cohérente avec les observations effectuées lors de l'activation par la Lys158, la liaison C_1-O_3 est affaiblie dans le premier modèle tandis que la liaison O_3-P est renforcée par rapport au phosphoacétate isolé. Cependant, au fur à mesure de la complexification du modèle, l'affaiblissement de la liaison C_1-O_3 est moins important. Par ailleurs, le renforcement de la liaison O_3-P n'est plus observé dans le modèle **1f'**. À partir de ces états de transition, les complexes intermédiaires des réactifs sont obtenus et observables à la Figure 7.17 (**2e'** et **2f'**). Ces complexes mettent en évidence que les réactifs de départ sont l'eau en présence de l'acétate. Les complexes ont des longueurs de liaison globalement semblables entre les deux modèles à l'exception de la liaison O_2-C_1 et la liaison O_3-P . Le renforcement de O_2-C_1 indique que l'eau se positionne plus près du carbonyle avant réaction lorsque le modèle est plus complexe. L'affaiblissement de la liaison O_3-P suit la tendance observée dans l'état de transition. Dans le réactif de départ, la liaison O_3-P est plus grande que dans le cas de la molécule isolée. En ce qui concerne les complexes des produits (**3e'** et **3f'**), le départ de l'hydrogène de l'eau vers l'acétate afin de former un l'hydroxyle, un acide acétique et le phosphoacétate permet l'activation du carbonyle du phosphoacétate, sous forme de O^- , par une attaque de l'hydroxyle généré. La liaison de l'hydroxyle au C_1 est observée par diminution de la longueur de liaison O_2-C_1 . Grâce à cette liaison, la liaison C_1-O_3 est affaiblie tandis que la liaison O_3-P est renforcée. L'hydrolyse est ainsi facilitée lors de la seconde étape de réaction.

Tableau 7.4 – Distances interatomiques optimisées (Å) des réactifs de départ, des produits, des états de transition et des complexes intermédiaires correspondants durant l'activation par l'Asp183.

| Type | Molécules | O_1-H_1 | H_1-O_2 | O_1-O_2 | O_2-C_1 | C_1-O_3 | O_3-P |
|--------------------|------------|-----------|-----------|-----------|-----------|-----------|---------|
| Molécule isolée | PAc^{2-} | / | / | / | / | 1,311 | 1,797 |
| Réactifs | 2e' | 1,759 | 0,984 | 2,742 | 3,582 | 1,302 | 1,807 |
| | 2f' | 1,707 | 0,999 | 2,610 | 3,343 | 1,297 | 2,023 |
| État de transition | 1e' | 1,064 | 1,425 | 2,487 | 1,750 | 1,398 | 1,697 |
| | 1f' | 1,268 | 1,148 | 2,415 | 1,816 | 1,358 | 1,829 |
| Produits | 3e' | 1,009 | 1,632 | 2,635 | 1,554 | 1,434 | 1,681 |
| | 3f' | 0,994 | 1,698 | 2,685 | 1,437 | 1,449 | 1,756 |

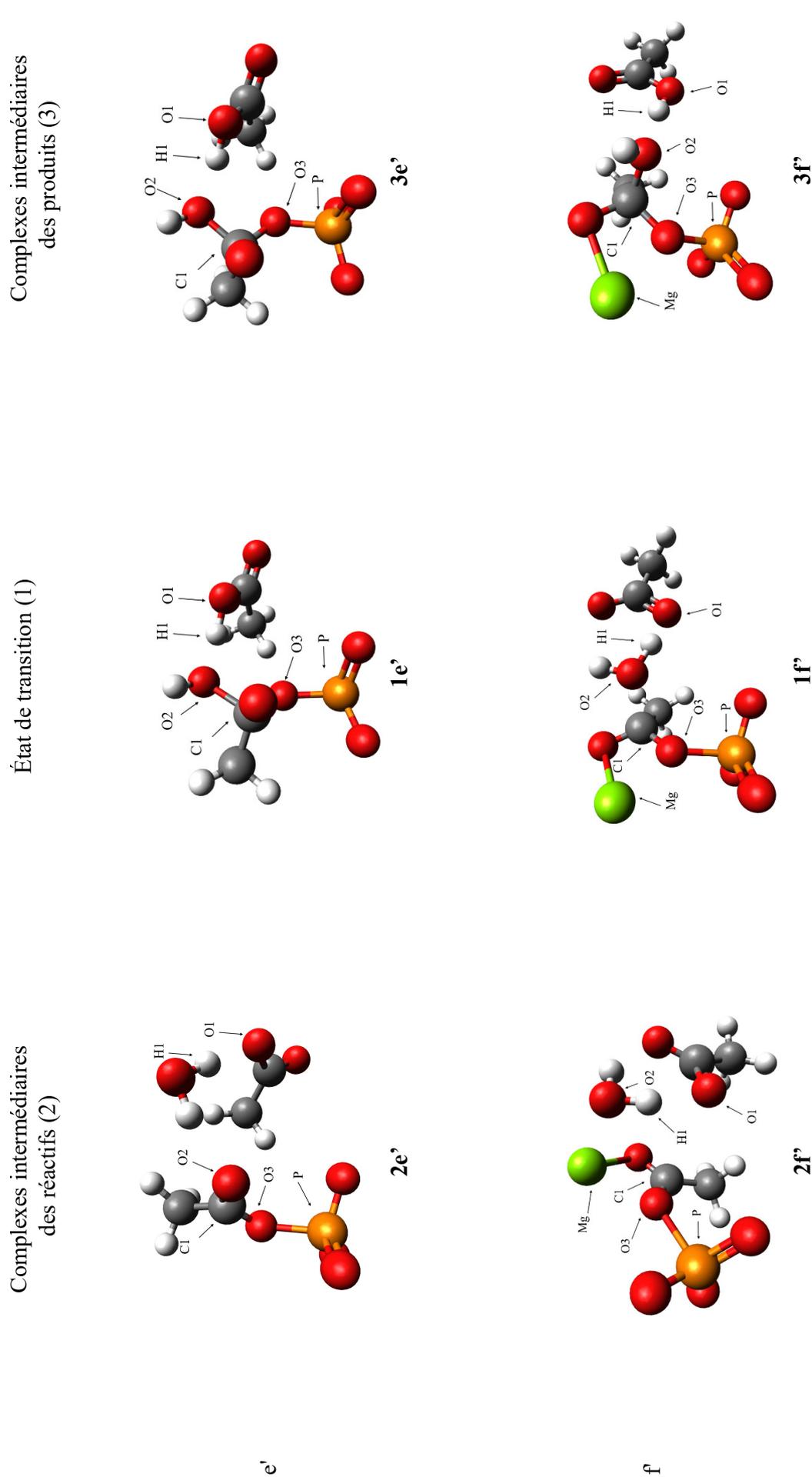


Figure 7.17 – Structures optimisées des complexes intermédiaires et des états de transition formés entre l'acétate protoné modélisant l'Asp183 et une molécule d'eau déprotonnée en fonction de la complexification du modèle (modèle **e'** et **f'**), calculées avec la méthode M06/6-311+G(d,p)/IEF-PCM.

7.3.2.2 Aspects thermodynamiques et cinétiques

Les quantités thermodynamiques (ΔG_r^0) correspondant aux différents modèles sont reprises à la Figure 7.18. Une nouvelle fois, les produits et les réactifs considérés sont les complexes intermédiaires. En effet, même si le phosphate est activé, il doit ensuite partir en tant que groupe partant.

De manière cohérente avec la réaction précédente, l'énergie de réaction se voit diminuée lorsque le modèle est complexifié, encourageant la réaction. Cependant, celle-ci reste fortement défavorable dans les deux premiers modèles (modèle **e'** avec un ΔG_r^0 de 162,1 kJ mol⁻¹ et **f'** avec un ΔG_r^0 de 100,3 kJ mol⁻¹, respectivement en bleu et rouge). La réaction n'est donc pas envisageable d'un point de vue thermodynamique dans ces deux modèles simplifiés. De plus, ces quantités thermodynamiques sont supérieures à celles obtenues dans le cas de la lysine (ΔG_r^0 de 133,7 kJ mol⁻¹ et de 51,2 kJ mol⁻¹ pour les modèles équivalents, par ordre de complexification). Malgré que la réaction soit défavorable, un abaissement de 49,1 kJ mol⁻¹ est observé entre le modèle **f'** de l'Asp183 et le modèle **f** de la Lys158. Cette nette diminution semble encourager l'activation par la lysine. Cela a du sens puisque la lysine est une meilleure base que l'aspartate. Bien qu'un modèle avec une sphère de coordination complète n'a pas été obtenu dans ces deux cas, étant donné que l'énergie du système tend à diminuer au cours de sa complexification, il est possible que si la réaction est réalisée dans ce troisième modèle, elle soit spontanée.

Toutefois, pour que cette réaction ait lieu, la Lys158 doit être sous une forme déprotonée avant réaction, ce qui ne correspond pas à sa forme à pH physiologique. Cette déprotonation pourrait être réalisée par l'environnement protéique, cela ayant déjà été mis en évidence dans d'autres mécanismes enzymatiques.⁷⁸⁻⁸⁰ Le pKa des chaînes latérales des acides aminés peut différer entre la solution et l'enzyme par la présence des acides aminés voisins. Notamment, la présence de l'Asp20 et l'Asp183 peuvent contribuer à cet environnement basique.

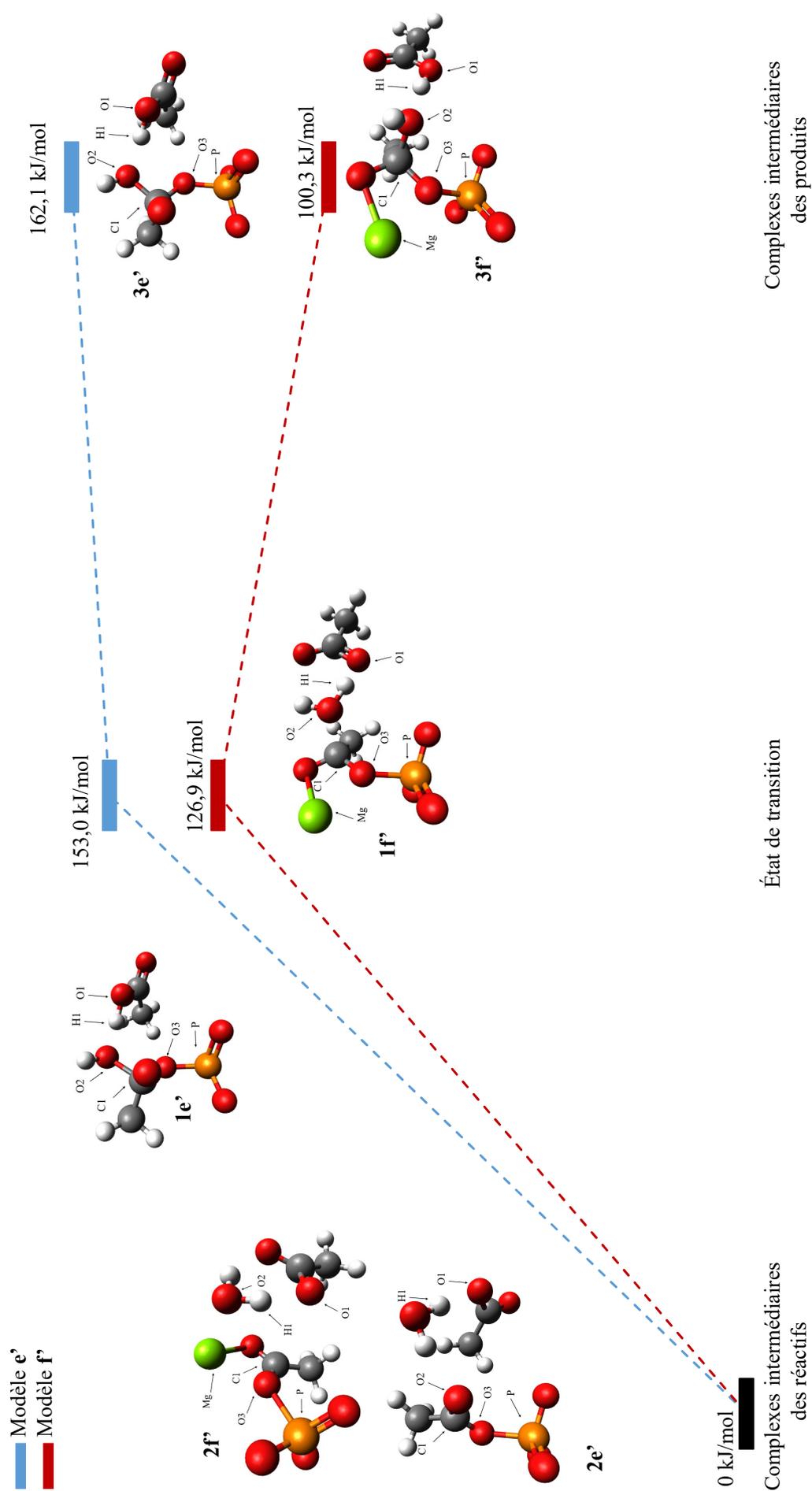


Figure 7.18 – Diagramme d'énergie libre de Gibbs (ΔG_p^0 , kJ/mol) de l'activation d'une molécule d'eau par l'Asp183 pour les différents modèles étudiés (modèle e' et modèle f').

D'un point de vue cinétique, l'énergie d'activation (ΔG_r^\ddagger) diminue également lors de la complexification du modèle. Toutefois, elle demeure élevée dans les deux modèles (153,0 kJ mol⁻¹ pour le modèle **e'** et 126,9 kJ mol⁻¹ pour le modèle **f'**, respectivement en bleu et rouge). En comparaison aux modèles équivalents dans le cas de la lysine, les énergies d'activation sont moins élevées dans le cas de l'activation par la lysine (141,9 kJ mol⁻¹ pour le modèle **e** et 86,8 kJ mol⁻¹ pour le modèle **f**). L'activation par la Lys158 est donc favorisée cinétiquement par rapport à une activation par l'Asp183. En conclusion, ayant un coût cinétique et thermodynamique moins important, l'activation de l'eau par la Lys158 est plus judicieuse d'un point de vue mécanistique.

7.3.3 Élimination du phosphate

7.3.3.1 Aspect structuraux

La dernière étape du mécanisme étudié durant ce mémoire est le départ du phosphate après activation. L'étude de cette réaction est réalisée dans différents modèles de complexification croissante. Par simplification du modèle et la flexibilité de la lysine ou de l'aspartate, l'acide aminé activateur n'est pas considéré dans le modèle. Ces modèles sont composés au minimum du phosphoacétate activé par insertion de l'hydroxyle et transformation du carbonyle en O⁻. Un premier modèle avec ce réactif est réalisé (modèle **h**). Ce modèle, dont l'état de transition (**1h**) est observable à la Figure 7.19, possède une fréquence imaginaire de 295,3i cm⁻¹. L'ajout de magnésium (modèle **i**) permet d'isoler l'état de transition (**1i**), de fréquence imaginaire de 377,7i cm⁻¹. Ces deux premiers modèles montrent déjà une différence entre les deux états de transition obtenus. Effectivement, bien que, dans les deux cas, la fréquence imaginaire témoigne de la rupture de la liaison O₂-C₁ au profit de la formation d'une double liaison O₃-C₁, l'hydrogène H₁ n'est pas à la même position dans les deux états de transition. Il montre une orientation préférentielle vers le phosphate dans le modèle contenant un magnésium. Les distances de liaisons reprises au Tableau 7.5 sont relativement semblables entre les deux modèles. Un troisième modèle (modèle **j**), avec une sphère de coordination complète pour le magnésium, n'a pas pu être réalisé par manque de temps afin d'explorer les différentes opportunités de coordination des molécules d'eau ou encore du phosphate.

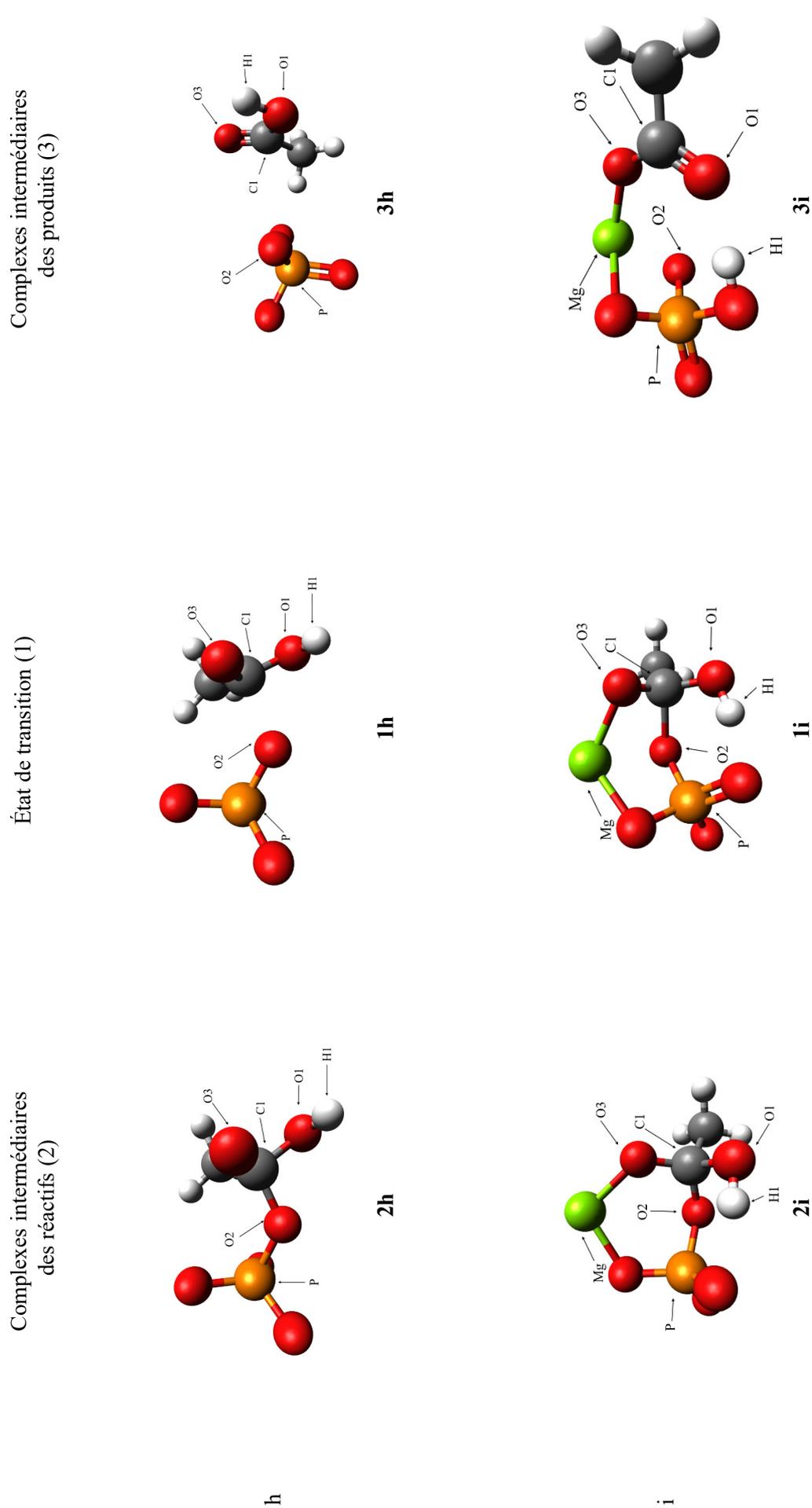


Figure 7.19 – Structures optimisées des complexes intermédiaires et des états de transition formés entre le carbonyle activé de l'acétate symbolisant l'Asp20 et le phosphate en groupe partant (modèle **h** et **i**), calculées avec la méthode M06/6-31 1+G(d,p)/IEF-PCM.

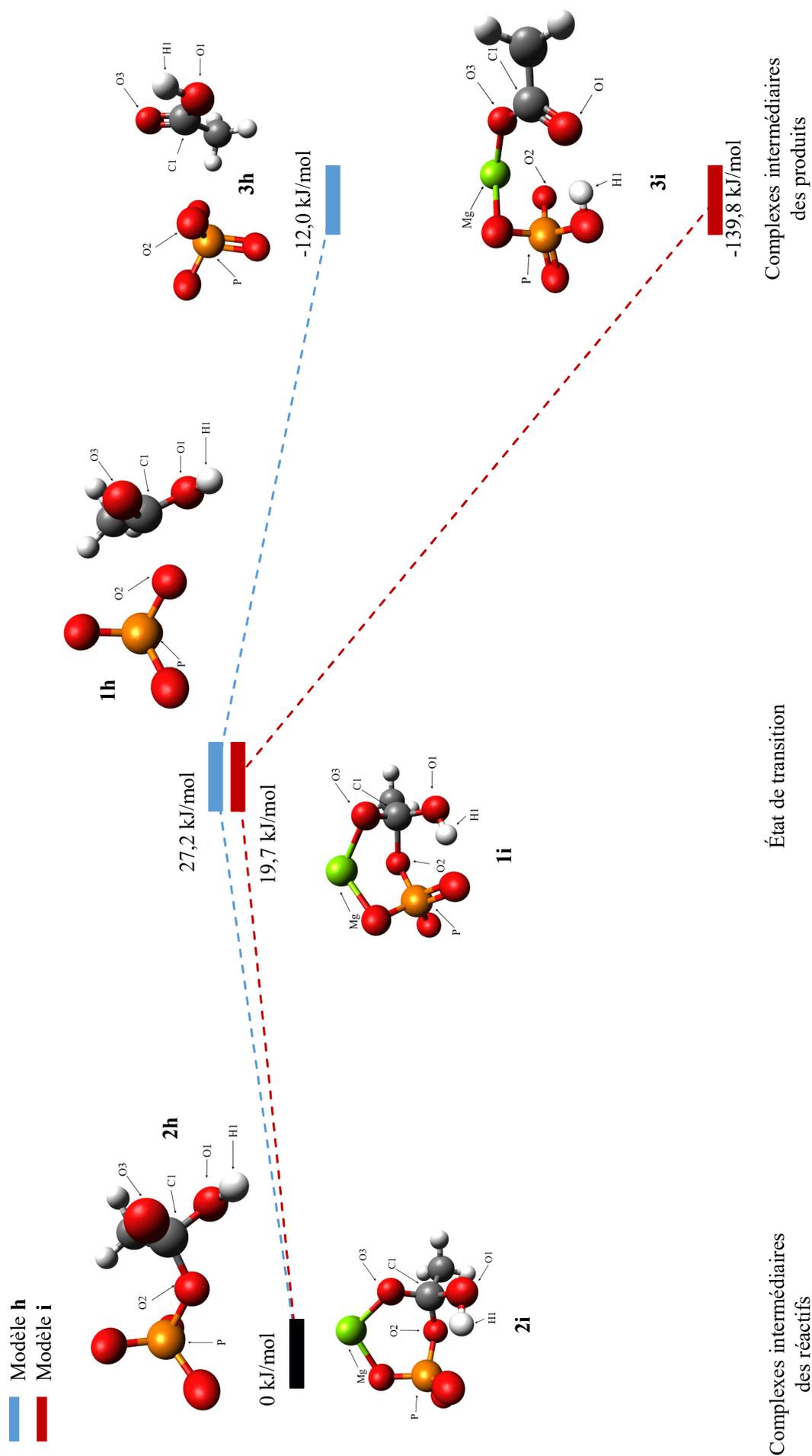
À partir des états de transition, les complexes intermédiaires des réactifs sont obtenus et repris à la Figure 7.19 (**2h** et **2i**). Dans le cas des deux premiers modèles, les distances de liaisons de ces réactifs, reprises au Tableau 7.5, montrent une liaison forte entre l'oxygène liant le phosphate et le carbone (C_1-O_2), cohérente avec les distances après activation par la Lys158. Par sa distance, la liaison O_3-C_1 est plus proche d'un caractère de simple liaison que dans l'état de transition. L'effet de la complexification du modèle est peu observable en termes de distance pour les réactifs. Les complexes intermédiaires des produits sont repris à la Figure 7.19 (**3h** et **3i**). L'analyse des distances de liaisons (Tableau 7.5) dans ces complexes souligne le départ du phosphate (liaison C_1-O_2) et formation d'une double liaison entre O_3-C_1 plus prononcée que dans l'état de transition. La liaison moins forte pour O_3-C_1 dans le modèle **3i** s'explique par une coordination de l'oxygène O_3 à l'ion Mg^{2+} . Des différences existent entre la position de l'hydrogène H_1 dans ces complexes. En effet, on peut remarquer une liaison de l'hydrogène en question à un oxygène lié au phosphate dans le deuxième modèle, augmentant la liaison O_1-H_1 . Cette liaison est plus cohérente avec la réalité à pH physiologique, permettant d'avoir un bon état de protonation pour l'Asp20 après hydrolyse.

Tableau 7.5 – Distances interatomiques optimisées (Å) des réactifs de départs, des produits, des états de transition et des complexes intermédiaires correspondants durant au départ du phosphate.

| Type | Molécules | O_1-H_1 | O_1-C_1 | C_1-O_2 | O_3-C_1 | O_2-P |
|--------------------|------------|-----------|-----------|-----------|-----------|---------|
| Molécule isolée | Ac^- | / | 1,254 | / | 1,254 | / |
| | PAc^{2-} | / | / | 1,311 | 1,222 | 1,797 |
| Réactifs | 2h | 0,963 | 1,476 | 1,469 | 1,297 | 1,672 |
| | 2i | 0,982 | 1,419 | 1,455 | 1,340 | 1,669 |
| État de transition | 1h | 0,965 | 1,390 | 1,904 | 1,239 | 1,601 |
| | 1i | 1,009 | 1,346 | 1,819 | 1,287 | 1,644 |
| Produits | 3h | 0,968 | 1,344 | 3,234 | 1,211 | 1,566 |
| | 3i | 1,788 | 1,240 | 3,082 | 1,276 | 1,562 |

7.3.3.2 Aspect thermodynamiques et cinétiques

Les quantités thermodynamiques (ΔG_r^0) de la réaction de départ du phosphate sont reprises à la Figure 7.20 en fonction des différents modèles étudiés. La réaction est favorable dans les deux modèles bien qu'une diminution de l'énergie libre est toujours observée au fur à mesure de la complexification du modèle. Cette forte diminution ($-127,8 \text{ kJ mol}^{-1}$) entre le premier (modèle **h**)

Figure 7.20 – Diagramme d'énergie libre de Gibbs (ΔG_r^0 , kJ/mol) de la réaction de départ du phosphate.

et le deuxième modèle (modèle **i**) s'explique par la protonation du phosphate améliorant son caractère de meilleur groupe partant. De plus, le second modèle possède un état de protonation cohérent avec un pH physiologique. La réaction est donc thermodynamiquement plus favorable.

La spontanéité de cette réaction rend le bilan énergétique global de la réaction d'hydrolyse favorable, la première étape mécanistique, bien que défavorable, est entraînée par la seconde, très spontanée. D'un point de vue cinétique, l'énergie d'activation (ΔG_r^\ddagger) est petite, soulignant une faible barrière énergétique de la réaction (27,2 kJ mol⁻¹ pour le modèle **h** et 19,7 kJ mol⁻¹ dans le modèle **i**). Cette faible énergie d'activation permet une réaction rapide. Au vu de la spontanéité de la réaction d'un point de vue thermodynamique et de la faible barrière énergétique, cette réaction de départ de phosphate à l'aide d'un intermédiaire tétraédrique formé par attaque de l'hydroxyle est tout à fait envisageable d'un point de vue mécanistique.

7.3.4 Conclusions de la réaction d'hydrolyse

Sur base de ces observations de quantités thermodynamiques, la réaction d'attaque nucléophile semble se dérouler en deux étapes. La première étape, l'activation de la molécule d'eau, est une réaction de déprotonation moins coûteuse d'un point de vue cinétique et thermodynamique si la réaction a lieu à l'aide de la Lys158 et non de l'Asp183. L'analyse souligne que, pour que cette réaction ait lieu, la Lys158 doit être sous une forme déprotonée avant réaction, ce qui ne correspond pas à sa forme à pH physiologique. Cette déprotonation pourrait être réalisée par l'environnement protéique, ce mécanisme ayant déjà été mis en évidence dans d'autres mécanismes enzymatiques.⁷⁸⁻⁸⁰ La seconde étape réactionnelle, l'hydrolyse, conduit au départ du phosphate. Après activation, l'hydroxyle, généré lors de l'activation, attaque le carbone du phosphoacétate et un intermédiaire est obtenu, correspondant au carbonyle activé. Ce carbonyle retrouve sa forme normale en éjectant le phosphate via une SN₂. Qu'importe le modèle, ce départ est favorisé. La différence d'énergie de réaction est d'autant plus grande si un oxygène du phosphate capte l'hydrogène de l'hydroxyle. Cet abaissement s'explique par un caractère de meilleur groupe partant du phosphate dans le cas où il est protoné. Dans le cas des deux étapes, il est possible qu'une complexification du système mette en évidence une réaction plus spontanée. Cependant, un modèle avec une coordination complète pour le magnésium n'a pas encore pu être obtenu.

7.4 Système protéique

Les systèmes décrits précédemment permettent de se rendre compte de la faisabilité de la réaction dans un modèle mimant le modèle protéique. Cependant, il est possible, à partir des fichiers de coordonnées obtenus par la cristallographie, d'élaborer un modèle encore plus complet de la réaction : un modèle protéique avec tout l'environnement structural du site actif. Ce modèle plus complexe permettrait donc d'établir l'évaluation de toutes les quantités thermodynamiques dans un même modèle, et donc d'avoir une vision continue de la réaction.

Afin de pouvoir construire ce modèle, les coordonnées des acides aminés clés du site actif obtenues par la cristallographie sont sélectionnées (Tableau 7.6) en fonction de leur importance. Les voisins de ces acides aminés ainsi que les résidus à moins de 4 Å du métal sont aussi ajoutés au modèle afin de donner une structure au site actif. L'ensemble de ces acides aminés forment des séries de peptides. Pour chaque fragment, l'extrémité N-terminale est mutée en une amide tandis que l'extrémité C-terminale est transformée en cétone afin de conserver l'hybridation du carbone. Les hydrogènes sont ajoutés aux acides aminés et aux molécules d'eau en respectant les conditions de pH physiologique. La structure du modèle final est reprise à la Figure 7.21. Un modèle de 532 atomes est ainsi obtenu. La base 6-311G(d) et la fonctionnelle M06 sont choisies afin d'optimiser le modèle. Pour réduire le temps de simulation, des contraintes sont imposées au système sur les carbones α des acides aminés terminaux pour les différents peptides du modèle. Cela fixe la position de ces carbones dans l'espace et rigidifie le système. Six modèles ont ainsi été optimisés en faisant varier le nombre de contraintes imposées de quatre à quatorze, les modèles plus élevés incluant les contraintes des modèles inférieurs^A. Les contraintes sont représentées à la Figure 7.21. La solvation est décrite à l'aide d'un modèle SAS. Au vu du temps de calcul nécessaire et de la cohérence du modèle obtenu, le modèle 5 est défini comme le modèle le plus approprié en ce qui concerne les contraintes pour l'optimisation. Toutefois, l'optimisation est toujours en cours. La convergence est actuellement de l'ordre de 10^{-4} Hartree pour 12 processeurs et 30Gb de mémoire. Une fois la convergence atteinte, l'étude de la réactivité dans un modèle proche de la protéine pourra être effectuée.

A. Par exemple, le modèle 5 inclut les contraintes sur les carbones α du modèle 4, 3, 2 et 1

Tableau 7.6 – Acides aminés clés choisis dans le modèle mimant le système enzymatique.

| Acides aminés | Rôle présumé | Mutation en N-terminale | Mutation en C-terminale |
|---------------|-------------------------------|-------------------------|----------------------------------|
| Phe 19 | Structure | Oui | Non |
| Asp20 | Nucléophile | Non | Non |
| Val21 | Structure | Non | Non |
| Asp22 | Acide-base | Non | Non |
| Ser23 | Structure | Non | Non |
| Tyr24 | Structure | Non | Non |
| Val25 | Structure | Non | Oui |
| Glu 28 | Structure | Oui | Non mais COOH en CH ₃ |
| Glu 29 | Liaisons ligands | Non | Non |
| Gly30 | Structure | Non | Oui |
| Ala51 | Boucle dynamique | Oui | Non |
| Met52 | Boucle dyna./Liaisons ligands | Non | Non |
| Gly53 | Boucle dynamique | Non | Oui |
| Pro57 | Boucle dynamique | Oui | Oui |
| Gly64 | Structure | Oui | Non mais COOH en CH ₃ |
| Arg65 | Liaisons ligands | Non | Oui |
| Leu66 | Structure | Non | Oui |
| Ser109 | Liaisons ligands | Oui | Non |
| Gly110 | Liaisons ligands | Non | Non |
| Gly111 | Liaisons ligands | Non | Oui |
| Glu154 | Structure | Oui | Non mais COOH en CH ₃ |
| Ser155 | Structure | Non | Non |
| Gly156 | Structure | Non | Non |
| Gly157 | Structure | Non | Non |
| Lys158 | Régénération | Non | Non |
| Gly159 | Structure | Non | Oui |
| Gly178 | Structure | Oui | Non |
| Asp179 | Coordination Mg ²⁺ | Non | Non |
| Gly180 | Structure | Non | Oui |
| Thr182 | Liaisons ligands | Oui | Non |
| Asp183 | Régénération | Non | Non |
| Met184 | Structure | Non | Oui |
| Ile201 | Structure | Oui | Non |
| Arg202 | Liaisons ligands | Non | Non |
| Gln203 | Structure | Non | Oui |

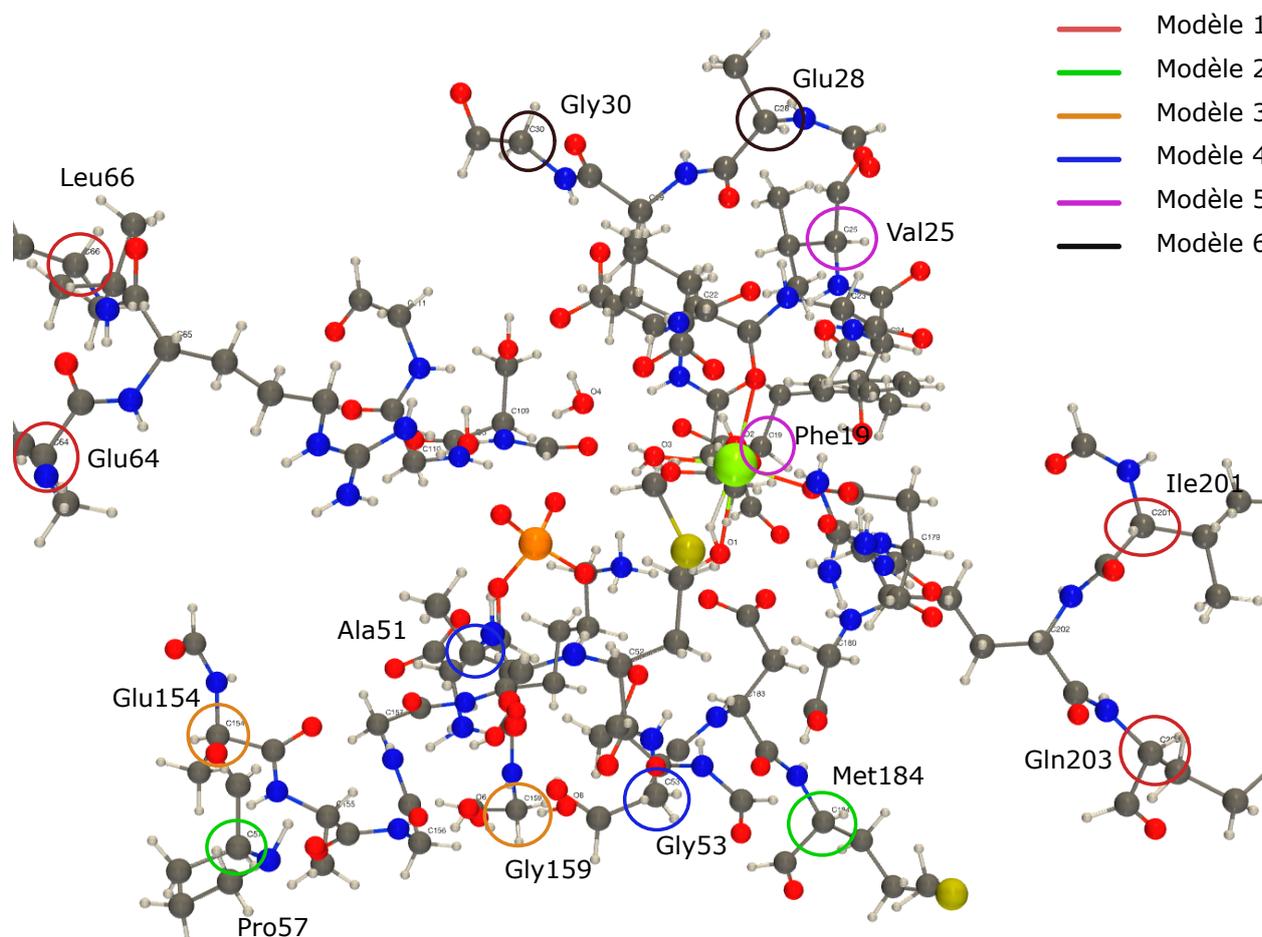


Figure 7.21 – Structure mimant le modèle enzymatique utilisée lors de l'optimisation. Contraintes imposées sur les carbones α en fonction du type de modèle. Le modèle 1 est mis en évidence par un rond rouge autour des carbones α fixés, le 2 est montré par des ronds verts, le 3 est indiqué par des ronds oranges, le 4 par des ronds bleus, le 5 par des ronds mauves et le 6 par des ronds noirs. La structure est calculée avec la méthode M06/6-311G(d)/IEF-PCM/SAS.

7.5 Conclusions intermédiaires

Au cours du présent chapitre, les trois réactions impliquées dans l'hydrolyse de la phospho-L-sérine sont analysées. Premièrement, l'analyse computationnelle a mis en évidence la spontanéité de la réaction d'attaque nucléophile lorsqu'un modèle proche du modèle protéique est utilisé. L'encombrement stérique ainsi que les interactions supplémentaires rendent possible la réaction. Un abaissement de l'énergie d'activation est aussi observé, soulignant que lorsqu'on se rapproche d'un système protéique, le rôle de catalyseur de ce dernier se révèle. L'attaque de la phospho-L-sérine par l'Asp20 est donc une réaction tout à fait raisonnable comme première étape mécanistique. Deuxièmement, le caractère spontané de la réaction de protonation est mis en évidence. Dès que la L-sérine déprotonée est à proximité de l'Asp22, le transfert de proton a lieu, quelle que soit la complexité du système, sans passer par un état de transition isolable. Toutefois, un système plus

complexe, reprenant un ion magnésium entouré d'une sphère de coordination à six ligands, permet un abaissement de plus grand l'énergie d'activation, explicable par une stabilisation de l'Asp22 comme base.

Troisièmement, en ce qui concerne la réaction d'hydrolyse de l'Asp20, les résultats semblent indiquer une réaction en deux étapes. La première étape, l'activation de l'eau, semble être réalisée par la Lys158 et non pas l'Asp183, cette étape étant moins défavorable dans le premier cas. À partir d'une Lys158 déprotonée et d'une molécule d'eau en présence d'un phosphoacétate modélisant l'Asp20 phosphorylé, il est ainsi possible d'obtenir une Lys158 protonée et le carbonyle de l'Asp20 activé par attaque de l'hydroxyle formé. Toutefois, pour que cette réaction ait lieu, la Lys158 doit être sous une forme déprotonée avant réaction, ce qui ne correspond pas à sa forme à pH physiologique. Cette déprotonation pourrait être réalisée par l'environnement protéique, ce mécanisme ayant déjà été mis en évidence dans d'autres mécanismes enzymatiques.⁷⁸⁻⁸⁰ Bien que non spontanée, cette première étape est entraînée par la seconde, largement favorable, quelle que soit la complexité du système. Cette seconde étape correspond au départ du phosphate comme groupe partant avec pour réactif de départ le carbonyle activé de l'Asp20 par attaque de l'hydroxyle formé à partir de l'eau activée. Cette réaction est encore plus favorisée dans le cas de modèles plus complexes, permettant, suite à une orientation plus favorable, le départ du phosphate sous forme de HPO_4^{2-} . Cela s'explique par un caractère de meilleur groupe partant de cette espèce par rapport à l'espèce PO_4^{3-} .

Enfin, la réactivité peut être étudiée dans un modèle plus complexe, plus proche du modèle protéique. Cette étude permettrait d'évaluer les quantités thermodynamiques des trois réactions dans un même modèle. Cela rendrait ainsi possible l'obtention d'une vision globale de l'évolution des quantités thermodynamiques dans la protéine. De plus, il serait ainsi possible d'en apprendre plus sur le mouvement des molécules d'eau. Cependant, l'optimisation de ce modèle est toujours en cours. Reprenant toutes les observations mécanistiques réalisées par l'analyse computationnelle, l'hypothèse mécanistique a pu être étoffée. Celle-ci est présentée à la Figure 7.22. Si la Lys158 est déprotonée avant l'activation de la molécule d'eau, il faut considérer ses interactions avec les ligands comme étant des interactions de type ponts hydrogène.

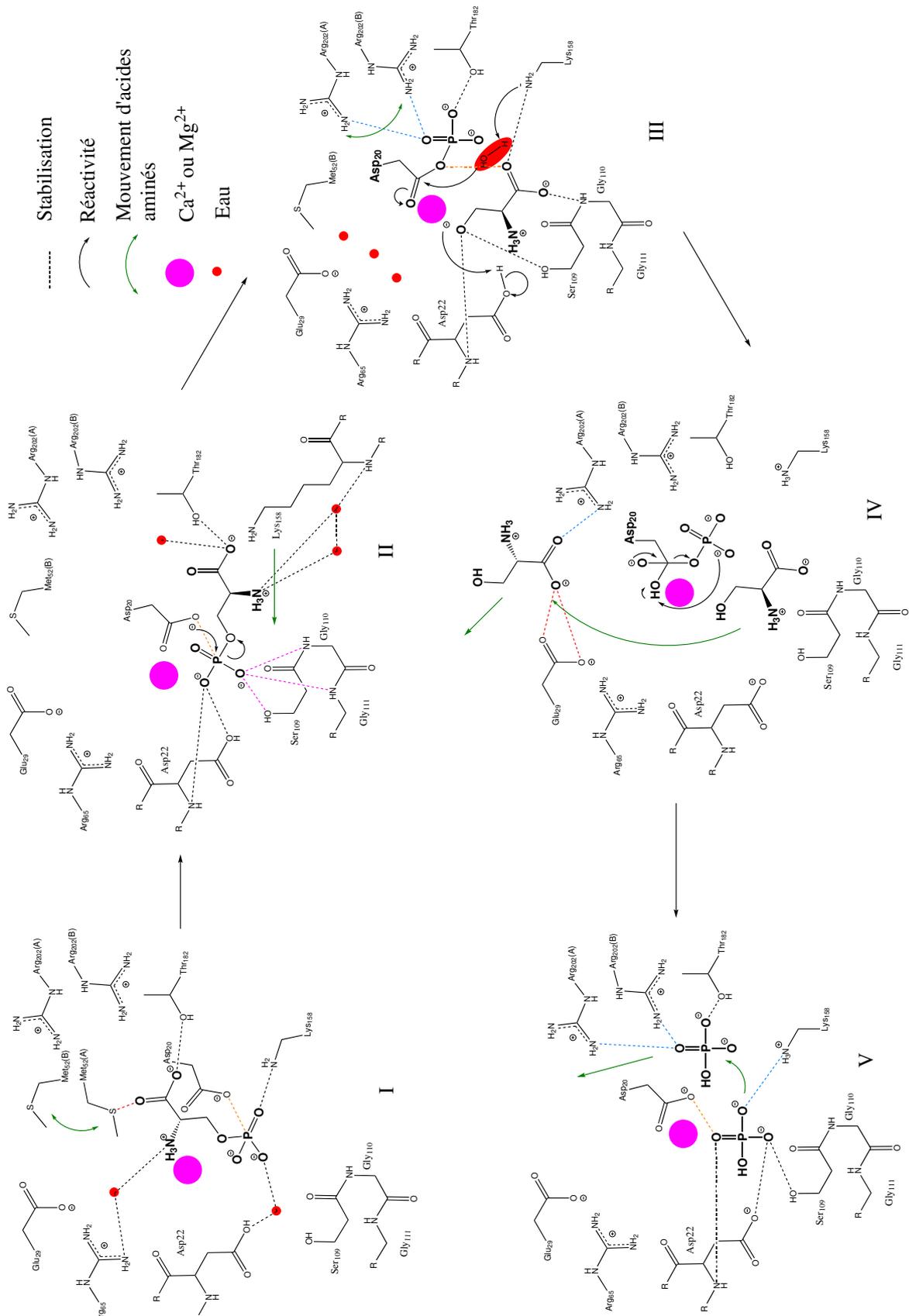


Figure 7.22 – Mécanisme d’HPSP appuyé sur base de l’analyse computationnelle. Les interactions ioniques sont en bleu, les interactions au sein de l’Oxyanion hole en rose et les ponts hydrogènes en noir. Les distances défavorables sont en rouge tandis que les distances pouvant montrer une possibilité de réactivité sont en orange.

Quatrième partie

Conclusions

Chapitre 8

Conclusions et perspectives

À l'heure actuelle, le cancer est une des causes majeures de décès au monde. En Belgique, le cancer du côlon est à la deuxième position chez la femme et à la troisième chez l'homme. Le traitement des patients résulte en une combinaison de chirurgie et de chimiothérapie. La chimiothérapie est réalisée au moyen de 5-Fluorouracile, administré par voie intraveineuse. Ce médicament est un dérivé d'uracile dont le mode d'action est une double stratégie. D'une part, il joue le rôle d'anti-métabolite et empêche la prolifération cellulaire des cellules cancéreuses. D'autre part, le 5-FU augmente la concentration en espèces oxygénées radicalaires dans les cellules tumorales, permettant de provoquer un stress oxydatif causant l'apoptose. Toutefois, un phénomène de résistance au 5-FU tend à apparaître chez les patients. L'implication de différentes enzymes, telles que la phosphosérine phosphatase humaine (HPSP) est mise en évidence dans ce mécanisme de résistance.

HPSP est une enzyme magnésio-dépendante impliquée dans la biosynthèse de la sérine. Elle est responsable de la dernière étape de cette voie métabolique, à savoir la dégradation de phospho-L-sérine en L-sérine. Malgré son caractère de cible thérapeutique, HPSP reste une protéine peu étudiée. Seulement trois structures de HPSP sont recensées dans la PDB et le mécanisme d'action de HPSP reste encore méconnu à l'heure actuelle. Pourtant, pour comprendre la résistance au traitement au 5-FU, il est nécessaire d'en apprendre plus sur cette enzyme qui provoque cette résistance. Ce mémoire consistait en l'étude mécanistique de HPSP, le but étant de mieux comprendre le fonctionnement de cette enzyme avant de pouvoir s'intéresser au problème de chimiorésistance qu'elle provoque, mais aussi de créer de nouveaux inhibiteurs sur base des conclusions mécanis-

tique. Pour ce faire, une stratégie en deux temps a été mise en place. Premièrement, une approche expérimentale axée sur une étude cristallographique de la protéine en présence de ligands a été réalisée afin de mettre en évidence les interactions clés entre les acides aminés du site actif et ces ligands. L'élaboration d'une hypothèse mécanistique a ainsi été faite. En second lieu, une approche computationnelle a été réalisée dans le but d'étudier l'hypothèse mécanistique dans des systèmes proches du système protéique afin de prouver sa faisabilité.

D'un point de vue expérimental, la protéine a d'abord été produite avec succès au moyen d'une induction à l'IPTG pour ensuite être purifiée par IMAC. La protéine récoltée est obtenue pure et en grandes quantités, permettant les essais de cristallisation. Ces essais se sont avérés fructueux, mettant en évidence l'obtention des conditions de cristallisation optimales pour HPSP en présence de phospho-L-sérine et d'un analogue, l'acide homocystéique. Les cristaux ainsi obtenus ont été analysés par diffraction au synchrotron à Paris. Trois structures protéiques sont obtenues après diffraction. La première structure possède deux molécules dans son unité asymétrique, nommées respectivement A et B. La seconde structure est composée d'un dimère, la molécule C et la molécule D. Dernièrement, une troisième structure possède également avec un dimère (molécule E et F). Un ion métallique est observé dans chaque monomère. Dans les molécules A, B, D, E et F, il s'agit de calcium tandis que du magnésium est retrouvé dans le site actif de la molécule C. Des ligands sont aussi observables, à savoir une phospho-L-sérine dans A, une L-sérine et une phospho-L-sérine dans B, une L-sérine et un phosphate dans C, un phosphate dans D et de l'acide homocystéique dans F. Ces six molécules possèdent également une hélice dynamique. Les molécules A, C et E se trouvent en conformation fermée, la boucle étant sous forme d'hélice α . À l'inverse, les molécules B, D et F se trouvent en conformation ouverte, totalement dépliée. Ces différentes conformations permettent aux acides aminés composant la boucle dynamique d'interagir avec les ligands et d'enclencher la réaction. La Met52 joue le rôle de loquet pour fermer l'hélice.

Sur base de l'analyse des interactions établies dans les sites actifs entre les ligands et les acides aminés, une hypothèse mécanistique, décrite à la Figure 8.1, a été envisagée. Au départ de la réaction, la phospho-L-sérine est à l'extérieur de l'enzyme et la boucle dynamique est ouverte. Le substrat entre à l'intérieur de l'enzyme, provoquant le changement de conformation de la boucle

dynamique. L'Asp22, la Lys158 et le Thr182 stabilisent le substrat. La Met52 change de conformation et ferme l'enzyme. Le substrat bouge légèrement et s'approche des aspartates catalytiques. La stabilisation a dès lors lieu grâce à la Ser109, les Gly110 et 111 et la Thr182. Un motif *oxyanion hole* est formé au moyen des résidus Ser109, Gly110 et Gly111. L'Asp20 attaque la phospho-L-sérine et l'hydrolyse a lieu, générant un phosphate lié à l'Asp20 et une molécule de L-sérine déprotonée dans la poche catalytique. Par la liaison du phosphate à l'Asp20, un vide se crée et la L-sérine déprotonée se positionne à proximité de l'Asp22, stabilisée par le Glu29, l'Arg65, la Ser109 et la Lys158. L'Asp22 protone la L-sérine. L'hydrolyse du phosphate par une molécule d'eau activée est réalisée au moyen de la Lys158 ou de l'Asp183. Par changement de conformations de l'hélice dynamique, les deux produits sortent de l'enzyme au moyen de l'Arg202 et, pour la L-sérine, du Glu29.

Cette hypothèse mécanistique a ensuite été étudiée à l'aide de la chimie computationnelle. Pour ce faire, différents modèles ont été analysés afin de mettre en évidence la nécessité d'un système ressemblant à la protéine pour que la réaction ait lieu. Effectivement, la faisabilité des différentes réactions a pu être mise en évidence pour autant que le système approche le système protéique. L'analyse computationnelle a également mis en évidence une activation de la molécule d'eau dans le cas de l'hydrolyse du phosphate plus favorable lorsqu'elle est réalisée par la Lys158 pour autant que cette lysine soit sous forme déprotonée. Cette déprotonation pourrait être réalisée au moyen de l'environnement protéique. Une étude computationnelle sur un modèle reprenant le site actif de HPSP est en cours afin d'évaluer l'évolution des quantités thermodynamiques dans ce modèle et d'apporter une vision plus globale du mécanisme.

Plusieurs perspectives sont envisageables. Premièrement, afin d'en apprendre plus sur l'état de protonation des différents acides aminés dans la poche catalytique, une analyse structurale par diffraction de neutrons pourrait être envisagée. Deuxièmement, l'étude computationnelle dans le modèle correspondant au site actif doit être poursuivie afin d'avoir une vision d'ensemble de l'évolution des quantités thermodynamiques dans la protéine. Enfin, une analyse de dynamique moléculaire sur l'hélice dynamique pourrait être entreprise afin d'étudier le dépliement de celle-ci.

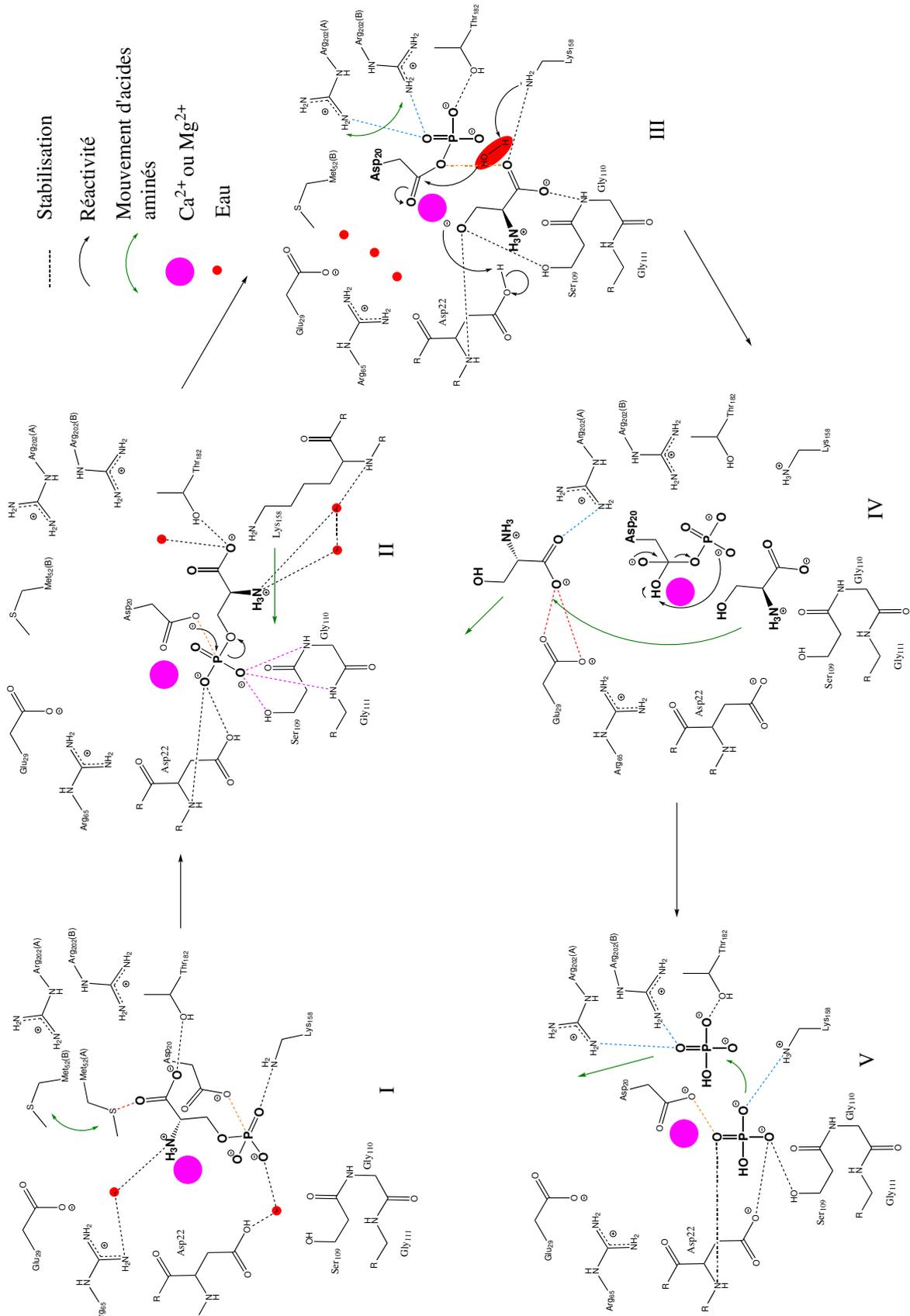


Figure 8.1 – Hypothèse mécanistique au terme de ce mémoire. Les interactions au sein de l'Oxyanion hole en rose et les ponts hydrogènes en noir. Les distances défavorables sont en rouge tandis que les distances favorables sont en orange.

Cinquième partie

Matériels et méthodes

Chapitre 9

Notions théoriques abordées

Au cours du présent mémoire, différentes notions liées à la diffraction de rayons X ou à la chimie computationnelle sont utilisées. Ce chapitre développe ces notions afin d'aider la compréhension des analyses.

9.1 Méthodes liées à la diffraction de rayon X

La diffraction de rayon X (DRX) est une technique d'analyse à partir d'ondes électromagnétiques de faible longueur d'onde, des rayons X (RX), interagissant avec les nuages électroniques des atomes constituant la protéine et provoquant un phénomène de diffusion de Rayleigh. Cette diffusion est un phénomène élastique selon lequel les électrons entrent en vibration et deviennent diffuseurs d'ondes électromagnétiques. Ce phénomène a lieu dans toutes les directions de l'espace et engendre un phénomène d'interférence soit constructive ou destructive, car les différentes ondes générées par les électrons se rencontrent et interfèrent pour créer une nouvelle onde. De par la périodicité du système cristallin, ce phénomène est discontinu et est régi par la loi de Bragg, décrite à l'équation 9.1. Cette loi représente le phénomène de diffraction comme une réflexion sur des plans de réseaux.

$$2d_{hkl} \sin \theta = n\lambda \quad (9.1)$$

où d est la distance entre les différents plans réticulaires (Å), θ est l'angle de diffraction (°), n est un nombre entier et λ est la longueur d'onde utilisée (Å).

Après une expérience de diffraction d'un cristal, un cliché de spots soulignant les angles 2θ de diffraction, tels qu'observable à la Figure 9.1, est obtenu. Ce cliché renseigne de la position des réflexions et de leur intensité. L'intensité donne accès aux facteurs F_o selon la relation 9.2. Le facteur de structure observé est déterminé à partir de ce dernier, selon l'équation 9.3. Il dépend ainsi de la nature et de la position des atomes dans l'unité asymétrique par l'intermédiaire de la phase et de l'amplitude. Après transformées de Fourier, ces informations donnent accès au motif diffractant, à savoir la distribution des électrons dans l'unité asymétrique au moyen d'une carte de densité.⁶⁸

$$I_{hkl} \propto |F_o|^2 \quad (9.2)$$

$$F_o = |F_o|e^{i\phi} \quad (9.3)$$

$$F_c = \sum_j f_j e^{-\frac{B_j \sin^2 \theta}{\lambda^2}} e^{2i\pi(hx_j + ky_j + lz_j)} \quad (9.4)$$

où $e^{i\phi}$ est la phase de la réflexion, f_j et B_j sont respectivement le facteur de forme et le facteur d'agitation thermique de l'atome j dans la maille cristalline. À partir des facteurs de structure observés, F_o , la carte de densité est construite et les électrons peuvent être associés aux atomes d'une molécule. Cependant, pour ce faire, il est nécessaire de connaître la phase des réflexions, donnée non accessible lors le phénomène de diffraction. De ce fait, un calcul de ces dernières par remplacement moléculaire est effectué si la structure possède au moins 30% d'identité avec une autre structure déjà résolue. À l'aide de ce calcul, les facteurs de structure calculés, F_c (équation 9.4), sont isolés ainsi que la carte de densité correspondante, après transformation de Fourier. De là, il est nécessaire d'affiner le modèle pour obtenir la structure protéique.^{67,68,81}

L'affinement d'une structure cristallographique est un processus qui utilise une alternance d'optimisations suivie de corrections manuelles cherchant à positionner au mieux les atomes dans les cartes de densité au moyen des coordonnées x , y et z ainsi que du facteur d'agitation thermique (facteur B) des atomes.⁶⁷ L'affinement permet d'améliorer la carte de densité construite à partir des facteurs de structures calculés, F_c , qui est comparée à la carte construite à partir des facteurs de structure observés, F_o . Une fois la concordance entre les différents types de facteurs de structures jugée satisfaisante, la structure est validée sur base d'indicateurs décrits en partie II.

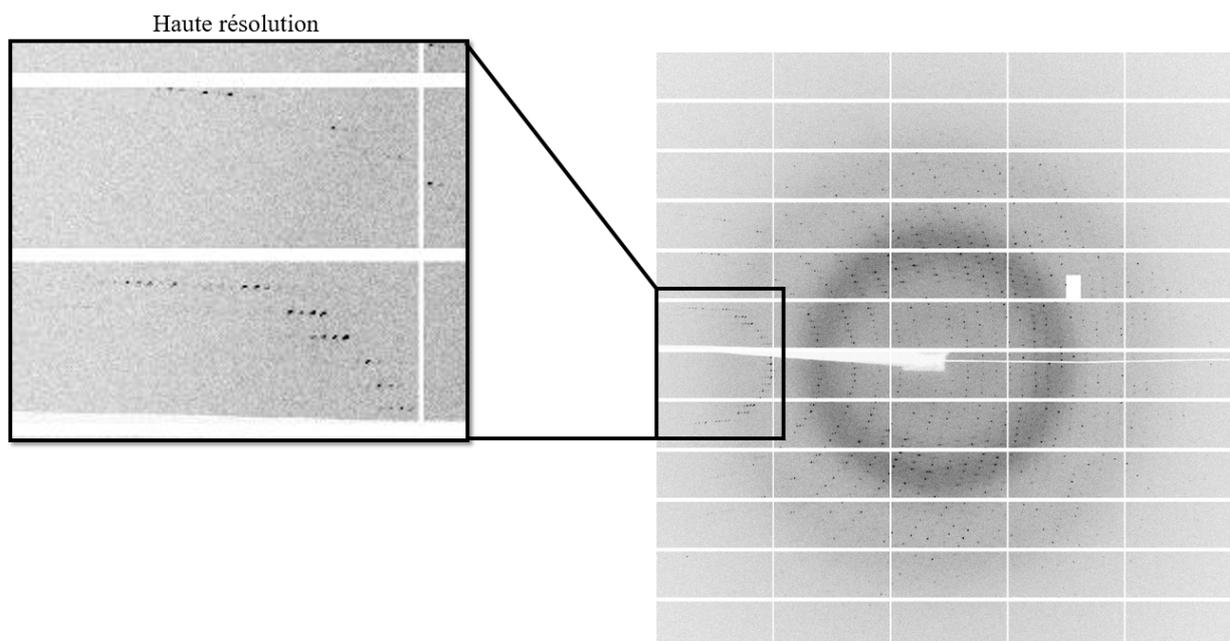


Figure 9.1 – Cliché de diffraction issu de la diffraction par les rayons X du cristal décrit au chapitre 6 obtenu par co-cristallisation avec la phospho-L-sérine, à pH 6,5 et à une concentration en CaCl_2 de 0,05 M dans un tampon cacodylate 0,1 M et de PEG 2000 20%. Résolution de 1,57 Å.

9.2 Méthodes calculatoires

9.2.1 Bases de la chimie quantique

Le point central de la chimie quantique est l'équation de Schrödinger. Cette équation permet de déterminer l'énergie d'un système et la fonction d'onde de ce dernier.^{82,83}

$$\hat{H}\Psi = E\Psi \quad (9.5)$$

où \hat{H} est l'opérateur hamiltonien soulignant les énergies du système à l'aide des opérateurs d'énergie cinétique et potentielle, E est l'énergie du système et Ψ est la fonction d'onde.

Différentes approximations nécessaires sont mises en place afin de résoudre cette équation telles que l'approximation de Born-Oppenheimer impliquant le fait que les mouvements des électrons du système sont plus rapides que les mouvements des noyaux. Ces noyaux sont donc considérés comme figés.⁸⁴ Cette approximation simplifie le calcul. Le théorème des variations,⁸⁵ explicité à l'équation 9.6, est un théorème mis en place pour résoudre cette équation. Ce dernier part du principe que si l'on détermine l'énergie du système avec une fonction d'onde approchée, $\Psi^{Approx.}$,

l'énergie obtenue est supérieure à celle de la fonction d'onde exacte, Ψ , de l'état fondamental. Il faut donc que la fonction d'onde utilisée soit la plus proche de la réalité pour que l'énergie se rapproche de l'énergie exacte.

$$E^{Exact} \leq E^{Approx.} = \langle \Psi^{Approx.} | \hat{H} | \Psi^{Approx.} \rangle \quad (9.6)$$

Dans le but de résoudre l'équation de Schrödinger, différentes méthodes calculatoires ont vu le jour, la plus connue étant la méthode Hartree-Fock (équation 9.7).^{86,87} Dans ces méthodes, la détermination des orbitales moléculaires recourt à l'approximation *Linear Combination of Atomic Orbitals* (LCAO).⁸⁸ Ainsi, les orbitales moléculaires, Φ_i , s'expriment sous la forme de combinaisons linéaires d'orbitales atomiques, χ_k (équation 9.8).

$$\hat{f}(\vec{r})\phi_i(\vec{r}) = \epsilon_i\phi_i(\vec{r}) \quad (9.7)$$

$$\phi_i(\vec{r}) = \sum_k^K C_{ki}\chi_k(\vec{r}) \quad (9.8)$$

où \hat{f} est l'opérateur de Fock, ϕ_i correspond aux orbitales moléculaires tandis que ϵ_i est l'énergie de ces orbitales moléculaires. Les coefficients C_{ki} sont les coefficients LCAO et K est le nombre d'orbitales atomiques. Il est dès lors nécessaire de déterminer ces coefficients LCAO afin d'obtenir une expression des orbitales moléculaires. Cette détermination a lieu au moyen d'un cycle itératif, connu sous le nom de cycle du champ autocohérent [ou *Self-Consistent Field*] (SCF). Afin de faciliter les calculs, les orbitales atomiques sont définies par des contractions de fonctions gaussiennes dont le nombre et la nature déterminent la performance du set de bases utilisé. Plus la base choisie est complète, plus fidèle sera la description des orbitales moléculaires. Dans le cadre de ce mémoire, la base utilisée pour les simulations est 6-311G+(d,p).⁸⁹ Celle-ci contient une contraction de 6 gaussiennes pour décrire les orbitales de cœur, trois contractions respectivement de trois, une et une gaussiennes pour décrire les orbitales de valence, un set de fonctions diffuses et un set de fonctions de polarisations, à savoir de nombre quantique supérieur aux électrons de valence des atomes isolés. Celles-ci sont symbolisées par un "d" et "p" : les fonctions "p" sont ajoutées sur les atomes d'hydrogène tandis que les "d" sont sur les atomes de la deuxième période et les suivantes. Vu que la méthode Hartree-Fock ne tient pas compte de la corrélation électronique, différentes mé-

thodes ont été développées pour pallier ce manque, comme la *density functionnal theory* (DFT).

9.2.2 Principe de la DFT

La méthode de la densité fonctionnelle, ou *density functionnal theory* (DFT) est une méthode variationnelle largement utilisée, recourant généralement à l'approximation LCAO et satisfaisant au théorème des variations.^{89,90} Cette méthode a pour avantage qu'elle permet d'introduire la corrélation électronique.

Le principe de cette théorie est de remplacer la fonction d'onde à N électrons, $\Psi(x_1, x_2, \dots, x_N)$, et l'équation de Schrödinger associée par la densité électronique et les équations nécessaires pour la déterminer.^{90,91} Cette densité se définit comme étant la probabilité de trouver un électron dans un volume dx_1 autour d'une position x_1 . On définit deux types de densité : la densité à un électron, $[\rho(x_1)]$, et à deux électrons, $[\rho(x_1, x_2)]$. La première s'obtient après une intégration de la fonction d'onde sur toutes les coordonnées de spin-espace sauf une (x_1) (équation 9.9). L'intégration de cette densité sur les variables $x_2 \dots x_N$ donne accès à la probabilité de trouver l'électron à la position x_1 . La probabilité à deux électrons correspond, de manière similaire, à la probabilité de trouver deux électrons dans un volume. Cela correspond à la corrélation entre deux électrons en mouvement.⁹²

$$N \int \Psi^*(x_1, x_2, \dots, x_N) \Psi(x_1, x_2, \dots, x_N) dx_2 \dots dx_N = \rho_1(x_1) \quad (9.9)$$

Cette théorie nait de l'approche de Kohn-Sham,⁹³ notamment réalisée à l'aide de différents théorèmes comme le premier théorème de Hohenberg-Kohn.⁹⁴ Le premier théorème de Hohenberg-Kohn relie le potentiel externe à la densité électronique de l'état fondamental ρ_0 . Ce potentiel s'exprime selon l'équation 9.10. Par le théorème des variations appliqué à la DFT, soit le second théorème de Hohenberg-Kohn, il est possible d'approcher l'énergie de l'état fondamental.^{90,94}

$$E_0[\rho_0] = T[\rho_0] + E_{ee}[\rho_0] + E_{ne}[\rho_0] \quad (9.10)$$

où E_0 est l'énergie de l'état fondamental, ρ_0 est la densité de l'état initial, T l'énergie cinétique,

E_{ee} l'énergie électrons-électrons et E_{ne} est l'énergie électrons-noyaux.

La méthode de Kohn-Sham prend son essence dans les théorèmes de Hohenberg-Kohn, mais elle considère un système de particules fictives indépendantes, des électrons, reproduisant la densité électronique du système réel. On définit ainsi des orbitales dites de Kohn-Sham correspondant à des fonctions d'onde à un électron. L'expression de l'énergie peut être décomposée en différents termes (équation 9.11). Cependant, E_{XC} n'étant pas connu, il est approximé au moyen des fonctionnelles d'échange-corrélation. Ce potentiel donne accès au potentiel effectif de Kohn-Sham (équation 9.12).

$$F[\rho(\vec{r})] = T[\rho(\vec{r})] + J[\rho(\vec{r})] + E_{XC}[\rho(\vec{r})] \quad (9.11)$$

$$v_{eff}(\vec{r}) = v(\vec{r}) + \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}' + \frac{\partial E_{XC}}{\partial \rho(\vec{r})} \quad (9.12)$$

où J est le terme de Coulomb et E_{XC} l'énergie d'échange-corrélation. Le terme v_{eff} est le potentiel effectif de Kohn-Sham.

L'équation de Kohn-Sham (équation 9.13) est résolue de manière itérative et donne accès à l'énergie du système. Elle possède une expression parallèle à l'équation d'Hartree-Fock (équation 9.7). La fonction d'onde obtenue permet de calculer une nouvelle valeur de densité jusqu'à satisfaction d'un critère de convergence.⁹³ Si E_{XC} était connu de manière exacte, le potentiel $v_{eff}(\vec{r})$ et l'énergie du système l'aurai été également.

$$\left[\frac{-1}{2} \nabla^2 + v_{eff}(\vec{r}) \right] \theta_n(x) = \epsilon_n \theta_n(x) \quad (9.13)$$

où θ_n est l'orbitale de Kohn-Sham.

9.2.3 Fonctionnelles d'échange-corrélation

Dans le but de calculer la valeur de l'énergie d'échange-corrélation, une fonctionnelle d'échange-corrélation est utilisée.^{90,91,95} Ces fonctionnelles sont classées en plusieurs ensembles en fonction des particularités de leurs spécificités. Dans cette classification, on retrouve, par exemple, les fonc-

tionnelles LDA, GGA ou mGGA. Les LDA, acronyme de *Local density approximation*, se basent sur le principe que les énergies des fonctionnelles d'échange-corrélation d'un système non uniforme peuvent être obtenues en décomposant le système en volumes infinitésimement petit. Selon les GGA, pour *Generalized Gradient Approximation*, les fonctionnelles dépendent de la densité et de son gradient. Dernièrement, les *meta-GGAs*, ou mGGA, introduisent une dépendance explicite de l'énergie cinétique. Au sommet de cette classification, on retrouve les fonctionnelles hybrides. Ces fonctionnelles introduisent un pourcentage d'échange Hartree-Fock exact par rapport aux fonctionnelles des niveaux inférieurs (*GGA*, *mGGA*, ...), par exemple, selon l'équation 9.14.^{90,95}

$$E_{XC} = aE_x^{M06} + (1 - a)E_x^{HF} + \text{Termes de corrélation} \quad (9.14)$$

où a est le pourcentage d'échange Hartree-Fock de la fonctionnelle et l'indice x indique l'échange. Contrairement à l'échange "DFT" des fonctionnelles LDA, GGA et mGGA, qui est local, l'échange Hartree-Fock est non local et présente un comportement asymptotique en $-1/r$. L'avantage de la DFT est qu'elle a un coût en ressources informatiques semblable à la méthode Hartree-Fock tout en introduisant la corrélation électronique. La fonctionnelle utilisée dans le cadre du mémoire est la fonctionnelle hybride M06 qui dérive de la mGGA et possède un terme d'échange Hartree-Fock valant 27%.⁹⁶

9.2.4 Optimisation de géométrie

Durant ce mémoire, différentes molécules chimiques ont été utilisées comme point de départ pour construire des modèles de plus en plus complexes. Ces espèces sont optimisées avec la méthode M06/6-311G+(d,p). Le procédé d'optimisation de géométrie consiste à ajuster les coordonnées atomiques jusqu'à l'obtention d'un minimum d'énergie sur la surface d'énergie potentielle.

9.2.5 Détermination des états de transition et des complexes intermédiaires

Les états de transition sont déterminés par la méthode dite des contraintes. Cela consiste à fixer les distances des liaisons impliquées dans la réaction ciblée à une distance intermédiaire proche de celles de l'état de transition supposé. L'optimisation procède initialement avec ces restrictions ce qui a pour effet de guider l'optimisation afin d'obtenir une première approximation de la géométrie

l'état de transition et une fréquence imaginaire associée. La géométrie de ce premier état sert de point de départ à l'obtention d'une seconde approximation, meilleure, de l'état de transition en suivant le déplacement associé à la fréquence imaginaire, cette fois sans contrainte. Les quantités thermodynamiques (G^0 et H^0) sont extraites après un calcul de fréquences issues de la dérivée seconde de l'énergie.

À partir de l'état de transition, il est possible de déterminer les complexes intermédiaires entre les réactifs ou les produits. Ces complexes témoignent du rapprochement des composés dans l'espace et de la stabilisation de ceux-ci au moyen d'interactions favorables. La détermination de ces complexes est réalisée selon la méthode de l'*Intrinsic Reaction Coordinate* (IRC).⁹⁷ Cela consiste à partir de la géométrie de l'état de transition et de minimiser son énergie pour retrouver les réactifs ou les produits de départ tout en suivant les déplacements atomiques associés à la fréquence imaginaire.⁹¹

9.2.6 Thermodynamique statistique

Par l'approximation de Born-Oppenheimer (noyaux figés), l'optimisation permet l'obtention de l'énergie du système telle une somme des répulsions entre noyaux et électrons. Toutefois, les atomes possèdent un mouvement résiduel, même à 0K. Il ne s'agit donc pas de la seule contribution énergétique au système : les énergies de vibration, de translation et de rotation ont aussi une incidence. L'étude thermodynamique est possible par l'évaluation d'une fonction de partition Ω , dont l'expression est mentionnée à l'équation 9.15. Cette fonction de partition caractérise la tendance du système à vouloir s'échapper à l'état fondamental. Les fonctions d'état telles que l'enthalpie (H), l'énergie libre (U), l'enthalpie libre de Gibbs (G) ou encore l'entropie (S) peuvent être déterminées à partir de cette dernière.⁸³

$$\Omega = \sum_i g_i e^{-\beta \epsilon_i} \quad \text{où } \beta = \frac{1}{k_B T} \quad (9.15)$$

où i décrit l'état i du système, ϵ_i est l'énergie de cet état, g_i la dégénérescence du système et β est un facteur dérivant de la constante de Boltzmann (k_B).

9.2.7 La solvation

La méthode dite de *Integral equation formalism of the polarizable continuum model*^{98,99} (IEF-PCM) a été employée dans le but d'introduire les effets du solvant lors des simulations. Cela permet de tenir compte des interactions entre le solvant et le soluté. Ce dernier est placé dans une cavité entourée d'un diélectrique, simulant le solvant. Par effet électrostatique, la cavité et le soluté vont se polariser (Figure 9.2) : une différence de potentiel apparait. La distribution des charges entre les deux milieux peut être reliée à une valeur d'énergie, correspondant aux interactions soluté-solvant.¹⁰⁰ La détermination des interactions soluté-solvant se fait via un processus itératif.

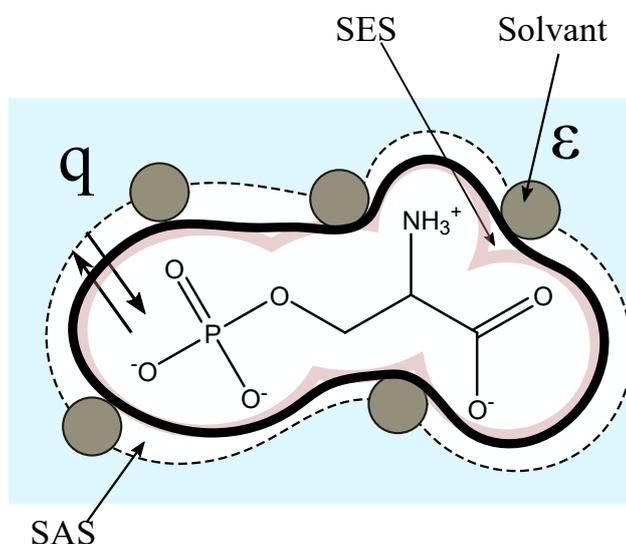


Figure 9.2 – Principe de la méthode IEF-PCM. Q représente la charge, ε le diélectrique du solvant, SAS et SES les méthodes de solvation.⁹⁹

Lors de l'élaboration de la cavité entourant les atomes, deux types de surface peuvent être utilisées, à savoir la SES et la SAS. La *Solvent Excluded Surface* (SES)⁹⁹ considère les murs de la cavité comme étant frontières entre le soluté-solvant. C'est la surface utilisée par défaut. Les molécules y sont considérées comme étant dans une cavité de solvant construite à partir de sphères de van der Waals placées sur chaque atome composant la molécule à l'aide d'un champ de force de type *Universal Force Field* (UFF).¹⁰¹ La *Solvent Accessible Surface* (SAS)⁹⁹ définit la surface accessible au solvant. Étant un peu moins réaliste, celle-ci va être utilisée dans le cas de grands systèmes, celui-ci simulant l'enzyme. Dans tous nos calculs, le solvant est l'eau. La constante diélectrique (ε) s'élève donc à 78,355300 à 25 °C tandis que la constante diélectrique à l'infini (ε_∞) est de 1,777849 à fréquence infinie et à 25 °C.

Chapitre 10

Matériel et méthodes

10.1 Production et purification

10.1.1 Préparation des solutions

Lors de la phase de production et de purification, différentes solutions sont réalisées. Il s'agit de la solution de lyse ainsi que des tampons A et B. Les informations portant sur la composition de ces solutions sont décrites au Tableau 10.1.

Tableau 10.1 – Récapitulatif des volumes des différentes solutions réalisées lors de la production et de la purification de HPSP. Informations sur les concentrations et les masses des différents composés utilisés. Le solvant est l'eau.

| Solutions | Volume (mL) | pH | Composé | Concentration (mM) | Masse (g) |
|-------------------|-------------|-----|--------------------|--------------------|-------------|
| Milieu de culture | 400 | / | LB liquide | / | 8 |
| Solution de lyse | 100 | 7,4 | NaCl | 100 | 0,584 |
| | | | HEPES | 25 | 0,596 |
| | | | EDTA-free protease | / | 2 pastilles |
| Tampon A | 250 | 7,4 | NaCl | 100 | 1,461 |
| | | | HEPES | 25 | 1,489 |
| | | | Imidazole | 10 | 0,170 |
| Tampon B | 250 | 7,4 | NaCl | 100 | 1,461 |
| | | | HEPES | 25 | 1,489 |
| | | | Imidazole | 250 | 4,255 |

10.1.2 Préculture et culture

Afin de produire HPSP de manière optimale, une préculture est réalisée en milieu stérile. Dans un falcon de 50 mL, 10 mL de la solution de milieu de culture (Tableau 10.1) sont mélangés à 10 μ L de Kanamycine™ à 50 mg mL⁻¹. Après cela, 200 μ L de souches BL21 (DE3) d'*E. coli*, transformées avec le plasmide codant pour HPSP, sont ajoutés au mélange. Le tout est laissé à 37 °C durant une nuit, sous agitation. Après une nuit, le falcon contenant la préculture est utilisé afin de réaliser la culture bactérienne. Pour ce faire, 200 mL de milieu de culture sont placés dans un erlenmeyer à défecteur en milieu stérile. À cela est ajouté 200 μ L de Kanamycine™ à 50 mg mL⁻¹. Une mesure de la densité optique est réalisée pour servir de blanc (temps 0). Dix microlitres de préculture sont ajoutés au contenu de l'erlenmeyer et une mesure de densité optique est réalisée afin de connaître l'absorbance au début de l'incubation (temps 1). Le tout est laissé sous incubation, à 37 °C, jusqu'à obtention d'une valeur comprise entre 0,6 et 0,8 pour la densité optique.

10.1.3 Induction

Une fois cette densité optique atteinte, l'induction de la protéine est provoquée par ajout, sous conditions stériles, d'IPTG 1 M au mélange afin d'atteindre une concentration finale de 4 mM. Le mélange est replacé dans l'incubateur, à 37 °C, pendant 3 heures, sous agitation. Une fois ce laps de temps écoulé, le contenu de l'erlenmeyer est centrifugé à l'aide d'une centrifugeuse Beckman à 4000 rpm, durant 30 minutes, à 4 °C. Les culots sont ensuite récupérés et congelés avant d'être lysés.

10.1.4 Lyse

Les culots sont suspendus dans un volume minimum de tampon à partir de la solution de lyse décrite au Tableau 10.1. La solution ainsi obtenue est lysée sur glace par sonification à l'aide d'une alternance de 30 secondes de sonification et 30 secondes de repos à 10 reprises. La puissance utilisée est de 10 watts. Une fois la lyse réalisée, la solution résultante est centrifugée à l'aide d'une centrifugeuse Beckman à 12 500 rpm, durant 60 minutes, à 4 °C. Le surnageant contenant HPSP non-purifié est récupéré. Il est filtré à l'aide d'un filtre *Whatman* 0,45 μ m et d'un filtre 0,2 μ m.

10.1.5 Purification par IMAC

La purification est réalisée par IMAC, à l'aide d'une colonne de type HiTrapTMFF Crude 1mL (GE Healthcare) et d'un appareillage de type *Fast Purification Liquid Chromatography* (FPLC) *Healthcare-FPLC BioRad*. Ce dernier permet un suivi de la purification à l'aide de l'absorbance et la conductivité.

10.1.5.1 Conditionnement de la colonne

Avant de commencer la purification, la colonne, chargée à l'aide d'une solution 0,1 M de NiSO₄, est placée, selon un débit de 1 mL min⁻¹, dans 100% de tampon A (Tableau 10.1) jusqu'à l'option d'un plateau de conductivité et d'absorbance sur le suivi par chromatogramme.

10.1.5.2 Éluion de la protéine

Le surnageant contenant le mélange protéique est chargé sur la colonne, toujours conditionnée avec le tampon A. Une fois l'éluion du *Flow Through* observée ainsi qu'une stabilisation de l'absorbance, le tampon conditionnant la colonne est modifié progressivement à l'aide d'un gradient de 0% à 100% en tampon B. Ce gradient est réalisé à l'aide de différents paliers. Durant les dix premières minutes, le pourcentage de tampon B va augmenter de 0% à 10% suivi de 20 minutes avec un pourcentage constant de 10%. Durant les 20 minutes qui suivent, le tampon B augmente de 10% à 100% jusqu'à la cinquantième minute. Un second palier est ensuite appliqué durant dix minutes, marquant la fin de la purification.

10.2 Suivi de purification

10.2.1 Électrophorèse SDS-PAGE

La réalisation d'une électrophorèse sur gel de type SDS-PAGE nécessite la préparation de ces gels d'électrophorèse. On en dénombre deux, à savoir le *running gel* et le *stacking gel* (décrits dans le tableau 10.2) ainsi qu'une série de solutions décrites dans le tableau 10.3. Le *running gel* est d'abord réalisé par ajout des composés décrits au Tableau 10.2, partie A. Le mélange est placé dans une cassette d'électrophorèse afin de la remplir aux trois quarts. Le tout est recouvert

d'une couche de norvanol afin d'éviter l'oxydation du gel durant la polymérisation. Une fois le gel polymérisé, le mélange du *stacking gel* est réalisé selon les conditions décrites au Tableau 10.2, partie B. Aussitôt le mélange effectué, le norvanol est évacué afin de pouvoir couler le second gel au-dessus du premier. Un peigne est ensuite placé au-dessus de ce dernier afin de permettre la formation de puits pour contenir les différents échantillons protéiques.

Tableau 10.2 – Composition des deux types de gels utilisés lors de la préparation d'un gel d'électrophorèse. L'ordre des composés est respecté durant la préparation du gel afin d'ajouter en dernier les deux catalyseurs de polymérisation (APS et TEMED) (**A** Running gel **B** Stacking gel).

| A | | B | |
|------------------------------|-------------|------------------------------|-------------|
| Running gel | Volume (mL) | Stacking gel | Volume (mL) |
| Eau distillée | 5,085 | Eau distillée | 6,00 |
| Tampon Tris-HCl pH 8,8 1,5 M | 3,75 | Tampon Tris-HCl pH 6,8 0,5 M | 2,52 |
| SDS | 0,150 | SDS | 0,100 |
| Acrylamide 30% | 5,84 | Acrylamide 30% | 1,320 |
| APS 10% | 0,075 | APS 10% | 0,050 |
| TEMED | 0,0075 | TEMED | 0,010 |

Tableau 10.3 – Composition des différentes solutions utilisées durant l'électrophorèse SDS-PAGE.

| Solutions | Composés | Volume (mL) | Masse (g) |
|--------------------------|--------------------------|-------------|-----------|
| Tampon d'électrophorèse | Eau distillée | 3,550 | / |
| | Tris 0,5 M pH = 6,8 | 2,500 | / |
| | SDS 10% | 2,000 | / |
| | Glycérol | 2,500 | / |
| | Bleu de bromophénol 0,5% | 0,200 | / |
| | β -mercaptoéthanol | 0,500 | / |
| Solution de coloration | Bleu de Coomassie | / | 0,25 |
| | Méthanol | 45 | / |
| | Acide acétique | 10 | / |
| | Eau distillée | 45 | / |
| Solution de décoloration | Méthanol | 300 | / |
| | Acide acétique | 100 | / |
| | Eau distillée | 400 | / |

Les différents échantillons protéiques sont préparés en prélevant 15 μL de chaque fraction issue de la purification et en les plaçant dans un eppendorf avec 5 μL de tampon d'électrophorèse (Tableau 10.3). Chaque échantillon est chauffé à 100 $^{\circ}\text{C}$, à l'aide d'un bain-marie, suivi d'un traitement sous agitation grâce au vortex. Après cela, 15 μL de chaque eppendorf sont placés dans un puit du gel. Le premier puit contient le marqueur de poids moléculaire. La migration s'effectue durant 40 minutes, sous 220 volts. Par la suite, les gels sont rincés à l'eau et placés dans une solution de coloration (Tableau 10.3). La coloration est réalisée durant une heure. Le gel est ensuite décoloré durant 20 minutes à l'aide de la solution de décoloration (Tableau 10.3).

10.2.2 Test d'activité

L'activité enzymatique des fractions est vérifiée après purification pour savoir si la protéine récoltée est obtenue sous forme active. Le principe du test d'activité est repris à la Figure 10.1. Dans un premier temps, les différentes fractions provenant de la purification sont ramenées le plus proche d'une concentration de 50 mg mL^{-1} . La composition de ce tampon est de 100 mM de tampon Tris-HCl à pH 7,4, 5 mM de MgCl_2 et 1 mM de DTT ajouté au dernier moment.

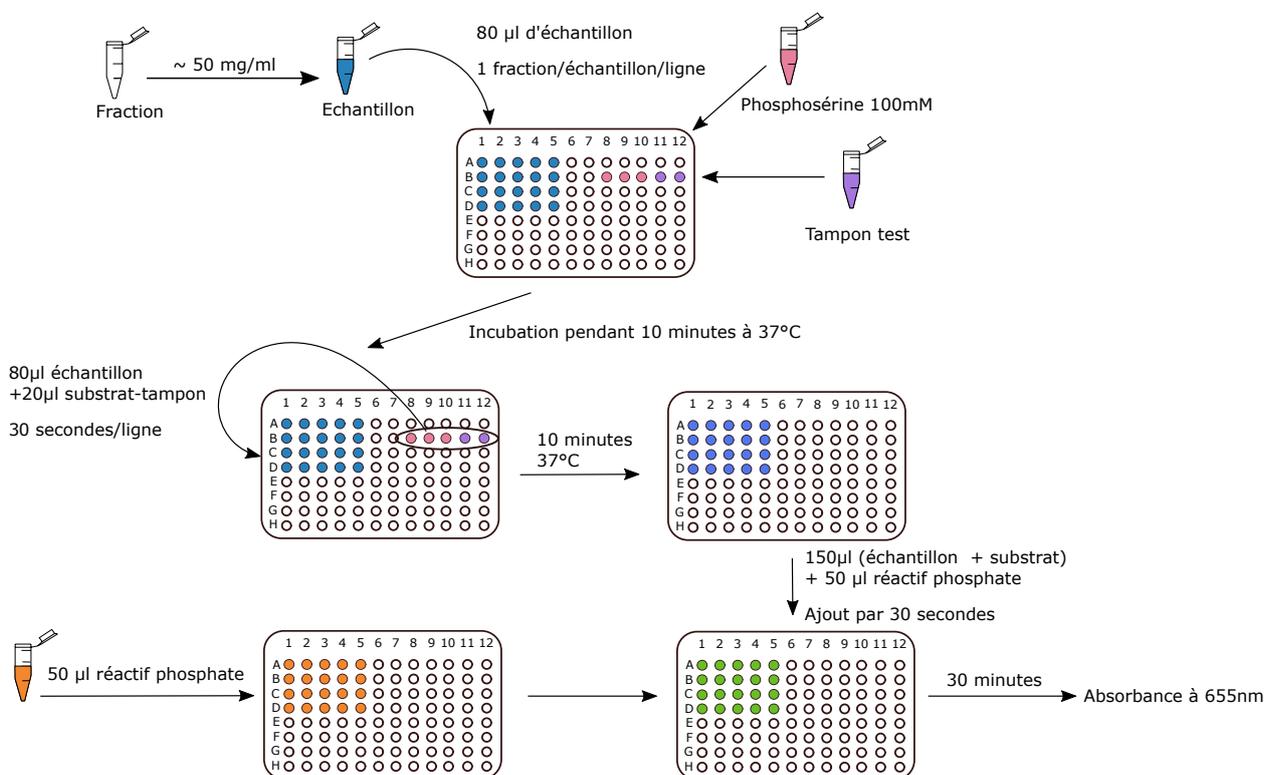


Figure 10.1 – Schéma représentatif du test d'activité de HPSP.

Sur une plaque de 96 puits, 80 μL de chaque échantillon sont placés dans cinq aliquots de manière à faire des réplicats. Une ligne composée d'un triplicat de phospho-L-sérine 100 mM ainsi que deux aliquots de tampon test, servant de blanc, est aussi réalisée, nommé ligne substrat-tampon. Le porte échantillon est placé dans un bain-marie durant 10 minutes à 37 °C. Après incubation, 20 μL de la ligne substrat-tampon sont ajoutés, par intervalle de 30 secondes, aux cinq réplicats d'une même ligne à l'aide d'une pipette multichannelle. Le tout est laissé dix minutes supplémentaires dans le bain, le temps que la réaction opère. À l'aide d'une pipette multipuits, 150 μL des cinq réplicats d'une même ligne sont prélevés et placés dans cinq puits d'une ligne d'une plaque 96 puits contenant chacun 50 μL de la solution "réactif phosphate", composée de 3,75 mL de vert de malachite et 1,25 mL de mobylate d'ammonium. Un intervalle de temps de 30 secondes entre le prélèvement de chaque ligne est respecté pour que le temps de réaction soit identique pour chaque ligne. Après 30 minutes de repos pour que la solution se stabilise, l'absorbance de la plaque est mesurée à 655 nm.

10.2.3 Dosage enzymatique

Le dosage de la concentration en protéine récoltée après purification est réalisé par mesure au Nanodrop™2000. Ceci est un spectromètre permettant une mesure d'absorbance à 280 nm à partir de 2 μL de la solution protéique. Afin de permettre le calcul de la concentration protéique, le poids moléculaire (25,007 kDa) ainsi que le coefficient d'extinction molaire de HPSP sont nécessaires (11585 L M cm⁻¹).^{102,103}

10.3 Cristallisation en présence de ligands

10.3.1 Préparation des solutions pour la cristallisation

Lors de la cristallisation de HPSP, différentes solutions stock sont nécessaires, à savoir une solution cacodylate 0,2 M, une solution de CaCl₂ 2 M ainsi qu'une solution de PEG 2000 monométhyl 50%. D'autre part, afin de permettre la co-cristallisation de HPSP avec différents ligands, il est nécessaire de préparer des solutions de ces derniers. Une solution stock de phospho-L-sérine 100 mM est effectuée dans l'eau ainsi qu'une solution avec un analogue non hydrolysable de sub-

strat, l'acide homocystéique dans le DMSO (100%). L'ensemble des solutions sont préparées avec les compositions décrites au Tableau 10.4.

Tableau 10.4 – Récapitulatif des volumes des différentes solutions réalisées lors de la cristallisation de HPSP. Informations sur les concentrations et les masses des différents composés utilisés. Le solvant est l'eau sauf pour l'acide homocystéique.

| Solutions | Volume (mL) | Composé | Concentration | Masse (g) |
|----------------------|-------------|--------------------|---------------|-----------|
| Cacodylate de sodium | 50 | $C_2H_7AsO_2$ | 0,2 M | 8,56 |
| Chlorure de calcium | 50 | $CaCl_2$ | 2 M | 11,10 |
| PEG 2000 monomethyl | 50 | $(C_2H_4O)_{2000}$ | 50% | 25,0 |
| Phospho-L-sérine | 10 | $C_3H_8NO_6P$ | 100 mM | 0,1851 |
| Acide homocystéique | 1 | $C_4H_9NO_5S$ | 200 mM | 0,0353 |

10.3.2 Préparation de la solution enzymatique : PD10 et aliquots

Après purification, les fractions contenant la protéine d'intérêt sous forme pure sont réunies et concentrées au moyen d'une centrifugeuse *Centrifuge 5430 R* à 7000 rpm et 4 °C par intervalles de 5 minutes afin d'atteindre un volume de 2,5 mL. Cela est possible à l'aide d'un filtre amicon *Vivaspin (10000 MW)* pouvant être centrifugé. Une fois le volume atteint, une PD10 est réalisée afin de placer la protéine dans le tampon utilisé lors de la cristallisation, à savoir un tampon glycine 25 mM à pH 9,0 contenant 10 mM de DTT. Dans un premier temps, la PD10 est conditionnée avec ce tampon au moyen de cinq rinçages de 5 mL. La solution protéique de 2,5 mL est ensuite placée sur la colonne suivie d'une fraction de 3,5 mL de tampon. Cette fraction est ensuite récoltée : elle contient la protéine dans le tampon adéquat. La colonne est ensuite conservée dans l'éthanol 20 % tandis que la fraction protéique récoltée est concentrée par centrifugation à l'aide d'une *Centrifuge 5430 R* et de filtres *Vivaspin Centrifugal Concentrator (10000 MW)* à 7000 rpm et 4 °C durant des intervalles de 5 minutes pour atteindre une concentration avoisinant 6 mg mL⁻¹.

Lors de la co-cristallisation avec la phospho-L-sérine, un aliquot contenant la protéine concentrée à 6 mg mL⁻¹, une solution de MgCl₂ 5 mM et une concentration en substrat, la phospho-L-sérine, 10 mM est réalisé. Une incubation de 30 minutes est respectée avant la réalisation des boîtes de cristallisation. À la différence du cas précédent, lors de la co-cristallisation avec l'acide homocystéique, l'aliquot contient la protéine concentrée à 6 mg mL⁻¹ et la solution de MgCl₂ 5 mM.

L'analogue non hydrolysable de substrat est placé directement dans le puits et mis en contact avec la protéine seulement lors de la formation de la goutte afin d'éviter un choc trop conséquent entre la protéine et le DMSO.

10.3.3 Compositions des puits / co-cristallisation avec la phospho-L-sérine

Lors de la co-cristallisation avec la phospho-L-sérine, la solution de cacodylate, de CaCl_2 ainsi que de PEG 2000 sont placées dans les puits selon les conditions suivantes :

- Cacodylate de sodium : 500 μL afin d'atteindre une concentration constante de 0,1 M. Le pH varie de 5,5 à 7,0.
- Chlorure de calcium : des volumes variables de 25 μL à 150 μL afin d'atteindre une concentration variant de 0,05 M à 0,30 M.
- PEG 2000 : 400 μL afin d'atteindre une concentration constante de 20%.
- De l'eau pour ajuster le volume à un millilitre lorsque cela s'avère nécessaire.

10.3.4 Compositions des puits/ co-cristallisation avec de l'analogue

Lors de la co-cristallisation avec l'analogue non hydrolysable de substrat, les puits de cristallisation ont été remplis selon les conditions précédemment décrites tout en ajoutant 4% de la solution d'acide homocystéique dans le puits. Les concentrations ainsi obtenues sont les suivantes :

- Cacodylate de sodium : 500 μL afin d'atteindre une concentration constante de 0,1 M. Le pH varie de 5,5 à 7,0.
- Chlorure de calcium : des volumes variables de 25 μL à 150 μL afin d'atteindre une concentration variant de 0,05 M à 0,30 M.
- PEG 2000 : 400 μL afin d'atteindre un pourcentage constant de 20%.
- De l'eau pour ajuster le volume à un millilitre lors que cela s'avère nécessaire.
- Acide homocystéique : 40 μL d'une solution 200 mM, 100% en DMSO, afin d'atteindre un pourcentage en DMSO constant de 4% dans le puits, soit une concentration en acide homocystéique de 8,0 mM.

10.4 La diffraction au synchrotron

10.4.1 Cryo-protection

Afin de préparer les différents échantillons pour l'analyse au synchrotron, une cryoprotection est nécessaire. Dans ce but, les cristaux sont piochés et placés dans une nouvelle goutte contenant 1,5 μL de leur solution respective de puit et 1,5 μL de PEG 400 50%. Les cristaux sont laissés dans cette goutte durant 5 minutes avant d'être congelés de manière rapide à l'aide de l'azote liquide.

10.4.2 Collecte des données de diffraction

Les données de diffraction sont collectées à 100 K en utilisant les radiations du synchrotron SOLEIL sur la ligne PROXIMA 1 (PX1) en utilisant un détecteur PILATUS 6M (Dectris Ltd.) à la longueur d'onde de 0,97857 Å. L'affinement est réalisé à l'aide des programmes PHENIX et Coot. L'attribution des éléments de structures secondaires a été réalisée avec le programme DSSP.^{71 69,70}

10.5 Analyse computationnelle

Les calculs quantiques ont été réalisés en utilisant la méthode de la fonctionnelle de densité (DFT) M06 implantée dans Gaussian16.¹⁰⁴ L'optimisation de géométrie a été réalisée avec la base 6-311+G(d,p) pour les modèles les plus simples et avec la base 6-311G(d) pour le modèle enzymatique. Les effets de solvation ont été introduits de manière implicite, selon la méthode *Integral equation formalism -polarizable continuum model* (IEF-PCM). La solvation du modèle enzymatique a été réalisée avec un modèle SAS. L'observation des structures est réalisée avec les programmes GaussView, Drawmol et Avogradro.¹⁰⁵⁻¹⁰⁷

Sixième partie

Bibliographie

Bibliographie

- [1] Reece, J. B.; Urry, L. A.; Cain, M. L.; Wasserman, S. A.; Minorsky, P. V.; Jackson, R. B., et al. *Campbell biology*; Pearson Boston, 2014.
- [2] Cancer Research UK. <http://www.cancerresearchuk.org>, site web en ligne. Consulté le 13 octobre 2018.
- [3] Organisation Mondiale de la Santé (OMS). <http://search.who.int>, site web en ligne. Consulté le 13 octobre 2018.
- [4] Belgian Cancer Registry. Les 10 tumeurs les plus fréquentes par sexe, <https://kankerregister.org>, site web en ligne. Consulté le 18 octobre 2018.
- [5] Larichesse.com, <http://larichesse.over-blog.com>, site web en ligne. Consulté le 22 octobre 2018.
- [6] Xu, J.; Zhang, S.; Wang, R.; Wu, X.; Zeng, L.; Fu, Z. *Bioscience Reports* **2017**, *37*, 1–9.
- [7] Sato, K. et al. *Anticancer Research* **2017**, *37*, 2365–2371.
- [8] Van Cutsem, E. et al. *New England Journal of Medicine* **2009**, *360*, 1408–1417.
- [9] Fondation contre le cancer. <https://www.cancer.be>, site web en ligne. Consulté le 13 octobre 2018.
- [10] Winawer, S. et al. *Gastroenterology* **1997**, *112*, 594–642.
- [11] Fondation ARC pour la recherche sur le Cancer. <https://www.fondation-arc.org>, site web en ligne. Consulté le 13 octobre 2018.
- [12] Alberts, S. R.; Horvath, W. L.; Sternfeld, W. C.; Goldberg, R. M.; Mahoney, M. R.; Dakhil, S. R.; Levitt, R.; Rowland, K.; Nair, S.; Sargent, D. J.; Donohue, J. H. *Journal of Clinical Oncology* **2005**, *23*, 9243–9249.
- [13] Gouvernement du Canada : Comment le cancer est-il traité? <https://www.canada.ca/fr/sante-publique>, site web en ligne. Consulté le 13 octobre 2018.
- [14] Longley, D. B.; Harkin, D. P.; Johnston, P. G. *Nature Reviews Cancer* **2003**, *3*, 330–338.
- [15] Grem, J. L. *Investigational New Drugs* **2000**, *18*, 299–313.
- [16] Goette, D. K. *Journal of the American Academy of Dermatology* **1981**, *4*, 633–649.
- [17] Assenat, E. *Future Medicine* **2012**, *1*, 181–184.
- [18] Hwang, P. M.; Bunzu, F.; Yu, J.; Rago, C.; Chan, T.; Murphy, M.; Kelso, G. F.; Smith, R. A. J.; Kinzler, K. W.; Vogelstein, B. *Nature Medecine* **2001**, *53*, 1111–1117.
- [19] Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Research* **2000**, *28*, 235–242.
- [20] Amelio, I.; Cutruzzolá, F.; Antonov, A.; Agostini, M.; Melino, G. *Trends in Biochemical Sciences* **2014**, *39*, 191–198.
- [21] Li, X.; Xun, Z.; Yang, Y. *Biochemical and Biophysical Research Communications* **2016**, *477*, 633–639.
- [22] Bahlis, N. J.; McCafferty-Grad, J.; Jordan-McMurry, I.; Neil, J.; Reis, I.; Kharfan-Dabaja, M.;

- Eckman, J.; Goodman, M.; Fernandez, H. F.; Boise, L. H.; Lee, K. P. *Clinical Cancer Research* **2002**, *8*, 3658–3668.
- [23] Hwang, I. T.; Chung, Y. M.; Kim, J. J.; Chung, J. S.; Kim, B. S.; Kim, H. J.; Kim, J. S.; Yoo, Y. D. *Biochemical and Biophysical Research Communications* **2007**, *359*, 304–310.
- [24] Chung, Y. M.; Park, S.; Park, J. K.; Kim, Y.; Kang, Y.; Yoo, Y. D. *Cancer Letters* **2000**, *159*, 95–101.
- [25] The Human Protein Atlas, PSPH, <https://www.proteinatlas.org>, site web en ligne. Consulté le 18 octobre 2018.
- [26] Bachelor, M. A.; Lu, Y.; Owens, D. M. *Journal of Dermatological Science* **2011**, *63*, 164–172.
- [27] WebPlotDigitizer, <https://apps.automeris.io/wpd/>, site web en ligne. Consulté le 20 octobre 2018.
- [28] Koch, G. A.; Eddy, R. L.; Haley, L. L.; Byers, M. G.; McAvoy, M.; Shows, T. B. *Cytogenetic and Genome Research* **1983**, *35*, 67–69.
- [29] Moro-Furlani, A. M.; Turner, V. S.; Hopkinson, D. A. *Annals of Human Genetics* **1980**, *43*, 323–334.
- [30] Jaeken, J. *Inborn metabolic diseases diagnosis and treatment [Internet]*; Saudubray, J. M. and Van Den Berghe, G. and Walter, J.H, 2012; pp 357–362.
- [31] De Koning, T. J.; Snell, K.; Duan, M.; Berger, R.; Poll-the, B.-T.; Surtees, R. *Biochemical Journal* **2003**, *371*, 653–661.
- [32] Douce, R.; Bourguignon, J.; Neuburger, M.; Rébeillé, F. *Trends in Plant Science* **2001**, *6*, 167–176.
- [33] Gottlieb, E.; Vousden, K. H. *Cell Death and Differentiation* **2017**, *24*, 193–194.
- [34] Veiga-da Cunha, M.; Collet, J. F.; Prieur, B.; Jaeken, J.; Peeraer, Y.; Rabbijns, A.; Van Schaftingen, E. *European Journal of Human Genetics* **2004**, *12*, 163–166.
- [35] El-Hattab, A. W. *Molecular Genetics and Metabolism* **2016**, *118*, 153–159.
- [36] Ros, R.; Muñoz-Bertomeu, J.; Krueger, S. *Trends in Plant Science* **2014**, *19*, 564–569.
- [37] Mullarky, E. et al. *Proceedings of the National Academy of Sciences* **2016**, *113*, 1778–1783.
- [38] Tabatabaie, L.; Klomp, L. W.; Berger, R.; de Koning, T. J. *Molecular Genetics and Metabolism* **2010**, *99*, 256–262.
- [39] Ho, C.-I.; Noji, M.; Saito, K. *Journal of Biological Chemistry* **1999**, *274*, 11007–11012.
- [40] Seifried, A.; Schultz, J.; Gohla, A. *FEBS Journal* **2013**, *280*, 549–571.
- [41] Locasale, J. W. *Nature Reviews Cancer* **2013**, *13*, 572–583.
- [42] Deberardinis, R. J. *Cell Metabolism* **2011**, *14*, 285–286.
- [43] Jaeken, J.; Detheux, M.; Fryns, J. P.; Collet, J. F.; Alliet, P.; Van Schaftingen, E. *Journal of Medical Genetics* **1997**, *34*, 594–6.
- [44] Peeraer, Y.; Rabijns, A.; Verboven, C.; Collet, J. F.; Van Schaftingen, E.; De Ranter, C. *Acta Crystallographica Section D : Biological Crystallography* **2002**, *58*, 133–134.
- [45] Peeraer, Y.; Rabijns, A.; Verboven, C.; Collet, J. F.; Van Schaftingen, E.; De Ranter, C. *Acta Crystallographica Section D : Biological Crystallography* **2003**, *59*, 971–977.
- [46] Peeraer, Y.; Rabijns, A.; Collet, J.-F.; Van Schaftingen, E.; De Ranter, C. *European Journal of Biochemistry* **2004**, *271*, 3421–3427.
- [47] Collet, J. F.; Stroobant, V.; Van Schaftingen, E. *Journal of Biological Chemistry* **1999**, *274*, 33985–33990.
- [48] Wang, W.; Cho, H. S.; Kim, R.; Jancarik, J.; Yokota, H.; Nguyen, H. H.; Grigoriev, I. V.; Wemmer, D. E.; Kim, S. H. *Journal of Molecular Biology* **2002**, *319*, 421–431.
- [49] Collet, J. F.; Gerin, I.; Rider, M. H.; Veiga-da Cunha, M.; Van Schaftingen, E. *FEBS Letters* **1997**,

- 408, 281–284.
- [50] Kim, H. Y. et al. *Journal of Biological Chemistry* **2002**, 277, 46651–46658.
- [51] Cho, H.; Wang, W.; Kim, R.; Yokota, H.; Damo, S.; Kim, S. H.; Wemmer, D.; Kustu, S.; Yan, D. *Proceedings of the National Academy of Sciences of the United States of America* **2001**, 98, 8525–30.
- [52] Danko, S.; Yamasaki, K.; Daiho, T.; Suzuki, H. *Journal of Biological Chemistry* **2004**, 279, 14991–14998.
- [53] Yang, L.; Liao, R. Z.; Ding, W. J.; Liu, K.; Yu, J. G.; Liu, R. Z. *Theoretical Chemistry Accounts* **2012**, 131, 1–12.
- [54] Allen, K. N.; Dunaway-Mariano, D. *Current Opinion in Structural Biology* **2009**, 19, 658–665.
- [55] Glasner, M. E.; Gerlt, J. A.; Babbitt, P. C. *Current Opinion in Chemical Biology* **2006**, 10, 492–497.
- [56] Lu, Z.; Dunaway-Mariano, D.; Allen, K. N. *Biochemistry* **2005**, 44, 8684–8696.
- [57] Stewart, A. J.; Schmid, R.; Blindauer, C. A.; Paisey, S. J.; Farquharson, C. *Protein Engineering* **2003**, 16, 889–895.
- [58] Ridder, I. S.; Dijkstra, B. W. *The Biochemical journal* **1999**, 339, 223–226.
- [59] Wang, W.; Kim, R.; Jancarik, J.; Yokota, H.; Kim, S. *Structure* **2001**, 9, 65–71.
- [60] Rowlett, R. S. *Protein X-Ray Crystallography Methods*; Protocol, 2005.
- [61] Qiagen, *Critical factors for successful protein crystallization*; Design, 2010.
- [62] Novagen, *pET System Manual Novagen*; Biosystems, 2003; pp 1–68.
- [63] GE Healthcare, *Strategies for Protein Purification Handbook*; 2010; p 167.
- [64] Block, H.; Maertens, B.; Spriestersbach, A.; Brinker, N.; Kubicek, J.; Fabis, R.; Labahn, J.; Schäfer, F. *Methods in Enzymology* **2009**, 463, 439–473.
- [65] Koichi, I.; Michio, U. **1965**, 14, 361–366.
- [66] Pierson, E. Étude de l'inhibition de SerB2, une phosphosérine phosphatase essentielle de Mycobactérium Tuberculosis : identification et évaluation de nouveaux inhibiteurs potentiels, Mémoire, UNamur, Janvier 2018.
- [67] Guiner, A. *Théorie et technique de la radiocristallographie*, troisième ed.; Dunod, 1964.
- [68] Wlodawer, A.; Minor, W.; Dauter, Z.; Jaskolski, M. *FEBS J* **2008**, 275, 1–21.
- [69] Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W., et al. *Acta Crystallographica Section D : Biological Crystallography* **2010**, 66, 213–221.
- [70] Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. *Acta Crystallographica Section D : Biological Crystallography* **2010**, 66, 486–501.
- [71] Touw, W. G.; Baakman, C.; Black, J.; Ah the beek, T.; Krieger, E.; Joosten, R. P.; Vriend, G. *Nucleic Acids Research* **2015**, D, 364–368.
- [72] Harding, M. M. *Acta Crystallographica Section D : Biological Crystallography* **2001**, D57, 401–411.
- [73] Allen, F. H. *Acta Crystallographica Section B : Structural Science* **2002**, 58, 380–388.
- [74] Ménard, R.; Storer, A. C. *Biological Chemistry Hoppe-Seyler* **1992**, 373, 393–400.
- [75] Ordentlich, A.; Barak, D.; Kronman, C.; Ariel, N.; Segall, Y.; Velan, B.; Shafferman, A. *Journal of Biological Chemistry* **1998**, 273, 19509–19517.
- [76] Mandrich, L.; Menchise, V.; Alterio, V.; De Simone, G.; Pedone, C.; Rossi, M.; Manco, G. *Proteins : Structure, Function, and Bioinformatics* **2008**, 71, 1721–1731.
- [77] Menard, R.; Plouffe, C.; Laflamme, P.; Vernet, T.; Tessier, D. C.; Thomas, D. Y.; Storer, A. C.

- Biochemistry* **1995**, *34*, 464–471.
- [78] Silverman, R. B. *The organic chemistry of enzyme-catalyzed reactions*; Academic Press, 2002; p 29.
- [79] Heightman, T. D.; Vasella, A. T. *Angewandte Chemie International Edition* **1999**, *38*, 750–770.
- [80] Westheimer, F. H.; Schmidt Jr, D. E. *Biochemistry* **1971**, *10*, 1249–1253.
- [81] Structural medicine, Protein Crystallography Course. <http://www-structmed.cimr.cam.ac.uk>, site web en ligne. Consulté le 15 aout 2018.
- [82] Schrödinger, E. *Physical Review* **1926**, *28*, 1049.
- [83] Atkins, P. W.; De Paula, J. *Chimie physique*; De Boeck Supérieur, 2013.
- [84] Baer, S. H., M. and Lin; Alijah, A.; Adhikari, S.; Billing, G. D. *Physical Review A* **2000**, *62*, 032506.
- [85] Epstein, S. *The variation method in quantum chemistry*; Elsevier, 2012; Vol. 33.
- [86] Slater, J. C. *Physical review* **1951**, *81*, 385.
- [87] Dirac, P. A. M. *Mathematical Proceedings of the Cambridge Philosophical Society*; Cambridge University Press, 1930; Vol. 26.
- [88] Mulliken, R. S. *The Journal of Chemical Physics* **1955**, *23*, 1833–1840.
- [89] Szabo, A.; Ostlund, N. S. *Modern quantum chemistry : introduction to advanced electronic structure theory*; Courier Corporation, 2012.
- [90] Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. *The Journal of Physical Chemistry A* **2007**, *111*, 10439–10452.
- [91] *Expanding the limits of computational chemistry*; Density Functional (DFT) Methods, site web (en ligne), <http://gaussian.com/dft/>, consulté le 7 octobre 2017.
- [92] Parr, R. G. *Annual Review of Physical Chemistry* **1983**, *34*, 631–656.
- [93] Kohn, W.; Sham, L. J. *Physical Review* **1965**, *140*, A1133.
- [94] Hohenberg, P.; Kohn, W. *Physical Review* **1964**, *136*, B864.
- [95] Car, R. *Nature chemistry* **2016**, *8*, 820–821.
- [96] PSI4, DFT Functionals. <http://www.pscod.org>, site web en ligne. Consulté le 7 octobre 2017.
- [97] Fukui, K. *Accounts of chemical research* **1981**, *14*, 363–368.
- [98] Tomasi, J.; Mennucci, B.; Cancès, E. *Journal of Molecular Structure : Theochem* **1999**, *464*, 211–226.
- [99] Tomasi, J.; Mennucci, B.; Cammi, R. *Chemical reviews* **2005**, *105*, 2999–3094.
- [100] Cossi, M.; Barone, V.; Mennucci, B.; Tomasi, J. *Chemical Physics Letters* **1998**, *286*, 253–260.
- [101] Cramer, C. J.; Truhlar, D. G. *Chemical Reviews* **1999**, *99*, 2161–2200.
- [102] Gasteiger, E.; Gattiker, A.; Hoogland, C.; Ivanyi, I.; Appel, R.; Bairoch, A. *Nucleic Acids Resources* **2003**, *31*, 3784–3788.
- [103] The UniProt Consortium, *Nucleic Acids Ressources* **2017**, *45*, D158–D169.
- [104] Frisch, M. J. et al. Gaussian~16 Revision B.01. 2016; Gaussian Inc. Wallingford CT.
- [105] Dennington, R.; Keith, T. A.; Millam, J. M. GaussView Version 6. 2016; Semichem Inc. Shawnee Mission KS.
- [106] Liegeois, V. DrawMol, UNamur. 2016; www.unamur.be/drawmol.
- [107] Hanwell, M. D.; Curtis, D. E.; Lonié, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. *Journal of Cheminformatics* **2012**, *4*, 17.

Septième partie

Annexes



Full wwPDB X-ray Structure Validation Report ⓘ

Oct 23, 2018 – 02:16 PM BST

PDB ID : 6HYY
Title : Human phosphoserine phosphatase with serine and phosphate
Deposited on : 2018-10-22
Resolution : 1.57 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report.

This report is produced by the wwPDB biocuration pipeline after annotation of the structure.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.7.3 (157068), CSD as539be (2018)
Xtriage (Phenix) : 1.13
EDS : rb-20031633
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20031633

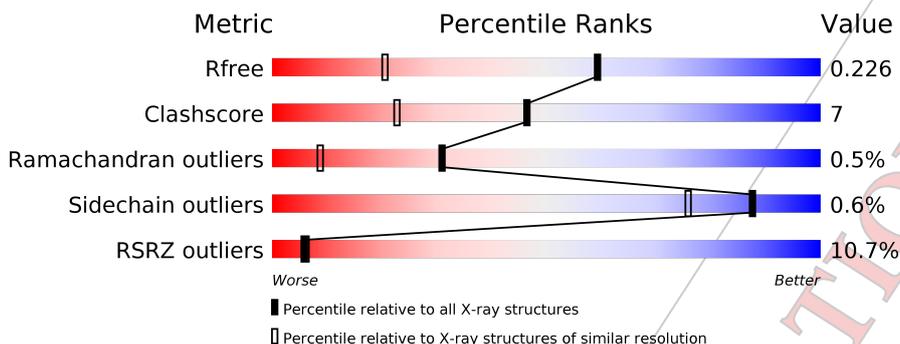
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.57 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



| Metric | Whole archive (#Entries) | Similar resolution (#Entries, resolution range(Å)) |
|-----------------------|-----------------------------|---|
| R_{free} | 111664 | 1224 (1.56-1.56) |
| Clashscore | 122126 | 1265 (1.56-1.56) |
| Ramachandran outliers | 120053 | 1240 (1.56-1.56) |
| Sidechain outliers | 120020 | 1238 (1.56-1.56) |
| RSRZ outliers | 108989 | 1207 (1.56-1.56) |

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

| Mol | Chain | Length | Quality of chain |
|-----|-------|--------|------------------|
| 1 | A | 221 | |
| 1 | B | 221 | |

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

| Mol | Type | Chain | Res | Chirality | Geometry | Clashes | Electron density |
|-----|------|-------|-----|-----------|----------|---------|------------------|
| 4 | SER | A | 303 | - | - | X | - |

CONFIDENTIAL VALIDATION REPORT



Full wwPDB X-ray Structure Validation Report i

Oct 23, 2018 – 10:41 AM BST

PDB ID : 6HYJ
Title : PSPH Human phosphoserine phosphatase
Deposited on : 2018-10-22
Resolution : 1.93 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report.

This report is produced by the wwPDB biocuration pipeline after annotation of the structure.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the i symbol.

The following versions of software and data (see [references](#) i) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.7.3 (157068), CSD as539be (2018)
Xtriage (Phenix) : 1.13
EDS : rb-20031633
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20031633

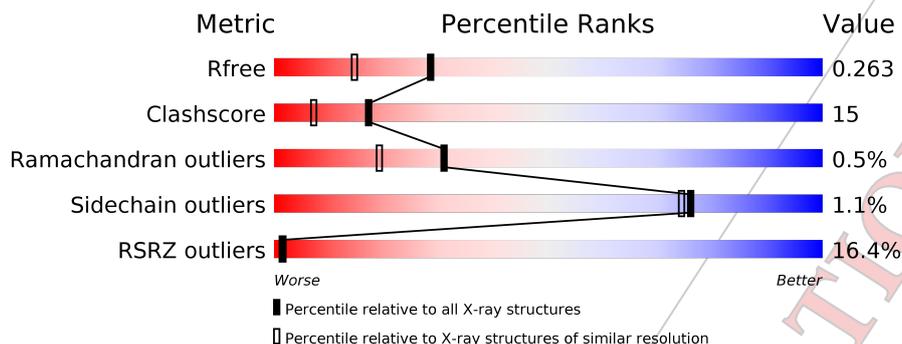
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.93 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



| Metric | Whole archive (#Entries) | Similar resolution (#Entries, resolution range(Å)) |
|-----------------------|-----------------------------|---|
| R_{free} | 111664 | 6904 (1.94-1.90) |
| Clashscore | 122126 | 7577 (1.94-1.90) |
| Ramachandran outliers | 120053 | 7491 (1.94-1.90) |
| Sidechain outliers | 120020 | 7491 (1.94-1.90) |
| RSRZ outliers | 108989 | 6759 (1.94-1.90) |

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

| Mol | Chain | Length | Quality of chain |
|-----|-------|--------|--|
| 1 | A | 224 | <div style="display: flex; align-items: center;"> <div style="width: 15%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 76%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 22%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 5%; height: 10px; background-color: grey; margin-right: 5px;"></div> <div style="width: 5%; height: 10px; background-color: grey;"></div> </div> |
| 2 | B | 223 | <div style="display: flex; align-items: center;"> <div style="width: 17%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 78%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 20%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 5%; height: 10px; background-color: grey; margin-right: 5px;"></div> <div style="width: 5%; height: 10px; background-color: grey;"></div> </div> |

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

| Mol | Type | Chain | Res | Chirality | Geometry | Clashes | Electron density |
|-----|------|-------|-----|-----------|----------|---------|------------------|
| 3 | SEP | A | 301 | - | - | X | X |
| 3 | SEP | B | 301 | - | - | - | X |
| 5 | SER | B | 303 | - | - | - | X |

CONFIDENTIAL VALIDATION REPORT



Full wwPDB X-ray Structure Validation Report i

Dec 11, 2018 – 11:50 AM GMT

PDB ID : 6Q6J
Title : Human phosphoserine phosphatase with substrate analogue homo-cysteic acid
Deposited on : 2018-12-11
Resolution : 1.99 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report.

This report is produced by the wwPDB biocuration pipeline after annotation of the structure.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the i symbol.

The following versions of software and data (see [references](#) i) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.7.3 (157068), CSD as539be (2018)
Xtriage (Phenix) : 1.13
EDS : rb-20031633
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20031633

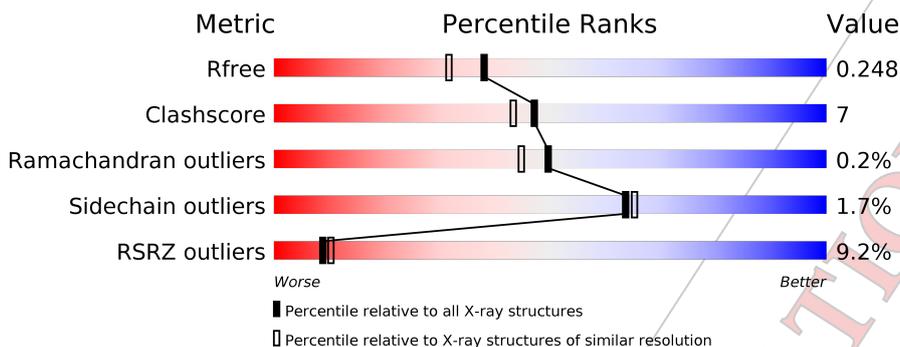
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.99 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



| Metric | Whole archive (#Entries) | Similar resolution (#Entries, resolution range(Å)) |
|-----------------------|-----------------------------|---|
| R_{free} | 111664 | 10189 (2.00-1.96) |
| Clashscore | 122126 | 11405 (2.00-1.96) |
| Ramachandran outliers | 120053 | 11281 (2.00-1.96) |
| Sidechain outliers | 120020 | 11280 (2.00-1.96) |
| RSRZ outliers | 108989 | 9953 (2.00-1.96) |

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

| Mol | Chain | Length | Quality of chain |
|-----|-------|--------|--------------------|
| 1 | A | 220 | 8% 86% 13% |
| 1 | B | 220 | 10% 83% 15% ** |

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

| Mol | Type | Chain | Res | Chirality | Geometry | Clashes | Electron density |
|-----|------|-------|-----|-----------|----------|---------|------------------|
| 3 | HJT | B | 302 | - | - | - | X |

CONFIDENTIAL VALIDATION REPORT