



**HAL**  
open science

**Caractérisation Structurale et Fonctionnelle de deux Enzymes de la Famille des Aldéhyde déshydrogénases : la Glycéraldéhyde-3-Phosphate Déshydrogénase de *B. stearothermophilus* et l'Erythrose-4-Phosphate Déshydrogénase d'*E. coli.* : structures cristallographiques d'intermédiaires réactionnels et de complexes enzyme-substrat**

Sébastien Moniot

► **To cite this version:**

Sébastien Moniot. Caractérisation Structurale et Fonctionnelle de deux Enzymes de la Famille des Aldéhyde déshydrogénases : la Glycéraldéhyde-3-Phosphate Déshydrogénase de *B. stearothermophilus* et l'Erythrose-4-Phosphate Déshydrogénase d'*E. coli.* : structures cristallographiques d'intermédiaires réactionnels et de complexes enzyme-substrat. Biochimie, Biologie Moléculaire. Université Henri Poincaré - Nancy 1, 2008. Français. NNT : 2008NAN10091 . tel-01748421

**HAL Id: tel-01748421**

**<https://hal.univ-lorraine.fr/tel-01748421>**

Submitted on 29 Mar 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



## AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : [ddoc-theses-contact@univ-lorraine.fr](mailto:ddoc-theses-contact@univ-lorraine.fr)

## LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

[http://www.cfcopies.com/V2/leg/leg\\_droi.php](http://www.cfcopies.com/V2/leg/leg_droi.php)

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>

## Thèse

présentée pour l'obtention du titre de  
**Docteur de l'Université Henri Poincaré, Nancy 1**  
en Enzymologie Moléculaire et Biologie Structurale  
par **Sébastien MONIOT**

**Caractérisation Structurale et Fonctionnelle de deux Enzymes de la Famille des Aldéhyde déshydrogénases : la Glycéraldéhyde-3-Phosphate Déshydrogénase de *B. stearothermophilus* et l'Erythrose-4-Phosphate Déshydrogénase d'*E. coli*.**  
*Structures cristallographiques d'intermédiaires réactionnels et de complexes enzyme-substrat.*

Soutenue le 20 octobre 2008

### Membres du Jury :

Monsieur Dominique BOURGEOIS	Directeur de Recherche, CNRS, UJF, Grenoble, rapporteur.
Monsieur Guy BRANLANT	Professeur des Universités, UHP, Nancy 1, examinateur.
Monsieur Gérard BRICOGNE	Directeur de Recherche, Global Phasing Limited, Cambridge, Royaume-Uni, examinateur.
Madame Catherine CORBIER	Professeur des Universités, UHP, Nancy 1, directeur de thèse.
Monsieur Claude DIDIERJEAN	Maître de Conférences - HDR, UHP, Nancy 1, co-directeur de thèse.
Madame Michèle REBOUD-RAVAUX	Professeur des Universités, UPMC, Paris VI, rapporteur.

### Membre Invité :

Monsieur Claude LECOMTE	Professeur des Universités, UHP, Nancy 1.
-------------------------	---

LABORATOIRE DE CRISTALLOGRAPHIE ET DE MODÉLISATION  
DES MATÉRIAUX MINÉRAUX ET BIOLOGIQUES

Groupe Biocristallographie, UMR CNRS-UHP 7036, Faculté des  
Sciences et Techniques, 54500 Vandoeuvre-lès-Nancy.



Ce travail a été réalisé au Laboratoire de Cristallographie et de Modélisation des Matériaux Minéraux et Biologiques (UMR CNRS-UHP 7036), dirigé par Monsieur le Professeur Claude Lecomte. Je tiens à lui exprimer mes profonds remerciements pour m'avoir accueilli dans son laboratoire où j'ai trouvé une très agréable ambiance de travail et dans lequel j'ai pu acquérir une formation scientifique en cristallographie des rayons X.

J'exprime ma respectueuse reconnaissance à Madame le Professeur Michèle Reboud-Ravaux et Monsieur le Docteur Dominique Bourgeois pour l'honneur qu'ils me font d'avoir accepté d'évaluer ce travail en tant que rapporteurs et à Messieurs le Docteur Gérard Bricogne et le Professeur Guy Branlant pour avoir accepté de juger cette thèse.

Je tiens à exprimer toute ma gratitude à Madame le Professeur Catherine Corbier qui a dirigé ce travail tout en me laissant une liberté de mouvement que j'ai beaucoup appréciée. Je la remercie pour l'encadrement de qualité dont elle m'a fait profiter, pour sa disponibilité, sa patience et sa grande générosité. Je lui suis particulièrement reconnaissant de m'avoir toujours accordé sa confiance et de m'avoir toujours soutenu au cours de ces années, et en particulier dans les moments plus difficiles. Je tiens également à remercier Monsieur le Docteur Claude Didierjean qui dirige l'équipe de Biocristallographie du LCM3B et qui a co-encadré ce travail. Son expérience dans l'obtention et l'analyse des structures cristallographiques de protéines a été très bénéfique sur la qualité de ces travaux. Je lui suis également reconnaissant de m'avoir accordé son amitié

Je tiens également à remercier les différentes personnes qui, par leur collaboration, ont contribué à ce travail de thèse. Les Professeurs Guy Branlant et Sandrine Boschi-Muller du MAEM pour les discussions scientifiques pertinentes lors de l'analyse de mes résultats et Alexandre Kriznick, pour m'avoir toujours fourni des échantillons protéiques de qualité ; Messieurs le Professeur Andrea Mozzarelli et le Docteur Stefano Bruno de l'Université de Parme, pour leur hospitalité et leur disponibilité lors de mes deux voyages en Italie et pour les discussions scientifiques fructueuses concernant la microspectrophotométrie et la GAPDH ; Monsieur le Docteur Clemens Vornrhein de Global Phasing pour son implication dans le délicat affinement des structures de GAPDH avec le logiciel Buster-TNT et pour ses remarques pertinentes concernant leurs interprétations.

J'exprime également ma profonde reconnaissance envers les Docteurs Frédérique Tete-Favier et Guillermo Mulliert pour m'avoir transmis leurs connaissances en cristallographie et en informatique, pour leur conseils éclairés tout au long de ce travail de thèse. Leur gentillesse et leur

disponibilité m'ont été très précieuses.

J'adresse mes profonds remerciements à Monsieur le Directeur de Recherche André Aubry, qui était directeur de l'équipe Biocristallographie au moment de mon arrivée au laboratoire, pour son précieux soutien et sa sympathie, ainsi qu'au Professeur Omar B. Shawkataly avec qui j'ai partagé de très agréables moments lors de son séjour en France.

Je tiens également à remercier Moutsé, Cha San, Hélène, William, Mickael, Bertrand, Brice Helena, Dorothée et El-Eulmi qui ont été des compagnons de thèse très agréables et précieux. J'ai partagé avec eux de formidables expériences et n'oublierai jamais leur soutien et leur amitié.

Je remercie Monsieur le Professeur Eric Chabrière, aujourd'hui à l'AFMB de Marseille, avec lequel j'ai travaillé sur le sujet de la protéine DING qui m'a permis d'acquérir une expérience en cristallographie à haute résolution.

Je remercie aussi toute l'équipe du LCM3B, chercheurs, techniciens et personnels administratifs qui ont contribué à faire de ces cinq années passées au laboratoire une expérience si bénéfique et assurément inoubliable. Je pense tout particulièrement à Sandrine Mathiot et Alexandre Bouché pour leur aide et leur disponibilité et au Docteur Benoît Guillot pour son aide dans la découverte de la haute résolution.

Je remercie également mes collègues enseignants et techniciens de l'IUT de Nancy-Brabois, au sein duquel j'ai pu acquérir une expérience de l'enseignement. Leur disponibilité et leurs conseils m'ont été très précieux et ont largement contribué aux très agréables souvenirs que me laisse cette expérience.

Je remercie les membres des lignes de lumière de l'ESRF (ID14, BM30A, BM14, ID23 et ID29) et de DESY (X11, X12, X13) sur lesquelles j'ai pu réaliser les différentes collectes de données de diffraction au cours de ma thèse.

J'exprime enfin ma plus profonde reconnaissance à ma famille : mes frères Thierry et Hervé et ma soeur Nadège pour avoir partagé ma vie durant ce travail et qui ont dû plus d'une fois supporter mon stress et mon emploi du temps ; mes parents qui m'ont permis de faire ces études, pour leurs perpétuels encouragements et leur confiance sans lesquels je n'aurais pu réussir.

Pour finir, un grand merci à Camille pour cette relecture estivale mais néanmoins attentive !

# Liste des Abréviations

( <i>Bst</i> )GAPDH	Glycéraldéhyde-3-phosphate déshydrogénase (de <i>Bacillus stearothermophilus</i> )
( <i>Eco</i> )E4PDH	Erythrose-4-phosphate déshydrogénase (d' <i>Escherichia coli</i> )
1,3-DPG (1,3DPG)	1,3-diphosphoglycérate
3-CAPAD	3-(chloroacétyl)pyridine adénine dinucléotide
3PG	3-phosphoglycérate
4PE	4-phosphoérythronate
Å	Angstroem
ALDH	Aldéhyde Déshydrogénase
apo (forme)	Forme de la protéine n'ayant pas fixé sa molécule de cofacteur
B (Facteur)	Facteur d'agitation thermique (en Å <sup>2</sup> )
C $\alpha$	Carbone $\alpha$
Da	Dalton
DESY	Deutschen Elektronen-Synchrotron
DTT	Dithiothréitol
E4P	Erythrose-4-phosphate
EDTA	Acide éthylène diamine-tétraacétique
EMBL	European Molecular Biology Laboratory
ESRF	European Synchrotron Radiation Facility
F <sub>c</sub> , F <sub>calc</sub>	Facteur de structure calculé
FIP	French beamline for Investigations of Proteins
F <sub>o</sub> , F <sub>obs</sub>	Facteur de structure observé
G3P	Glycéraldéhyde-3-phosphate
holo	Forme de la protéine ayant fixé sa molécule de cofacteur
IAA	Acide iodoacétique
k <sub>cat</sub>	Constante de vitesse
K <sub>M</sub>	Constante de Michaëlis
M	Molaire (mol.L <sup>-1</sup> )
NAD(P)	Nicotinamide adénine dinucléotide (phosphate)
PDB	Protein Data Bank (banque de données des structures de protéines de Brookhaven)

PEG	Polyéthylène glycol
PLP	Pyridoxal-5'-phosphate
rmsd	"root mean square deviation" écart quadratique moyen
SA	Sulfate d'ammonium
TF	Transformée de Fourier
Tris	Tris(hydroxyméthyl)aminométhane hydrochloride



# Table des matières

## Chapitre I : Étude Bibliographique.....9

### Partie A : La Glycéraldéhyde-3-Phosphate Déshydrogénase (GAPDH)....2

<b>A.1. Les différentes classes de GAPDH.....</b>	<b>2</b>
<b>A.2. Rôles physiologiques de la GAPDH.....</b>	<b>3</b>
<b>A.3. Structures des GAPDH phosphorylantes.....</b>	<b>4</b>
A.3.1. Structure primaire.....	4
A.3.2. Structure tridimensionnelle.....	4
A.3.2.1. Structure globale.....	4
A.3.2.2. Structure d'un monomère.....	5
A.3.2.2.1. Le domaine de fixation du cofacteur.....	5
Le sous-site de fixation de l'adénosine.....	6
Le pyrophosphate du NAD.....	6
Le sous-site de fixation du nicotinamide.....	6
Transition apo-holo et conséquence sur la réactivité de l'enzyme.....	7
A.3.2.2.2. Le domaine catalytique.....	8
<b>A.4. Le mécanisme catalytique de la GAPDH.....</b>	<b>9</b>
A.4.1. L'acylation.....	9
A.4.2. La désacylation.....	10
<b>A.5. Apport des études enzymologiques et structurales à la compréhension du mécanisme réactionnel.....</b>	<b>11</b>
A.5.1. Rôles des résidus du site actif.....	11
A.5.1.1. Les résidus essentiels : cystéine et histidine.....	11
A.5.1.2. Les sites de reconnaissances anionique.....	12
A.5.1.2.1. Le site Ps.....	13
A.5.1.2.2. Le site Pi.....	14
A.5.1.2.3. Apport de la mutagenèse-dirigée à l'étude des sites de reconnaissance anionique.....	15
A.5.2. Structures cristallographiques de complexes enzyme-substrat.....	16
A.5.2.1. Structure de l'intermédiaire hémithioacétal de la GAPDH d'E. coli sous forme apoenzyme.....	17
A.5.2.2. Structure d'un complexe ternaire de la GAPDH de T. cruzi avec le NAD et un inhibiteur covalent de la cystéine catalytique.....	18
A.5.2.3. Structure du complexe ternaire de type Michaelien de la BstGAPDH (GAPDH - NAD - G3P).....	19

***Partie B : L'érythrose-4-Phosphate déshydrogénase (E4PDH).....21***

**B.1. Introduction.....21**

**B.2. Réaction catalysée par l'E4PDH et mécanisme réactionnel.....22**

**B.3. Structures de l'E4PDH.....23**

B.3.1. Structure primaire.....23

B.3.2. Structure tridimensionnelle.....24

B.3.2.1. Structure quaternaire.....24

B.3.2.1.1. Structure des monomères.....25

Le domaine de fixation du cofacteur.....25

Le domaine catalytique.....26

B.3.2.1.2. Le site actif de l'enzyme.....26

Les résidus catalytiques.....26

Le(s) site(s) de reconnaissance anionique.....27

***Objectifs des travaux de thèse.....28***

***Chapitre II :Matériel et Méthodes.....30***

***Partie A : Principe de l'étude des protéines par cristallographie des rayons X.....31***

**A.1. Préambule.....31**

**A.2. Principes généraux de la cristallisation des macromolécules biologiques.....31**

A.2.1. Qu'est ce qu'un cristal ?.....31

A.2.2. Préparation de la protéine.....32

A.2.3. Cristallogenèse.....32

A.2.3.1. La diffusion de vapeur.....33

A.2.3.2. La méthode "batch".....33

A.2.3.3. Un processus multi-paramétrique.....34

**A.3. Collecte et traitement des données de diffraction.....35**

A.3.1. Acquisition des données.....35

A.3.1.1. Le rayonnement X.....35

A.3.1.1.1. Les Synchrotrons.....36

A.3.1.2. Préparation des cristaux : la cryocristallographie.....37

A.3.1.3. Protocole de collecte.....38

A.3.2. Le traitement des données.....39

A.3.2.1. Indexation : détermination des paramètres de maille.....39

A.3.2.2. Intégration des réflexions.....40

A.3.2.3. Mise à l'échelle et réduction du jeu de données.....40

A.3.2.4. Évaluation de la qualité du jeu de données.....41

**A.4. Résolution de structures par la cristallographie.....43**

A.4.1. Le problème de la phase.....43

A.4.1.1. Les différentes méthodes de phasage.....	43
Méthode du remplacement isomorphe.....	44
Diffusion anormale.....	44
Méthode de phasage direct.....	45
Le remplacement moléculaire.....	46
A.4.2. Affinement du modèle.....	47
A.4.3. Les cartes de densité électronique.....	48
A.4.4. Validation du modèle.....	49

## ***Partie B : Procédures expérimentales.....51***

<b>B.1. Généralités.....</b>	<b>51</b>
B.1.1. Les échantillons protéiques.....	51
B.1.2. Les essais de cristallisation.....	52
<b>B.2. Cristallisation de la GAPDH.....</b>	<b>53</b>
B.2.1.1. Trempage des cristaux.....	54
<b>B.3. Cristallisation de l'E4PDH.....</b>	<b>55</b>
<b>B.4. Collecte des données de diffraction.....</b>	<b>56</b>
<b>B.5. Phasage et affinement des structures cristallographiques.....</b>	<b>57</b>

## ***Chapitre III : Contribution à l'étude du mécanisme réactionnel de la Glycéraldéhyde-3-Phosphate Déshydrogénase.....59***

### ***Partie A : Résultats.....60***

<b>A.1. Contexte et principaux résultats.....</b>	<b>60</b>
<b>A.2. Structure de l'intermédiaire réactionnel thioacylenzyme de la GAPDH de type sauvage.....</b>	<b>62</b>
A.2.1. Stratégie.....	62
A.2.2. Cristallogenèse et essais préliminaires de trempage des cristaux d'holoenzyme en présence de substrat (G3P).....	63
A.2.3. Expériences de microspectrophotométrie - Optimisation des conditions de trempage .....	64
A.2.3.1. Préambule.....	64
A.2.3.2. Résultats des expériences de microspectrophotométrie.....	65
A.2.3.2.1. Expériences d'alkylation des cristaux de GAPDH.....	65
A.2.3.2.2. Incubation des cristaux de GAPDH holoenzyme en présence du substrat. ....	65
A.2.4. Enregistrement des données de diffraction.....	67
A.2.4.1. Cristaux PEG.....	67
A.2.4.2. Cristaux SA.....	68
A.2.5. Présentation de la structure et qualité du modèle.....	69
A.2.6. Étude du site actif de l'enzyme.....	71
A.2.6.1. Conformation de l'intermédiaire réactionnel GAPDH-G3P-NADH.....	71

A.2.6.2. La molécule de cofacteur.....	72
A.2.6.3. Les sites de reconnaissance anionique.....	72
A.2.6.4. Conformations alternatives de la boucle 206-212.....	73
A.2.7. Comparaison avec quelques structures représentatives des GAPDH.....	73
<b>A.3. Structure de complexes entre la GAPDH de <i>B. stearothermophilus</i>, le NAD<sup>+</sup> et le 1,3-DPG .....</b>	<b>75</b>
A.3.1. Stratégie.....	75
A.3.2. Structure des complexes ternaires C149S/A-1,3DPG-NAD <sup>+</sup> .....	75
A.3.2.1. Présentation des structures et qualité des modèles.....	75
A.3.2.1.1. Structure du mutant C149S.....	75
A.3.2.1.2. Structure du mutant C149A.....	76
A.3.2.2. Comparaison des structures ternaires et quaternaires des complexes GAPDH-1,3DPG-NAD <sup>+</sup> .....	77
A.3.2.2.1. Structure du mutant C149S.....	77
A.3.2.2.2. Structure du mutant C149A.....	77
A.3.2.3. Conformation du complexe ternaire C149S/A-1,3DPG-NAD <sup>+</sup> .....	78
A.3.2.3.1. La molécule de cofacteur.....	78
A.3.2.3.2. La molécule de 1,3-DPG.....	78
A.3.2.4. Conformation de la boucle 206-212.....	80
A.3.2.5. Comparaison avec les autres structures de GAPDH en complexe avec des substrats ou analogues de substrat.....	81
A.3.2.5.1. Comparaison avec la structure de la GAPDH de <i>T. cruzi</i> en complexe avec un analogue du 1,3-DPG.....	81
A.3.2.5.2. Comparaison avec les structures du complexe michaelien (C149S-G3P-NAD <sup>+</sup> ) et de l'intermédiaire thioacylenzyme de la GAPDH de <i>B. stearothermophilus</i> .....	82
<b>A.4. Structure d'un intermédiaire réactionnel acylenzyme-NAD<sup>+</sup> .....</b>	<b>83</b>
A.4.1. Objectif initial et stratégie.....	83
A.4.2. Présentation de la structure et qualité du modèle.....	83
A.4.3. Conformation du site actif.....	85
A.4.3.1. L'intermédiaire réactionnel acylenzyme.....	85
A.4.3.2. Conformation de la boucle 206-212.....	87
A.4.3.3. La molécule de cofacteur.....	87

***Partie B : Discussion.....88***

***Chapitre IV : étude structurale et fonctionnelle de l'Erythrose-4-phosphate déshydrogénase d'Escherichia coli.....96***

***Partie A : Résultats.....97***

A.1. Contexte et principaux résultats.....	97
A.2. Présentation des structures et qualité des modèles.....	99

A.2.1. Structure cristallographique de la forme apoenzyme de l'E4PDH d' <i>E. coli</i> (APO1).....	99
A.2.2. Nouvelle forme cristalline de l'E4PDH d' <i>E. coli</i> sous forme apoenzyme (APO2) .....	100
A.2.3. Structure cristallographique du complexe entre l'E4PDH d' <i>E. coli</i> et l'inhibiteur 3-(chloroacétyl) pyridine adénine dinucléotide (ECAP).....	101
<b>A.3. Analyse des structures tertiaire et quaternaire de l'E4PDH d'<i>E. coli</i>.....</b>	<b>101</b>
A.3.1. Analyse topologique de la structure.....	101
A.3.2. Comparaison de la structure des monomères d'E4PDH.....	102
A.3.3. Structure du tétramère de l'E4PDH d' <i>E. coli</i> .....	103
<b>A.4. Analyse des sites fonctionnels de l'E4PDH d'<i>E. coli</i>.....</b>	<b>104</b>
A.4.1. Conformation du site actif.....	104
A.4.2. Le site de fixation du cofacteur.....	105
A.4.2.1. Le complexe E4PDH - 3(chloroacétyl) pyridine adénine dinucléotide (ECAP) .....	105
A.4.2.2. Comparaison avec le site de fixation du cofacteur de la GAPDH de <i>B. stearothermophilus</i> .....	106
A.4.3. Les sites de reconnaissance anionique .....	108
A.4.3.1. Site anionique du modèle APO1.....	108
A.4.3.2. Site anionique du modèle ECAP.....	109
A.4.3.3. Recherche d'un site équivalent au site Ps des GAPDH.....	110
<b>A.5. Les boucles de structure variable.....</b>	<b>111</b>
A.5.1. Conformation de la boucle S (181-204).....	111
A.5.1.1. Préambule.....	111
A.5.1.2. Conformation de la boucle S (181-204) dans le modèle APO2.....	111
A.5.2. La boucle 96-102.....	112
A.5.3. La boucle 48-52.....	113
<b>A.6. Construction d'un modèle de l'intermédiaire hémithioacétal entre l'E4PDH et son substrat (E4P).....</b>	<b>113</b>

<b><i>Partie B : Discussion.....</i></b>	<b><i>115</i></b>
--	-------------------

<b><i>Conclusion et Perspectives.....</i></b>	<b><i>123</i></b>
---	-------------------

<b><i>Bibliographie.....</i></b>	<b><i>128</i></b>
----------------------------------	-------------------

<b><i>Annexe 1 : Publication : "Trapping of the Thioacyl-GAPDH intermediate"</i></b>
--

<b><i>Annexe 2 : Autres travaux réalisés pendant la thèse : la protéine DING</i></b>
--



# Liste des Figures et Tableaux

(Les figures ont été préparées avec le logiciel Pymol et annotées avec The GIMP)

## Chapitre I

### Figures

Figure I-1 : Réaction catalysée par la GAPDH.....	3bis
Figure I-2 : Structure tridimensionnelle d'une GAPDH phosphorylante.....	4bis
Figure I-3 : Représentation en ruban d'un monomère de la GAPDH.....	5bis
Figure I-4 : Représentation en ruban du domaine de fixation du cofacteur.....	5ter
Figure I-5 : Profil de liaisons hydrogène entre le NAD <sup>+</sup> et la <i>Bst</i> GAPDH.....	5ter
Figure I-6 : Représentation en ruban du domaine catalytique.....	8bis
Figure I-7 : Schéma du mécanisme de type ping-pong de la GAPDH.....	9bis
Figure I-8 : Détail du mécanisme catalytique.....	9ter
Figure I-9 : Vue stéréoscopique des sites de reconnaissance anionique Pi et Ps.....	13bis
Figure I-10 : Comparaison des deux conformations de la boucle $\beta$ 2- $\alpha$ 2 (207-210).....	15bis
Figure I-11 : Structure de l'intermédiaire hémithioacétal de la GAPDH d' <i>E. coli</i> .....	17bis
Figure I-12 : Formule développée de l'inhibiteur 3c.....	18bis
Figure I-13 : Structure d'un complexe ternaire de la GAPDH de <i>T. cruzi</i> .....	18bis
Figure I-14 : Structure du complexe ternaire de type michaelien de la GAPDH.....	20bis
Figure I-15 : Alignement de la séquence de l' <i>Eco</i> E4PDH avec une séquence consensus GAPDH.....	22bis
Figure I-16 : Alignement des séquences d'E4PDH de six microorganismes.....	24bis
Figure I-17 : Présentation du tétramère de l' <i>Eco</i> E4PDH.....	25bis
Figure I-18 : Superposition des monomères de l'E4PDH.....	25bis
Figure I-19 : Superposition des résidus catalytiques de la <i>Bst</i> GAPDH et de l' <i>Eco</i> E4PDH.....	27bis

### Tableaux

Tableau I-1 : Comparaison des sites de fixation du cofacteur de la <i>Bst</i> GAPDH et de l' <i>Eco</i> E4PDH.....	25ter
--	-------

## Chapitre II

### Figures

Figure II-1 : Représentation schématique d'un cristal de protéine.....	32bis
--	-------

<b>Figure II-2</b> : Représentation schématique de deux méthodes de cristallisation des protéines.....	33bis
<b>Figure II-3</b> : Diagramme de phase théorique d'un mélange protéine-agent précipitant.....	33bis
<b>Figure II-4</b> : Vue aérienne de l'ESRF (European Synchrotron Radiation Facility).....	36bis
<b>Figure II-5</b> : Les différents aimants utilisés pour la production de lumière sur synchrotron.....	36bis
<b>Figure II-6</b> : Représentation schématique d'une ligne de lumière sur synchrotron.....	37bis
<b>Figure II-7</b> : Influence de l'utilisation d'agent cryoprotecteur.....	38bis
<b>Figure II-8</b> : Illustration de la notion de résolution en cristallographie des rayons X.....	41bis
<b>Figure II-9</b> : Illustration de l'importance de l'information contenue dans les phases.....	43bis
<b>Figure II-10</b> : Schéma général de la méthode de phasage direct (Shake-and-Bake).....	45bis
<b>Figure II-11</b> : Illustration du concept de phasage par la méthode du remplacement moléculaire.....	47bis

## **Tableaux**

<b>Tableau II-1</b> : Cristallisation de la GAPDH de <i>B. stearothermophilus</i> .....	53bis
<b>Tableau II-2</b> : Cristallisation de l'E4PDH d' <i>E. coli</i> .....	55bis

## ***Chapitre III***

### **Figures**

<b>Figure III-1</b> : Représentation schématique du mécanisme réactionnel postulé des GAPDH phosphorylantes. .....	60bis
<b>Figure III-2</b> : Spectres d'absorption des cristaux de GAPDH obtenus en présence de PEG ou de sulfate d'ammonium en condition native et après alkylation par l'IAA.....	65bis
<b>Figure III-3</b> : Spectres d'absorption mesurés sur les cristaux de GAPDH holoenzyme obtenus en présence de PEG ou de sulfate d'ammonium et incubés en présence de G3P. ....	65bis
<b>Figure III-4</b> : Représentation graphique de la fraction de NADH formé en fonction de la concentration en <i>D,L</i> -G3P utilisé pour le trempage des cristaux.....	66bis
<b>Figure III-5</b> : Vue stéréoscopique des deux états, holoenzyme et thioacylenzyme, du site actif.....	70bis
<b>Figure III-6</b> : Vue stéréoscopique de l'intermédiaire réactionnel de la sous-unité R dans sa carte de densité électronique.....	71bis
<b>Figure III-7</b> : Vues stéréoscopiques de l'intermédiaire thioacylenzyme dans des cartes de densité électronique "omit".....	71ter
<b>Figure III-8</b> : Conformation de l'intermédiaire thioacylenzyme et interactions avec le site actif de la GAPDH.....	71quater
<b>Figure III-9</b> : Vue stéréoscopique de la molécule de cofacteur de la structure de l'intermédiaire thioacylenzyme de la <i>Bst</i> GAPDH.....	72bis
<b>Figure III-10</b> : Vue stéréoscopique des deux conformations alternatives de la boucle 206-212.....	73bis
<b>Figure III-11</b> : Superposition de la région 205-218 avec des structures représentatives.....	73bis



<b>Figure III-12</b> : Superposition des intermédiaires NADH - thioacylenzyme et apo-hémithioacétal de la GAPDH.....	73ter
<b>Figure III-13</b> : Superposition de l'intermédiaire thioacylenzyme - NADH et du complexe michaelien de la <i>Bst</i> GAPDH.....	74bis
<b>Figure III-14</b> : Contenu de l'unité asymétrique de la forme monoclinique <i>C</i> 2.....	76bis
<b>Figure III-15</b> : Vue stéréoscopique de la molécule de cofacteur NAD <sup>+</sup> dans sa carte de densité électronique.....	78bis
<b>Figure III-16</b> : Vue stéréoscopique de la molécule de 1,3-DPG dans des cartes de densité électronique "omit".....	78ter
<b>Figure III-17</b> : Vue stéréoscopique du réseau de liaisons hydrogène formé entre la protéine et la molécule de 1,3-DPG en conformation 1 (mutant C149S, monomère Q).....	79bis
<b>Figure III-18</b> : Vue stéréoscopique du réseau de liaisons hydrogène formé entre la protéine et la molécule de 1,3-DPG en conformation 2 (mutant C149S, monomère Q).....	79bis
<b>Figure III-19</b> : Vue stéréoscopique de la superposition de la région 206-215 des différentes sous-unités du complexe ternaire C149S-1,3DPG-NAD <sup>+</sup> .....	80bis
<b>Figure III-20</b> : Superposition du complexe ternaire C149S-1,3DPG-NAD <sup>+</sup> et de la structure de la GAPDH de <i>T. cruzi</i> en complexe avec un analogue du 1,3-diphosphoglycérate.....	81bis
<b>Figure III-21</b> : Superposition du complexe ternaire C149S-1,3DPG-NAD <sup>+</sup> avec la structure de l'intermédiaire thioacylenzyme.....	82bis
<b>Figure III-22</b> : Superposition des complexes ternaires C149S-1,3DPG-NAD <sup>+</sup> et C149S-G3P-NAD <sup>+</sup> .....	82bis
<b>Figure III-23</b> : Vue des 4 tétramères du modèle colorés en fonction de la valeur du facteur B.....	84ter
<b>Figure III-24</b> : Vue stéréoscopique de l'intermédiaire acylenzyme de la GAPDH.....	85bis
<b>Figure III-25</b> : Vue stéréoscopique de l'intermédiaire acylenzyme dans le site actif (confo $\alpha$ ).....	85ter
<b>Figure III-26</b> : Vue stéréoscopique de l'intermédiaire acylenzyme dans le site actif (confo $\beta$ ).....	86bis
<b>Figure III-27</b> : Vue du site actif dans lequel est lié l'intermédiaire acylenzyme (carte différence).....	86ter
<b>Figure III-28</b> : Superposition des molécules de cofacteur des différentes formes de la <i>Bst</i> GAPDH.....	87bis

## **Tableaux**

<b>Tableau III-1</b> : Statistiques des jeux de données de diffraction collectés sur des cristaux de <i>Bst</i> GAPDH obtenus en présence de PEG et incubés en présence de G3P.....	67bis
<b>Tableau III-2</b> : Statistiques des jeux de données de diffraction collectés sur des cristaux de <i>Bst</i> GAPDH obtenus en présence de sulfate d'ammonium et incubés en présence de G3P.....	68bis
<b>Tableau III-3</b> : Statistiques du jeu de données de diffraction et d'affinement de la structure de l'intermédiaire thioacylenzyme de la <i>Bst</i> GAPDH.....	68ter
<b>Tableau III-4</b> : Valeurs de rmsd résultant de la superposition des monomères de la structure de l'intermédiaire thioacylenzyme et comparaison avec la structure holoenzyme.....	69bis
<b>Tableau III-5</b> : Taux d'occupation et facteur B du modèle affiné (Buster-TNT).....	70bis
<b>Tableau III-6</b> : Distances interatomiques des liaisons hydrogène formées entre l'intermédiaire	

thioacylenzyme et l'enzyme.....	71quater
<b>Tableau III-7</b> : Statistiques des jeux de données de diffraction et d'affinement des structures de complexes ternaires entre la <i>Bst</i> GAPDH, le NAD <sup>+</sup> et le 1,3-DPG.....	75bis
<b>Tableau III-8</b> : Analyse des valeurs moyennes des facteurs B par monomère et par tétramère du complexe ternaire C149A-1,3DPG-NAD <sup>+</sup> .....	77bis
<b>Tableau III-9</b> : Valeurs de rmsd (Å) résultant de la superposition des carbones α des différentes sous-unités des modèles C149S- et C149A-1,3DPG-NAD <sup>+</sup> .....	77bis
<b>Tableau III-10</b> : Paramètres cristallographiques de la molécule de 1,3-DPG dans les deux complexes ternaires (C149S/A-1,3DPG-NAD <sup>+</sup> ).....	78bis
<b>Tableau III-11</b> : Interactions de type liaisons hydrogène entre la molécule de 1,3-DPG et les résidus du site actif du mutant C149S.....	79ter
<b>Tableau III-12</b> : Facteurs d'occupation et d'agitation thermique moyens du résidu Ser149 et des résidus de la boucle 206-212 de la structure du mutant C149S.....	79ter
<b>Tableau III-13</b> : Statistiques du jeu de données de diffraction et d'affinement de la structure de l'intermédiaire acylenzyme/NAD <sup>+</sup> de la <i>Bst</i> GAPDH.....	84bis
<b>Tableau III-14</b> : Valeur de rmsd résultant de la superposition des différents monomères sur le monomère O.....	84ter
<b>Tableau III-15</b> : Étude des facteurs B moyens de la structure par sous-unité et par tétramère.....	84ter
<b>Tableau III-16</b> : Facteurs d'occupation et facteurs B moyen des intermédiaires acylenzymes et des molécules de 1,3-DPG dans les différentes sous-unités du modèle. ....	85bis
<b>Tableau III-17</b> : Interactions de l'intermédiaire (conformation α) avec la structure protéique.....	85ter
<b>Tableau III-18</b> : Interactions de l'intermédiaire (conformation β) avec la structure protéique. ....	86bis
<b>Tableau III-19</b> : Paramètres cristallographiques de la boucle 206-212 de la structure de l'acylenzyme-NAD <sup>+</sup> .....	87bis

## *Chapitre IV*

### Figures

<b>Figure IV-1</b> : Vue de l'empilement cristallin de la forme quadratique de l' <i>Eco</i> E4PDH (APO2).....	100bis
<b>Figure IV-2</b> : Présentation de la structure de l'E4PDH d' <i>E. coli</i> .....	102bis
<b>Figure IV-3</b> : Profils de déviations moyennes de la position des atomes de Ca découlant de la superposition des tétramères APO1 et APO2 de l'E4PDH d' <i>E. coli</i> .....	103bis

<b>Figure IV-4</b> : Mise en évidence de la variabilité conformationnelle de la région 75-145 entre les formes APO1 et APO2 de l'E4PDH. ....	103bis
<b>Figure IV-5</b> : Comparaison de la conformation des résidus catalytiques dans le site actif des structures APO1 de l'E4PDH d' <i>E. coli</i> et apoGAPDH de <i>B. stearothermophilus</i> .....	104bis
<b>Figure IV-6</b> : Vue stéréoscopique de la molécule de 3-CAPAD en liaison covalente avec le résidu Cys149 .....	106bis
<b>Figure IV-7</b> : Superposition des molécules de 3-CAPAD.....	106bis
<b>Figure IV-8</b> : Vue stéréoscopique du site de fixation de la molécule de 3-CAPAD.....	106ter
<b>Figure IV-9</b> : Comparaison des sites de fixations du 3-CAPAD et du NAD.....	106quater
<b>Figure IV-10</b> : Comparaison de la conformation des boucles 233-237 de la <i>Bst</i> GAPDH et de l' <i>Eco</i> E4PDH .....	107bis
<b>Figure IV-11</b> : Vue du réseau de liaisons hydrogène du site de reconnaissance anionique de l'E4PDH apoenzyme (APO1 monomère B.....	109bis
<b>Figure IV-12</b> : Comparaison de la position des sites de reconnaissance anionique Pi et "new Pi" de la GAPDH avec le site identifié dans la structure APO1 de l'E4PDH d' <i>E. coli</i> .....	109bis
<b>Figure IV-13</b> : Vue stéréoscopique des boucles 181-190 des sous-unités B et C.....	112bis
<b>Figure IV-14</b> : Vue stéréoscopique de la superposition des boucles S (181-204) des quatre sous-unités du modèle APO2.....	112bis
<b>Figure IV-15</b> : Vue stéréoscopique de la superposition des boucles S de l' <i>Eco</i> E4PDH avec celles de la forme apoenzyme de la <i>Bst</i> GAPDH.....	112ter
<b>Figure IV-16</b> : Graphique représentant l'écart (en Å) entre la position des C $\alpha$ des boucles S de l'E4PDH par rapport à la boucle S de la <i>Bst</i> GAPDH sous forme apoenzyme.....	112ter
<b>Figure IV-17</b> : Comparaison des conformations de la boucle 96-102 dans les monomères B et C.....	113bis
<b>Figure IV-18</b> : Vue stéréoscopique du coeur du tétramère d'E4PDH.....	113bis
<b>Figure IV-19</b> : Modèle de l'intermédiaire hémithioacétal de l'E4PDH.....	114bis

## **Tableaux**

<b>Tableau IV-1</b> : Statistiques de collecte, d'affinement et qualité géométrique des structures APO1, APO2 et ECAP.....	98bis
<b>Tableau IV-2</b> : Valeurs de rmsd résultant des superpositions des atomes de C $\alpha$ des sous-unités composant chaque modèle.....	102ter
<b>Tableau IV-3</b> : Distances des liaisons hydrogène entre le 3-CAPAD et l'E4PDH.....	106ter
<b>Tableau IV-4</b> : Distances des liaisons hydrogène du site de reconnaissance anionique de l'E4PDH (APO1). .....	109bis



**CHAPITRE I : ÉTUDE**

**BIBLIOGRAPHIQUE**

# ***PARTIE A : LA GLYCÉRALDÉHYDE-3-PHOSPHATE DÉSHYDROGÉNASE (GAPDH)***

## ***A.1. Les différentes classes de GAPDH***

La *D*-glycéraldéhyde-3-phosphate déshydrogénase (GAPDH : EC 1.2.1.12) est une enzyme ubiquitaire dans le monde vivant. Elle appartient à la famille des aldéhyde déshydrogénases (ALDH) à cofacteur NAD(P). Trois classes de GAPDH existent et se distinguent par leur fonction enzymatique, leur localisation cellulaire et leur spécificité de cofacteur :

- **Les GAPDH glycolytiques** (EC 1.2.1.12), cytoplasmiques et NAD-dépendantes, sont impliquées dans la voie de la glycolyse et de la néoglucogenèse chez la plupart des bactéries et des eucaryotes (Harris and Waters, 1976). La GAPDH glycolytique catalyse, de manière réversible, la phosphorylation oxydative du *D*-glycéraldéhyde-3-phosphate (G3P) en 1,3-diphosphoglycérate (1,3-DPG) en réduisant une molécule de cofacteur et en utilisant une molécule de phosphate inorganique.
- **Les GAPDH chloroplastiques** (EC 1.2.1.13) sont, elles, retrouvées spécifiquement dans les chloroplastes des organismes photosynthétiques où elles interviennent dans le cycle de Calvin-Benson. Cette classe présente une dualité de spécificité de cofacteur mais possède

toutefois une préférence pour le NADP, cohérente avec son implication dans une voie anabolique (Ferri *et al.*, 1978). Les enzymes de cette classe catalysent la déphosphorylation réductive du 1,3-DPG en G3P en parallèle avec l'oxydation d'une molécule de cofacteur.

- La troisième classe regroupe les **GAPDH non phosphorylantes** cytosoliques (EC 1.2.1.9). Ces enzymes catalysent l'oxydation irréversible du G3P en 3-phosphoglycérate (3PG) en présence du cofacteur NADP et d'une molécule d'eau (Iglesias and Losada, 1988). Cette troisième classe de GAPDH se distingue très clairement des deux précédentes tant au niveau des séquences protéiques (environ 18 % d'identité de séquence) qu'au niveau des structures tridimensionnelles et est apparentée à certaines classes d'aldéhyde déshydrogénases (Michels *et al.*, 1994 ; Cobessi *et al.*, 2000).

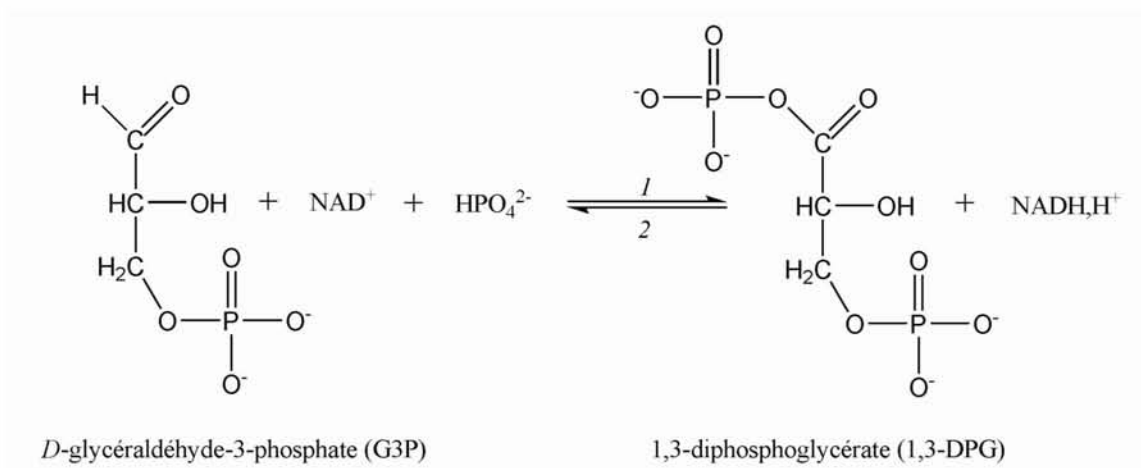
Les GAPDH phosphorylantes issues d'archaea ont une faible homologie de séquence avec les GAPDH isolées à partir de bactéries ou d'eucaryotes (16-20 %). Elles présentent également une dualité de spécificité de cofacteur avec, dans la plupart des cas, une préférence pour le NADP (Fothergill-Gilmore and Michels, 1993).

Au niveau structural, les GAPDH sont toutes des homotétramères à l'exception de certaines GAPDH chloroplastiques qui sont, elles, des hétérotétramères de type A<sub>2</sub>B<sub>2</sub>, la sous-unité B présentant, par rapport à la sous-unité A, une extension C-terminale qui possède une fonction régulatrice (Brinkmann *et al.*, 1989, Fermani *et al.*, 2007).

Par la suite, nous nous intéresserons essentiellement aux GAPDH phosphorylantes glycolytiques des bactéries et eucaryotes, et plus particulièrement à celle issue de *Bacillus stearothermophilus* (*BstGAPDH*) sur laquelle porte une partie de ces travaux de thèse.

## ***A.2. Rôles physiologiques de la GAPDH***

La D-glycéraldéhyde-3-phosphate déshydrogénase catalyse, en présence de phosphate inorganique, la phosphorylation oxydative réversible du D-G3P en 1,3-DPG, en utilisant le nicotinamide adénine nucléotide (NAD<sup>+</sup>) en tant qu'accepteur d'électrons. La GAPDH joue un rôle essentiel au sein des voies métaboliques de la glycolyse (sens 1) ou de la néoglucogenèse (sens 2) Figure I-1. Au niveau physiologique, la GAPDH a été très étudiée en tant qu'enzyme métabolique impliquée dans la glycolyse. Récemment, un grand nombre de publications ont suggéré que cette enzyme possédait de multiples fonctions chez les cellules de mammifères (pour une revue, se



**Figure I-1 : Réaction catalysée par la GAPDH.**



reporter à Sirover, 1999). Ces recherches ont mis en évidence, qu'en fonction de sa localisation cellulaire (cytosolique, membranaire ou nucléaire), la GAPDH serait associée aux mécanismes d'endocytose et de fusion membranaire, dans le transport par sécrétion vésiculaire, et dans le contrôle de la traduction, ou dans les mécanismes de transport nucléaire des ARN de transfert, de réplication et de réparation de l'ADN ou encore d'apoptose (pour une revue se reporter à Sirover, 2005). Ces fonctions encore putatives des GAPDH de mammifères n'étant pas en lien direct avec les travaux réalisés au cours de cette thèse qui concerne une enzyme d'origine bactérienne dédiée à la glycolyse, elle ne seront pas développées dans ce manuscrit.

## ***A.3. Structures des GAPDH phosphorylantes***

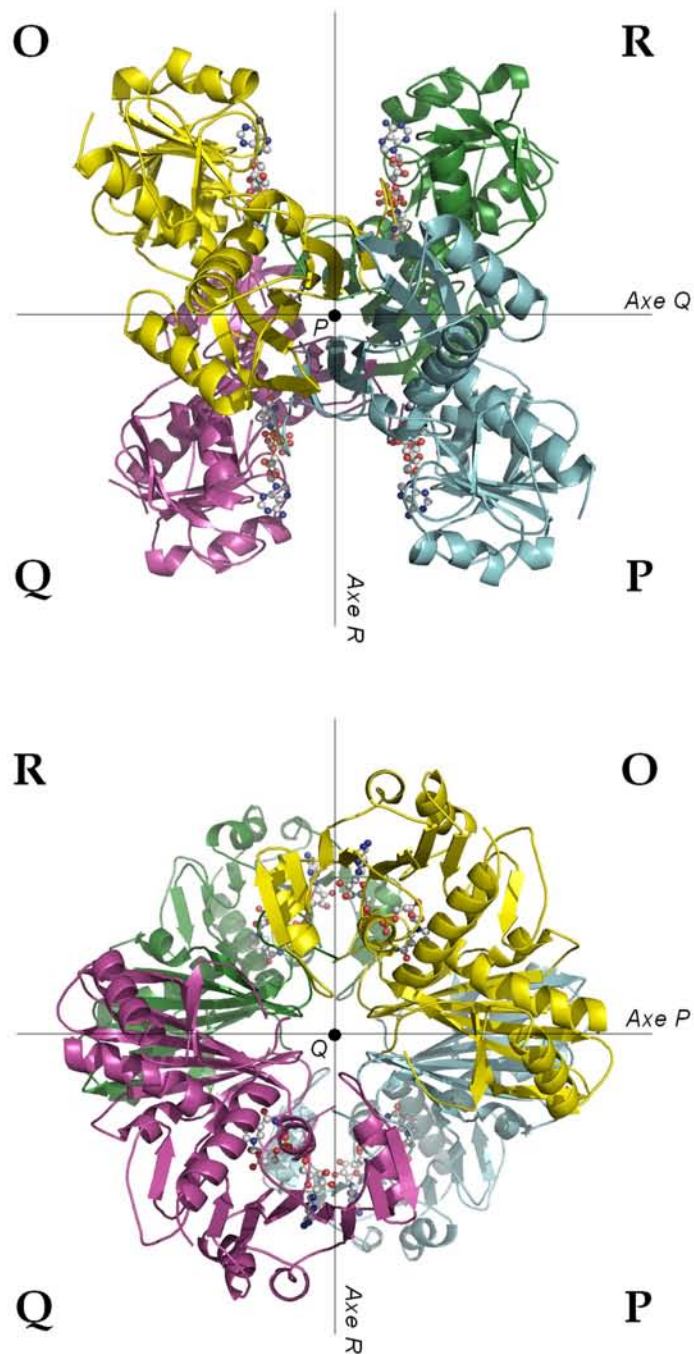
### ***A.3.1. Structure primaire***

L'alignement des séquences protéiques de GAPDH actuellement connues montre l'existence d'une forte identité entre les GAPDH glycolytiques et chloroplastiques issues des eubactéries et des eucaryotes (au minimum 40 %). Les séquences des GAPDH d'archaea divergent, elles, clairement et ne présentent qu'une faible identité de séquence avec les autres GAPDH phosphorylantes (environ 15 %).

### ***A.3.2. Structure tridimensionnelle***

#### ***A.3.2.1. Structure globale***

La GAPDH phosphorylante est une enzyme homotétramérique dont les sous-unités ont une masse moléculaire variant de 34 000 à 38 000 Da. Chacune des sous-unités fixe une molécule de cofacteur de manière coopérative. Cependant, cette coopérativité a été décrite comme positive ou négative en fonction de l'organisme dont est issue l'enzyme (Kirschner *et al.*, 1971 ; Corbier *et al.*, 1990a). Les quatre sous-unités du tétramère sont arrangées selon une symétrie 222 apparente et sont décrites dans un système de coordonnées orthogonales d'axes *P*, *Q* et *R*, initialement défini par Rossmann pour la lactate déshydrogénase (LDH) (Rossmann *et al.*, 1973) (Figure I-2). A la première sous-unité *O*, arbitrairement choisie, on fait correspondre, par ce système d'axes *P*, *Q* et *R*, les sous-unités *P*, *Q* et *R* respectives. La structure de la GAPDH a été résolue par cristallographie des rayons



**Figure I-2: Structure tridimensionnelle d'une GAPDH phosphorylante (EC 1.2.1.12).** Le modèle présenté correspond à l'enzyme issue de *B. stearothermophilus* sous forme holoenzyme (Igd1 : Skarzynski et al., 1987). Les axes de symétrie P, Q et R, internes à la molécule, permettent à partir d'une sous-unité O choisie arbitrairement de définir les sous-unités P, Q, et R, respectivement. La molécule de cofacteur est représentée en mode sphère et bâton et colorée en blanc. Les sous-unités O, P, Q et R sont colorées respectivement en jaune, bleu, violet et vert.

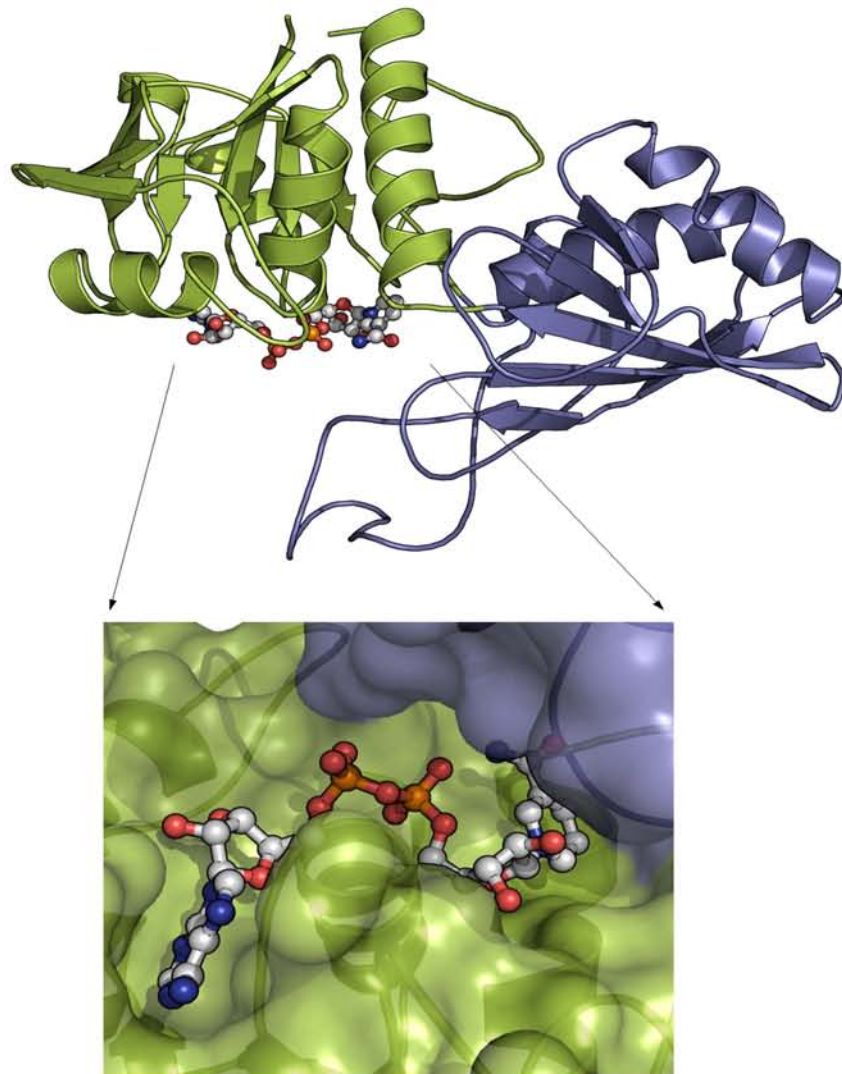
X pour une grande variété d'organismes vivants et sous différents états. Plus de 50 structures cristallographiques de GAPDH phosphorylantes glycolytiques (EC 1.2.1.12) ou chloroplastiques (EC 1.2.1.13) sont ainsi répertoriées dans la Protein DataBank. Les GAPDH de structures connues sont issues d'une vingtaine d'organismes sources appartenant aux trois règnes du vivant. La structure de certaines GAPDH a, par ailleurs, été déterminée sous ses formes apo et holoenzyme, c'est-à-dire en absence ou en présence du cofacteur. Les études structurales de Wonacott et collaborateurs comparant les deux formes de l'enzyme issue de *B. stearothermophilus* ont notamment permis d'étudier très précisément les bases moléculaires de la fixation du cofacteur et son implication sur la structure globale de l'enzyme (Skarzynski *et al.*, 1987 ; Skarzynski and Wonacott, 1988). Les structures tridimensionnelles de différents complexes, essentiellement avec des analogues de substrat, sont également disponibles dans la Protein DataBank et seront présentées plus loin dans ce manuscrit (*cf.* en A.5.2.).

### ***A.3.2.2. Structure d'un monomère***

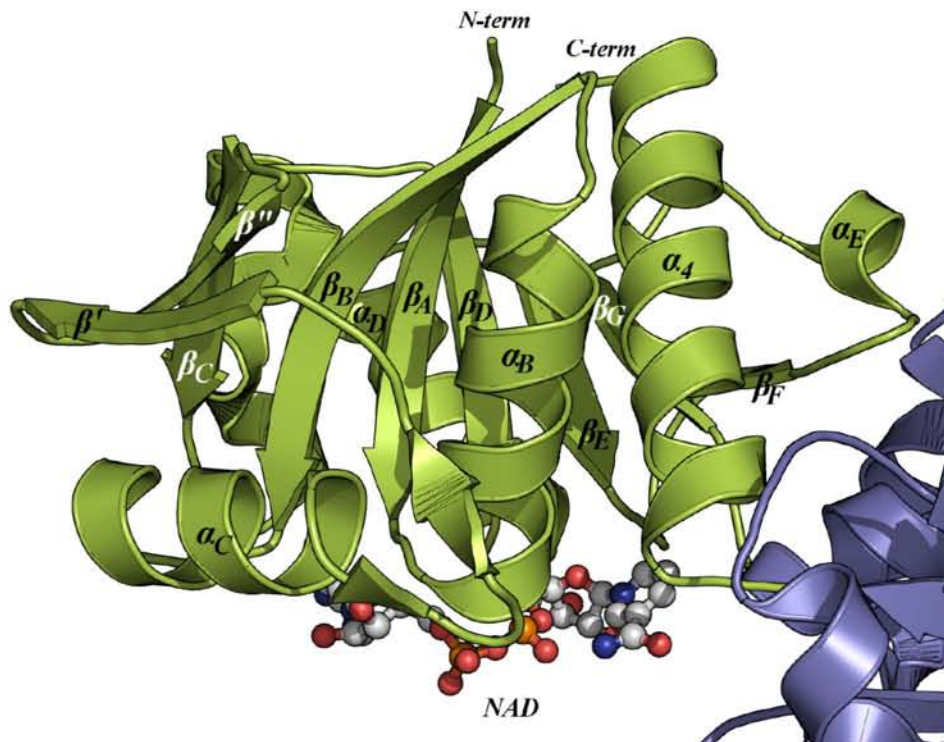
Une sous-unité est composée d'environ 330 acides aminés qui se replient en formant deux domaines ayant chacun une fonction spécifique (Figure I-3). Le premier, comprenant les résidus 1-148 et 312-333 (numérotation de Biesecker *et al.*, 1977 ; Branlant *et al.*, 1989), est impliqué dans la fixation de la molécule de cofacteur. Le second, appelé domaine catalytique, renferme les acides aminés 149-311 et inclut les résidus impliqués dans la fixation des substrats et dans la catalyse.

#### ***A.3.2.2.1. Le domaine de fixation du cofacteur***

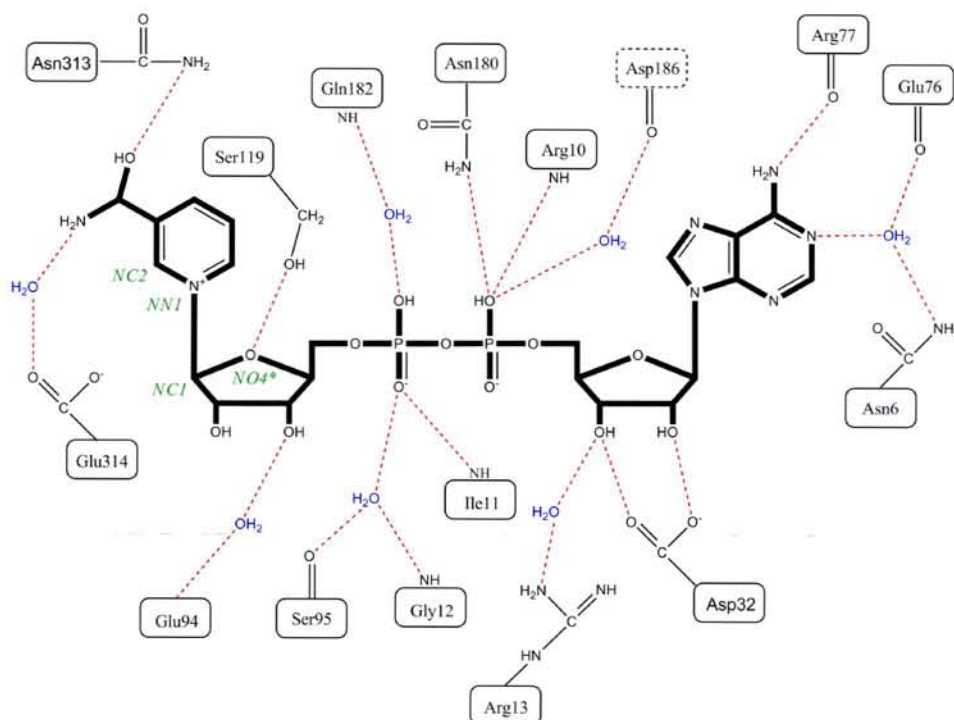
Le domaine de fixation du cofacteur est constitué des 148 acides aminés N-terminaux auxquels s'associent 22 résidus de l'extrémité C-terminal (312-333). La structure de ce domaine renferme 5 hélices  $\alpha$  (de  $\alpha_B$  à  $\alpha_E$  et  $\alpha_4$ ) et 9 brins  $\beta$  (de  $\beta_A$  à  $\beta_G$  et  $\beta'$ ,  $\beta''$ ) qui adoptent un repliement caractéristique de la fixation de cofacteur NAD(P) (Figure I-4). Les neuf brins  $\beta$  s'organisent selon un long feuillet  $\beta$  parallèle central (brins  $\beta_A$ ,  $\beta_B$ ,  $\beta'$ ,  $\beta''$ ,  $\beta_C$ ,  $\beta_D$ ,  $\beta_E$ ,  $\beta_F$  et  $\beta_G$ ) présentant une forme torsadée thermodynamiquement stable. Les deux noyaux du cofacteur, l'adénosine et le nicotinamide, sont reconnus par les deux motifs de Rossmann  $\beta_A \alpha_B \beta_B \alpha_C \beta_C$  et  $\beta_D \alpha_E \beta_E \alpha_4 \beta_F$ , respectivement (Rossmann *et al.*, 1974). La molécule de coenzyme est enfouie le long d'un sillon formé à la surface du domaine de fixation du cofacteur (Figure I-3). La Figure I-5 présente le réseau de liaisons hydrogène formé entre la protéine et le NAD dans la structure de *B. stearothermophilus* (Skarzynski *et al.*, 1987).



**Figure I-3 : Représentation en ruban d'un monomère de la GAPDH phosphorylante et vue de la poche de fixation du cofacteur.** (haut) La structure tridimensionnelle présentée correspond à un monomère de la GAPDH de *B. stearothermophilus* sous forme holoenzyme (1gd1, Skarzynski et al., 1987). Le domaine de fixation du cofacteur et le domaine catalytique sont colorés respectivement en vert et bleu. La molécule de cofacteur est représentée en mode sphère et bâton et colorée en blanc. (bas) La surface de l'enzyme est représentée avec la molécule de cofacteur. La surface en jaune appartient au domaine de fixation du cofacteur tandis que la surface en bleu appartient au domaine catalytique.



**Figure I-4 :** Représentation en ruban du domaine de fixation du cofacteur (en vert) avec la molécule de cofacteur (en blanc) et la nomenclature de la structure secondaire.



**Figure I-5 :** Profil de liaisons hydrogène entre le  $\text{NAD}^+$  et la GAPDH de *B. stearothermophilus*. La molécule de cofacteur est représentée en gras, les molécules d'eau en bleu. Le résidu dans le cadre en pointillé appartient à la sous-unité adjacente. (D'après Skarzynski et al., 1987).

## **Le sous-site de fixation de l'adénosine**

Le sous-site de fixation de l'adénosine est constitué d'une poche hydrophobe formée par les chaînes latérales des résidus Leu33 et Thr96. Deux liaisons hydrogène existent également entre la protéine et les atomes d'azote du cycle adénine. La première, au travers d'une molécule d'eau, relie l'azote AN1 aux résidus Asn6(N $\delta$ 2) et Glu76(O). La seconde est directe et relie l'azote AN6 à la fonction carbonyle de la chaîne principale du résidu Arg77. Les deux fonctions hydroxyle en position 2' et 3' du cycle ribose sont reconnues spécifiquement par le carboxylate de la chaîne latérale du résidu Asp32. Ce résidu, très conservé dans les séquences, participe à la spécificité pour le cofacteur NAD (Clermont *et al.*, 1993).

## **Le pyrophosphate du NAD**

Le pyrophosphate central, de part sa localisation dans l'alignement du N-terminal de l'hélice  $\alpha_B$ , est stabilisé par des interactions électrostatiques (Wierenga *et al.*, 1986). Une série de liaisons hydrogène directes avec la protéine ou par l'intermédiaire de molécules d'eau participe également à la fixation des deux groupements phosphates. Les interactions directes impliquent notamment les deux atomes d'azote peptidique des résidus Arg10 et Ile11 et la chaîne latérale de l'Asn180 (N $\delta$ 2).

## **Le sous-site de fixation du nicotinamide**

Le groupe nicotinamide est enfoui à l'intérieur de la structure à proximité du site actif. La fonction carboxamide portée par le cycle forme, au travers de son atome d'oxygène, une liaison hydrogène avec l'atome N $\delta$ 2 du résidu Asn313. Le noyau pyridinium est quant à lui stabilisé par des interactions hydrophobes avec les chaînes latérales des résidus Ile11 et Thr317. L'atome d'oxygène (NO4\*) du ribose adjacent au groupement nicotinamide forme une liaison hydrogène avec la chaîne latérale de la Ser119 et sa fonction hydroxyle en position 2' participent à la formation d'un site de reconnaissance anionique (Figure I-5). Cette partie du cofacteur NAD est essentielle pour le transfert d'hydrure ( $H^+ + 2e^-$ ) entre le substrat et l'atome C4 du nicotinamide et son positionnement correct est donc déterminant pour l'efficacité de la catalyse. Les GAPDH phosphorylantes fixent la molécule de cofacteur dans une conformation *syn* et appartiennent, de ce fait, à la classe B des déshydrogénases. Cette conformation *syn* du nicotinamide par rapport au ribose adjacent est définie par l'angle de torsion NO4\*-NC1-NN1-NC2 qui adopte une valeur de 0° à  $\pm 90^\circ$  (Eyschen *et al.*,

1999). Elle est stabilisée par une interaction intramoléculaire entre les atomes NN7 et OP1 du NAD et implique une stricte stéréospécificité du transfert d'hydrure qui met en jeu exclusivement le proton Pro-S (Eyschen *et al.*, 1999).

### **Transition apo-holo et conséquence sur la réactivité de l'enzyme**

La fixation de la molécule de cofacteur sur la GAPDH induit d'importants changements conformationnels mis en évidence grâce notamment à la comparaison des structures à haute résolution des formes apo- et holoenzyme de la GAPDH de *B. stearothermophilus* (Skarzynski and Wonacott, 1988). L'équipe de Wonacott propose un modèle séquentiel de la fixation du NAD, dont tous les éléments en partant du cycle adénine jusqu'au noyau nicotinamide, vont peu à peu développer des interactions favorables avec la protéine et conduire au réarrangement progressif de la structure du domaine de fixation du cofacteur.

La fixation de la partie ADP, dans le sous-site constitué des résidus Arg77 (O – AN6), Arg10 (N – AO3\*) et Ile11 (N – AO2\*) et de la poche hydrophobe formée par les résidus Leu33 et Thr96, s'effectue sans changement notable par rapport à la structure de l'apoenzyme. Néanmoins, cette fixation permettrait l'orientation et le positionnement corrects du reste de la molécule de NAD pour compléter la fixation. La liaison hydrogène intramoléculaire entre les atomes NN7 et NO1 du cofacteur est ensuite formée, positionnant les atomes d'oxygène NO4\*, du cycle ribose, et NO7, de l'amide du nicotinamide, à bonne distance pour former des liaisons hydrogène avec les chaînes latérales des résidus Ser119 et Asn313, respectivement. Les interactions hydrophobes avec les chaînes latérales de l'Ile11 et de la Tyr317 complètent le positionnement optimal du cofacteur à proximité des acides aminés catalytiques. Pour se fixer, le nicotinamide doit s'enfoncer profondément à l'intérieur de la structure de l'enzyme et ceci nécessite en retour des réarrangements substantiels du domaine de fixation du cofacteur. Globalement, la transition vers la forme holoenzyme entraîne une rotation de 4° du domaine de fixation du cofacteur par rapport au domaine catalytique et la structuration plus compacte du domaine de fixation du cofacteur autour du NAD.

De plus, le positionnement correct de l'ensemble de la molécule conduit à des modifications légères mais néanmoins essentielles du site actif. Il entraîne, en effet, le rapprochement des résidus catalytiques (Cys149 et His176) qui forment une paire d'ions exacerbant les propriétés nucléophiles de la cystéine. L'une des fonctions hydroxyle (en 2' du ribose du nicotinamide) participe également à la formation du site de reconnaissance anionique Ps. Les sites Pi et Ps de l'apoenzyme n'étant que partiellement occupés, contrairement à ceux de l'holoenzyme, par des ions sulfates, la présence du

cofacteur semble essentielle dans la constitution de sites de reconnaissance anioniques complètement fonctionnels.

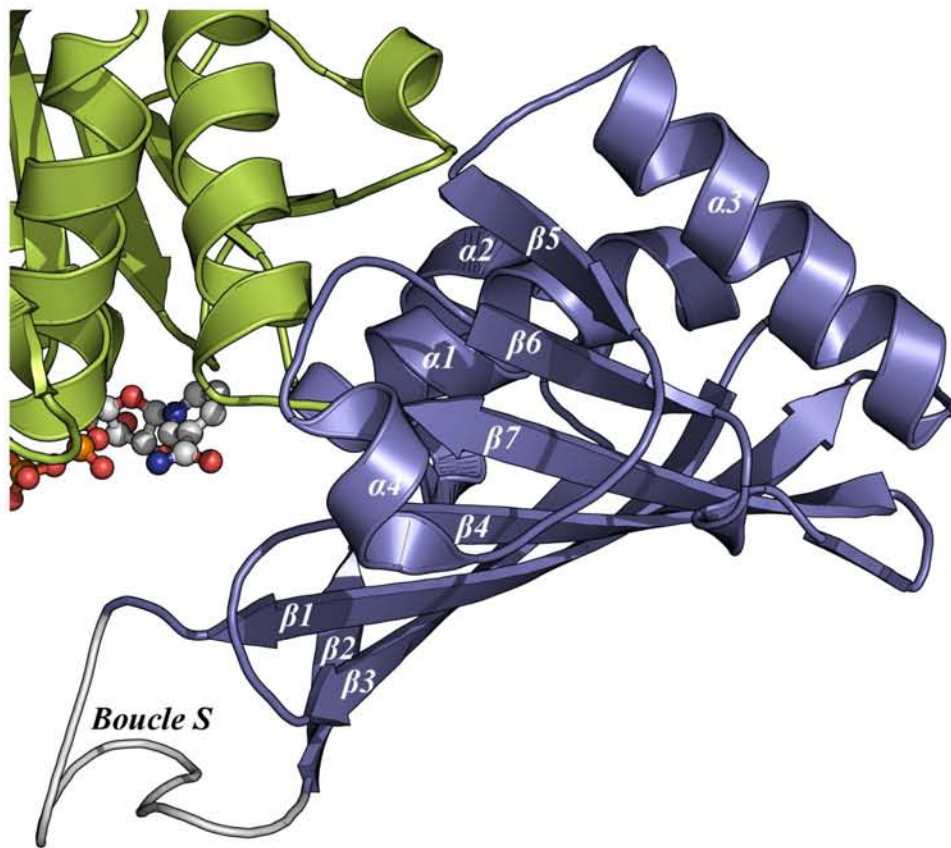
Une propriété spectroscopique particulière émerge de la proximité entre la charge positive globale du noyau pyridinium et la cystéine activée sous forme thiolate ( $-S^-$ ). Découverte par Racker et Krimsky (1952), cette bande d'absorption particulière (autour de 365 nm) apparaît lors de la fixation du NAD par l'apoenzyme et confère aux solutions protéiques de forte concentration une coloration jaune caractéristique. L'origine chimique exacte de ce phénomène est un transfert de charge qui s'établit entre la cystéine et le noyau pyridinium. Cette propriété spectrale a été utilisée de nombreuses fois dans l'étude de cette enzyme. En effet, l'apparition tout autant que la disparition de ce signal constituent des informations précises sur l'état de l'enzyme. La présence du signal indique que les deux partenaires sont idéalement positionnés pour l'acte catalytique, alors que l'absence de la bande de Racker indique soit que la cystéine est sous forme acylée (par le substrat), alkylée (par exemple par l'acide iodoacétique (IAA) ou oxydée, soit que la transition apo-holo est incomplète (suite à une mutation, par exemple) (Trentham, 1971a ; Harrigan and Trentham, 1973, Mougín *et al.*, 1988).

#### ***A.3.2.2.2. Le domaine catalytique***

Ce domaine est constitué des résidus 149 à 310 qui forment un feuillet  $\beta$  antiparallèle central de 7 brins (brins  $\beta_1$  à  $\beta_7$ ) flanqué par 4 hélices  $\alpha$  ( $\alpha_1$  à  $\alpha_4$ ) et une longue boucle très peu structurée appelée boucle S qui comprend les résidus 181 à 204 (Figure I-6). Cette boucle participe tout d'abord au maintien de la structure quaternaire de l'enzyme grâce aux interactions qu'elle crée avec les autres sous-unités composant le tétramère. Elle est également impliquée dans la spécificité de cofacteur au travers des résidus Leu187 et Pro188, qui, situés à proximité du ribose de la partie adénosine de la sous-unité adjacente, créent un encombrement stérique défavorable à l'accommodation du groupement phosphate du NAD(P) (Skarzynski *et al.*, 1987 ; Corbier *et al.*, 1990b ; Clermont *et al.*, 1993).

Le site actif se situe au fond d'une large poche formée entre les deux domaines et bordée par la boucle S du monomère relié par l'axe R. Le site renferme, entre autres, les deux résidus catalytiques Cys149 et His176, ainsi que les résidus impliqués dans la fixation des substrats et produits de la réaction (le *D*-G3P, le phosphate inorganique et le 1,3-DPG). Deux sites de reconnaissance anionique sont nécessaires à leur fixation et ont été identifiés dès la résolution de la structure cristallographique de la GAPDH de homard par l'équipe de Rossmann (Moras *et al.*,





*Figure I-6 : Représentation en ruban du domaine catalytique (en bleu) avec la nomenclature des structures secondaires. La boucle S (résidus 181 à 204), impliquée notamment dans le maintien de la structure quaternaire et dans la spécificité de cofacteur est représentée en blanc.*

1975). En effet, cette structure, dérivant de cristaux obtenus en présence de forte concentration de sulfate d'ammonium (agent précipitant), présentait deux ions sulfate proches du site actif qui ont permis de localiser ces deux sites. A partir de la modélisation de l'intermédiaire thiohémiacétal dans la structure de la GAPDH de homard, les deux sites de reconnaissance anionique ont été attribués respectivement à la fixation du phosphate inorganique et du groupement phosphate du substrat et dénommés Pi et Ps. Cependant, Moras et collaborateurs n'excluaient pas la possibilité que le groupement C3-phosphate appartenant au substrat puisse passer d'un site à l'autre au cours des différentes étapes de la catalyse.

## ***A.4. Le mécanisme catalytique de la GAPDH***

Le mécanisme réactionnel le plus couramment accepté pour la GAPDH a été établi au début des années 70, en particulier grâce aux travaux de Segal et Boyer (1953) puis de Trentham et collaborateurs (Trentham, 1968, 1971a, 1971b ; Harrigan and Trentham 1973, 1974). Les différentes études cinétiques à l'état stationnaire et pré-stationnaire concluent à un mécanisme de type ping-pong, présenté dans la Figure I-7, dans lequel l'holoenzyme et le thioacylenzyme, un intermédiaire réactionnel covalent entre l'enzyme et le substrat, sont les formes les plus stables. Le mécanisme se décompose en deux étapes dites d'acylation et de désacylation (phosphorolyse), ponctuées par la libération du cofacteur réduit et l'entrée d'une nouvelle molécule de  $\text{NAD}^+$ .

### ***A.4.1. L'acylation***

Comme nous l'avons vu en A.3.2.2.1, l'enzyme, lorsqu'elle a fixé le cofacteur, se trouve dans une conformation activée, la formation d'une paire d'ions entre la Cys149 et l'His176 stabilisant la forme thiolate de la cystéine (abaissement de son pKa) tout en exacerbant son caractère nucléophile (Soukri *et al.*, 1989). A l'approche du carbone aldéhydique C1 du G3P ( $sp^2$ ), l'attaque nucléophile de la cystéine essentielle conduit à la formation d'un intermédiaire tétraédrique ( $sp^3$ ), appelé thiohémiacétal. Cet intermédiaire, particulièrement instable, est stabilisé par une assistance acide du noyau imidazole de l'His176, qui cède son proton à l'oxygène porté par le C1 du G3P (Figure I-8a). L'étape suivante est le transfert d'hydrure ( $\text{H}^+ + 2e^-$ ) depuis le carbone C1 vers le carbone C4 du noyau pyridinium du NAD qui conduit à la réduction du cofacteur et à la formation de l'intermédiaire thioacylenzyme métastable. Ce transfert d'hydrure est assisté par l'His176, (déprotonée à ce stade de la réaction) qui arrache le proton de la fonction

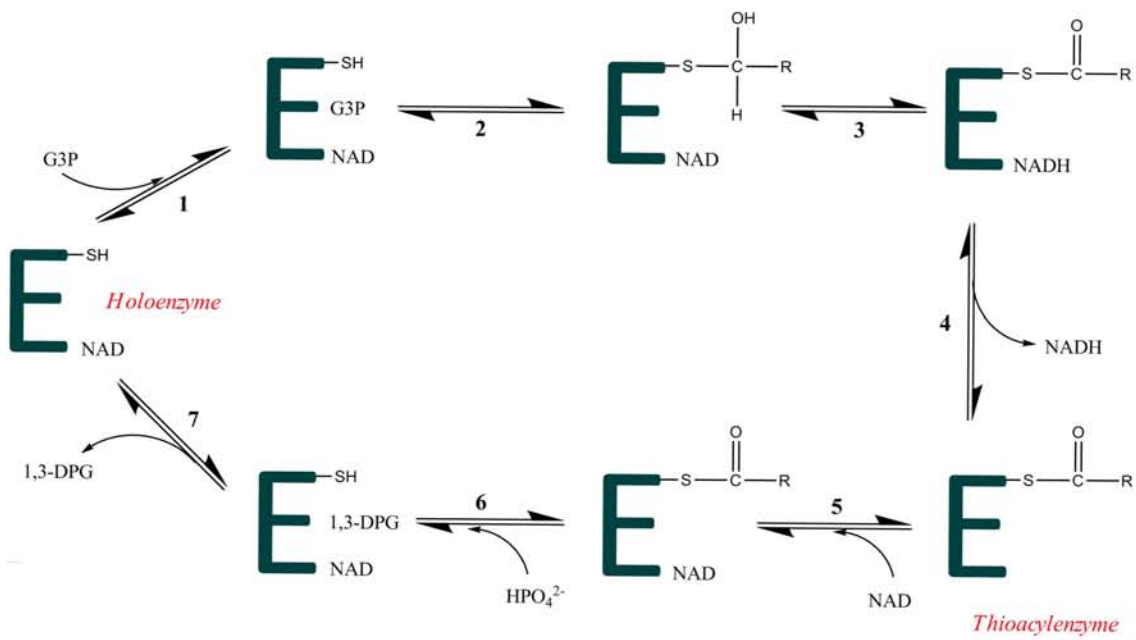
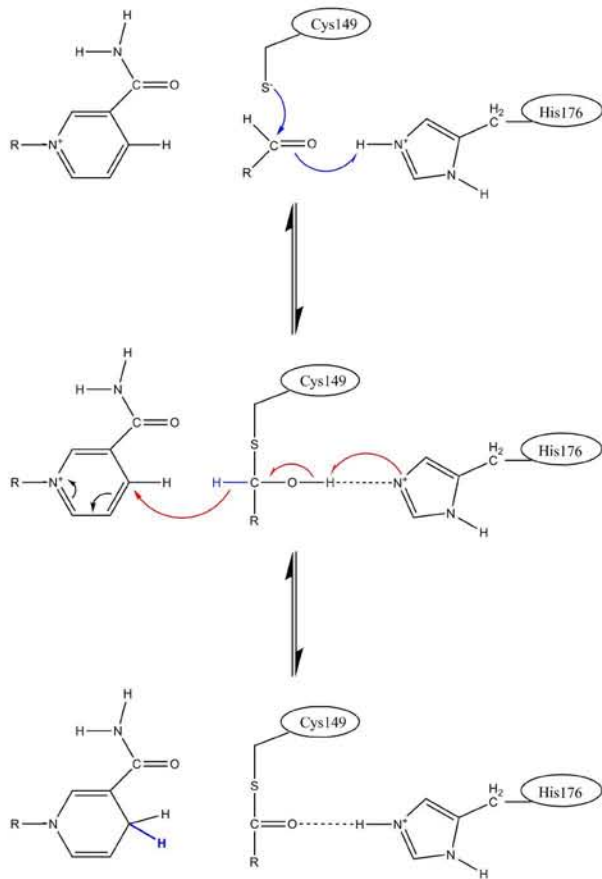
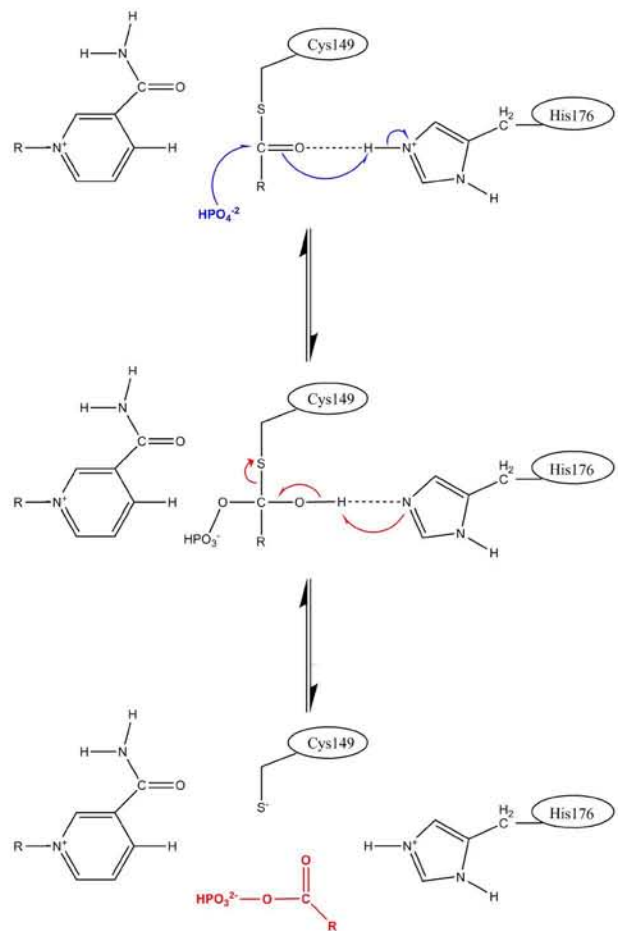


Figure I-7 : Schéma du mécanisme de type ping-pong de la GAPDH. (D'après Segal and Boyer, 1953)



**a. Acylation.** La réaction débute par l'attaque nucléophile de la cystéine catalytique sur le carbone aldéhydique du substrat, avec l'assistance acide de l'histidine catalytique (bleu). Le transfert d'hydrure est facilité par l'assistance basique de l'His176 (rouge). A l'issue de l'acylation, le cofacteur est réduit et l'intermédiaire thioacylenzyme est formé.

**b. Désacylation.** La seconde étape débute par l'entrée d'un phosphate inorganique qui réalise une attaque nucléophile sur le thioacylenzyme, assistée par l'His176 (bleu). Le produit de la réaction (1,3-DPG) est finalement libéré (rouge).



**Figure I-8 : Détail du mécanisme catalytique.** Les étapes d'acylation et de désacylation sont ponctuées par l'étape d'échange de cofacteur (libération de NADH et entrée de NAD<sup>+</sup>). D'après Boschi-Muller and Branlant, 1999.

hydroxyle en C1. L'existence d'un effet isotopique de substrat pour le mutant H176N de la GAPDH, non observé pour l'enzyme sauvage, a permis de démontrer l'assistance basique de l'His176 au cours de ce transfert d'hydrure (Talfournier *et al.*, 1998). A l'issue de l'étape d'acylation, l'enzyme est liée au substrat de manière covalente par une liaison de type thioester.

Le cofacteur sous forme réduite se dissocie alors du complexe ternaire pour être remplacé par une molécule de cofacteur oxydée. Des études cinétiques ont montré que la présence du  $\text{NAD}^+$  augmentait l'efficacité de l'étape de désacylation (Trentham, 1968 ; Biesecker *et al.*, 1977 ; Corbier *et al.*, 1989).

### ***A.4.2. La désacylation***

Dans les conditions physiologiques, la seconde étape de la réaction nécessite la fixation d'une molécule de phosphate inorganique ( $\text{HPO}_4^{2-}$ ) dans le site actif de l'enzyme (Armstrong and Trentham, 1976). L'attaque nucléophile de celle-ci sur le carbone plan de la liaison thioester pourrait à nouveau être facilitée par la présence de l'His176 sous forme protonée et jouant le rôle d'acide. Le mode de stabilisation de ce second intermédiaire thiohémiacétal ( $\text{sp}^3$ ) ainsi formé est supposé identique à celui de l'étape précédente, son groupement hydroxyle en position C1 étant stabilisé grâce à une liaison H avec l'His176 (Figure I-8b).

Comme indiqué précédemment, la vitesse de cette étape de phosphorolyse est fortement accélérée par la présence du  $\text{NAD}^+$ , probablement parce que la charge partielle positive portée par le carbone C4 du nicotinamide polarise la liaison C-S avec le substrat et favorise, avec l'assistance basique de l'His176, la formation de l'ester phosphorique (Trentham, 1968 ; Harrigan and Trentham, 1974 ; Biesecker *et al.*, 1977). A la fin de la réaction, le produit (1,3-DPG) est libéré et l'enzyme, ayant préalablement fixé une molécule de cofacteur oxydée, se retrouve directement sous forme active, prête pour un nouveau cycle catalytique.

Les études cinétiques réalisées sur des GAPDH issues de divers organismes sources (Trentham, 1971a et 1971b ; Corbier *et al.*, 1990b ; Liu and Huskey, 1992 ; Corbier *et al.*, 1994) suggèrent que les étapes limitantes de la réaction sont :

- dans le sens de l'oxydation phosphorylante : la phosphorolyse à pH acide et une étape associée au relargage du NADH à pH basique ;
- dans le sens de la déphosphorylation : le relargage du *D*-G3P à pH acide et un changement conformationnel lié à la fixation du NADH à pH basique.

## ***A.5. Apport des études enzymologiques et structurales à la compréhension du mécanisme réactionnel***

### ***A.5.1. Rôles des résidus du site actif***

Situé au fond d'une cavité formée entre les domaines de fixation du cofacteur et catalytique, le site actif de l'enzyme est composé des deux résidus catalytiques, la Cys149 et l'His176, du groupement nicotinamide du cofacteur ainsi que de deux sites de reconnaissance anionique fixant le groupement phosphate du substrat et le phosphate inorganique.

#### ***A.5.1.1. Les résidus essentiels : cystéine et histidine***

La conformation et le positionnement de la cystéine essentielle sont équivalents dans l'ensemble des structures tridimensionnelles connues de GAPDH, y compris pour les enzymes isolées d'archaea. La substitution de ce résidu par les acides aminés Ser, Ala ou Gly, diminue drastiquement, voire totalement, l'activité enzymatique de la protéine (Mougin *et al.*, 1988 ; Boschi-Muller and Branlant, 1999). En revanche, une certaine variabilité a été observée concernant l'histidine catalytique en fonction de l'organisme dont est issue l'enzyme. Dans la plupart des structures tridimensionnelles de GAPDH résolues à ce jour, la conformation de l'histidine est conservée. Celle-ci pointe, en effet, vers la cystéine essentielle et présente l'atome N $\epsilon_2$  de son cycle à l'atome S $\gamma$  de la cystéine à une distance oscillant autour 3,5Å. Le positionnement et l'orientation corrects du noyau imidazole sont assurés grâce à une liaison hydrogène entre l'atome N $\delta$  du cycle et la fonction carbonyle de la chaîne peptidique en position 177. La distance entre l'histidine et la cystéine essentielle varie avec l'état apo ou holoenzyme de la protéine. Si la fixation du cofacteur induit un rapprochement (de 0,4 Å) entre les deux résidus dans le cas de la GAPDH de *B. stearothermophilus* (Skarzynski and Wonacott, 1988), elle semble induire une très légère augmentation de cette distance (de l'ordre du 10<sup>e</sup> d'angström), dans le cas d'*Escherichia coli* (Yun *et al.*, 2000). Le cas de la structure cristallographique de la GAPDH de homard reste à ce jour un cas unique, dans lequel les sites actifs des différentes sous-unités ne sont pas dans la même conformation. Ainsi, alors que deux des quatre sites actifs répondent à la description faite ci-dessus et sont donc présentés comme étant dans une conformation active, les deux autres montrent une conformation différente dans laquelle l'atome N $\epsilon_2$  de l'histidine ne pointe pas vers la cystéine. Néanmoins, la faible résolution de la structure cristallographique (2,9Å) rend incertaine

l'interprétation de la densité et du modèle protéique qui en découle. Les GAPDH isolées d'archaea constituent un cas particulier. En effet, dans ce cas, l'histidine essentielle ne se situe pas, dans les structures primaire et tertiaire, dans une position équivalente à l'histidine essentielle des enzymes isolées à partir de bactéries ou d'eucaryotes. Les deux structures tridimensionnelles connues de GAPDH issues d'archaea montrent toutefois qu'une histidine située au niveau du brin  $\beta_4$  (contre  $\beta_1$  pour les GAPDH d'autres sources) en position 219 de la séquence de la GAPDH de *Sulfolobus solfataricus* adopte un positionnement comparable et jouerait un rôle équivalent à l'His176 de la GAPDH de *B. stearothermophilus* (Isupov *et al.*, 1999, Charron *et al.*, 2000).

Les études de mutagenèse dirigée qui ont eu pour objet la substitution de cette histidine par une asparagine ont conduit à des observations similaires quant aux propriétés des GAPDH mutantes (Soukri *et al.*, 1989 ; Talfournier *et al.*, 1998 et 1999 ; Charron *et al.*, 2000) :

- un accroissement du pKa de la cystéine catalytique variant d'une à plus de deux unités pH et donc une diminution de son caractère nucléophile et de sa réactivité (pKa de 5,92 pour la Cys149 de l'enzyme de type sauvage isolée de *B. stearothermophilus*, contre 8,00 pour le mutant H176N (Talfournier *et al.*, 1998))
- un changement de l'étape limitante de la réaction qui est désormais associée à la formation de l'acylenzyme (à pH basique et dans le sens de l'oxydation phosphorylante) et non plus à une étape ultérieure à l'acylation. Plus précisément, l'apparition d'effet isotopique avec un substrat deutéré pour les mutants H176N suggère l'implication directe de l'His176 dans l'étape de transfert d'hydrure (Talfournier *et al.*, 1998).

#### *A.5.1.2. Les sites de reconnaissances anionique*

La réaction catalysée par les GAPDH phosphorylantes fait intervenir alternativement le glycéraldéhyde-3-phosphate et le phosphate inorganique ou le 1,3-diphosphoglycérate. De ce fait, pour accueillir les deux groupements phosphate, deux sites de reconnaissance anionique sont nécessaires et ont d'ailleurs été mis en évidence dès la résolution de la première structure cristallographique de GAPDH phosphorylante (enzyme de homard) par l'équipe de Rossmann en 1975. Comme précédemment présenté en A.3.2.2.2, la modélisation de l'intermédiaire hémithioacétal dans la structure de la GAPDH de homard a conduit les auteurs à proposer la dénomination Pi et Ps pour chaque site, les indices i et s renvoyant respectivement au phosphate inorganique et au phosphate du substrat. Néanmoins, la topologie exacte de chacun de ces deux

sites ainsi que les acides aminés impliqués dans leur formation n'ont été précisés que plus tard, grâce à la structure tridimensionnelle de l'enzyme de *B. stearothermophilus* obtenue à haute résolution (1,8 Å) (Skarzynski *et al.*, 1987). Le site actif de cette structure (code pdb 1gd1) avec les deux anions sulfate indiquant la position des sites Pi et Ps est présenté en Figure I-9. Il faut noter, qu'à nouveau, les auteurs soulèvent des interrogations quant au véritable rôle de chacun des deux sites de reconnaissance anionique au cours de la catalyse. Comme nous le verrons ensuite (*cf.* en A.5.1.2.3.), de nombreux travaux ont été réalisés à ce sujet sans toutefois pouvoir attribuer de manière univoque leur fonction respective. Néanmoins, de manière à faciliter les comparaisons, les dénominations Pi et Ps ont été conservées.

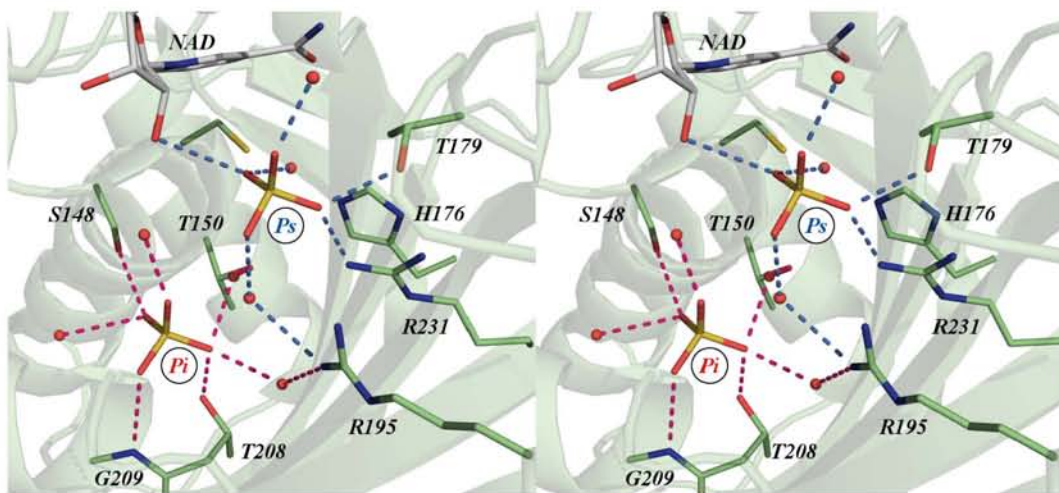
#### ***A.5.1.2.1. Le site Ps***

Le site Ps, dans la structure de la GAPDH de *B. stearothermophilus*, est occupé par un anion sulfate qui forme des liaisons hydrogène avec les chaînes latérales des résidus conservés Thr179 et Arg231. La fonction hydroxyle portée par le ribose adjacent au nicotinamide (NAD) en position 2' forme, elle aussi, une liaison hydrogène avec la molécule de sulfate. La contribution de la partie ribose du cofacteur semble essentielle à la constitution d'un site de reconnaissance anionique Ps complètement fonctionnel. En effet, la densité électronique observée au niveau de ce site dans la structure de l'apoenzyme est plus faible, signe d'une occupation partielle de l'anion sulfate en l'absence du cofacteur (Skarzynski *et al.*, 1987).

La localisation exacte du site Ps ainsi que sa géométrie sont très conservés dans l'ensemble des structures de GAPDH de bactéries et d'eucaryotes. Seule l'Arg231 présente, dans certains cas, une variabilité conformationnelle. Ainsi, la chaîne latérale de cet acide aminé, dans les structures de la GAPDH isolée à partir d'*E. coli* (Duée *et al.*, 1996) ou de *Trypanosoma cruzi* (Souza *et al.*, 1998 : structure non disponible dans la PDB) adopte une conformation distincte de celle rencontrée dans la structure de *B. stearothermophilus* et forme un pont salin avec le résidu Asp192 (ou, en position équivalente, l'Asp210 pour *T. cruzi*). Toutefois, d'autres structures de GAPDH issues de ces deux organismes montrent une conformation tout à fait classique de l'Arg231, comparable à celle observée dans la structure de *B. stearothermophilus* et ce, en l'absence ou en présence d'un ligand dans le site Ps (Yun *et al.*, 2000 pour la GAPDH d'*E. coli* : 1dc3, 1dc4, 1dc5 et 1dc6 ; Castilho *et al.*, 2003 et Ladame *et al.*, 2003 pour la GAPDH de *T. cruzi* : code pdb 1ml3 et 1qxs, respectivement).

La position et la composition du site Ps dans la structure des GAPDH d'archaea ne sont pas





*Figure I-9 : Vue stéréoscopique des sites de reconnaissance anionique Pi et Ps de la structure de B. stearothermophilus selon la position de deux ions sulfate (Skarzynski et al., 1987).*

les mêmes que celles rencontrées dans les autres GAPDH. Ce site est constitué de deux résidus arginine en position 166 et 167 de la séquence de *Sulfolobus solfataricus* (position 167-168 dans celle de *Methanothermus fervidus*) qui sont conservés dans l'ensemble des GAPDH d'archaea. Néanmoins, ce site n'ayant été identifié que dans une structure sous forme apoenzyme (Isupov *et al.*, 1999), on ne peut exclure que le NAD, une fois présent, puisse contribuer à la formation de ce site de reconnaissance anionique et influencer sa position.

#### ***A.5.1.2.2. Le site Pi***

Dans la structure de l'holoenzyme de *B. stearothermophilus*, le site Pi est partiellement occupé par un ion sulfate. Celui-ci est stabilisé par l'établissement de liaisons hydrogène avec les chaînes latérales des résidus conservés Ser148 et Thr208 ainsi qu'avec l'azote de la chaîne principale de la Gly209. Les chaînes latérales des résidus Arg195, Thr150 et la fonction hydroxyle en 2' du ribose sont indirectement impliquées dans la stabilisation de l'anion sulfate au travers de molécules d'eau cristallographiques. L'Arg195, contrairement à la Thr150, n'est pas conservée dans les séquences de GAPDH phosphorylantes mais est une singularité des GAPDH issues d'organismes thermophiles. Il faut noter que le résidu Arg194, invariant chez les organismes mésophiles, n'est pas structuralement équivalent à l'Arg195 des organismes thermophiles puisque la superposition des différentes structures exclut toute contribution de ce résidu à la formation du site Pi (Duée *et al.*, 1996).

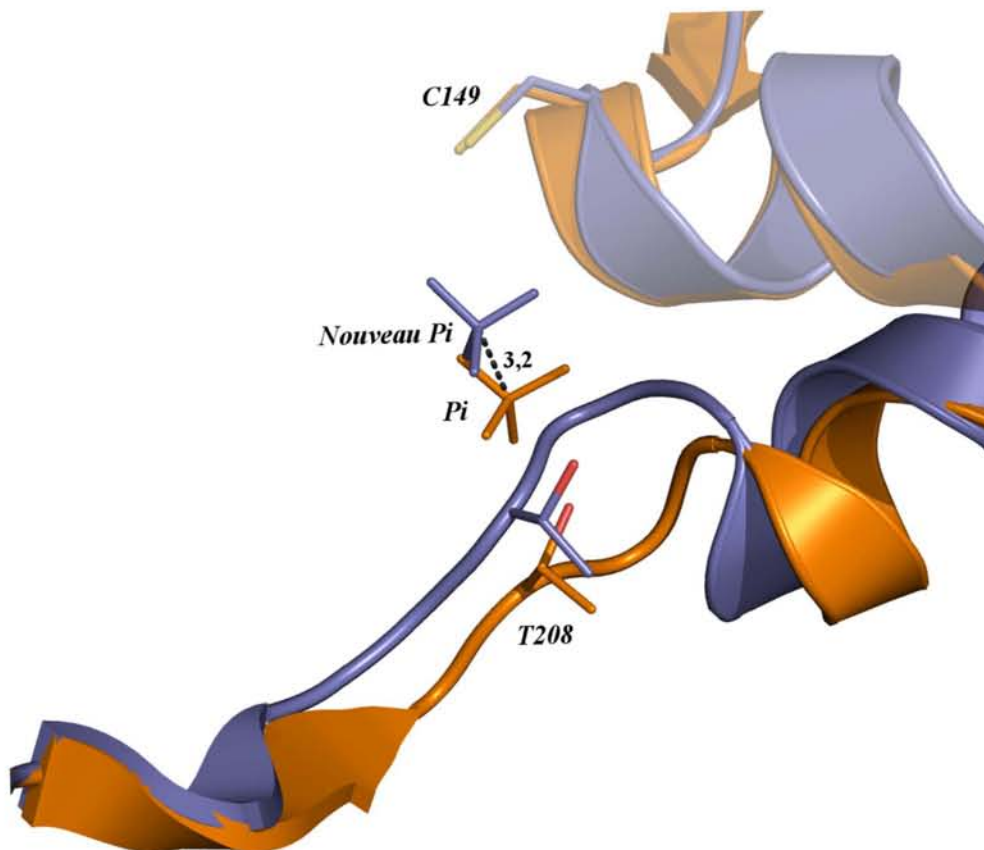
Si la position exacte du site Ps est très similaire dans l'ensemble des structures de GAPDH, le site Pi, lui, semble avoir une localisation moins précise, variable d'une structure à l'autre. Ainsi, la position de ce site de reconnaissance anionique semble liée à la conformation d'une boucle du site actif, la boucle  $\beta_2$ - $\alpha_2$  (Thr207-Ala210) qui contient les résidus Thr208 et Gly209 et borde le site Pi. Cette boucle a été observée sous deux conformations différentes. La première, celle présente dans la structure de l'enzyme de *B. stearothermophilus* et considérée comme la conformation classique, est la plus fréquemment rencontrée. La seconde est observée dans les structures de la GAPDH de *Thermotoga maritima* (code pdb 1hdg), *Leishmania mexicana* (1gyp et 1a7k), *Palinurus versicolor* (avec la cystéine carboxyméthylée : 1dss), et *E. coli* (avec la cystéine acylée par le G3P : 1dc4) (Korndörfer *et al.*, 1995 ; Kim *et al.*, 1995 et 1998 ; Song *et al.*, 1999 ; Yun *et al.*, 2000). Bien que cette conformation particulière ait été initialement obtenue en présence de sulfate pour l'enzyme de *T. maritima*, Kim *et al.* (1998) ont suggéré que cette conformation de la boucle  $\beta_2$ - $\alpha_2$  soit associée, dans la structure de l'enzyme de *L. mexicana*, à la présence d'un anion phosphate plus représentatif de l'état physiologique que l'ion sulfate. Les deux structures de *P. versicolor* et d'*E. coli* obtenues

ensuite, avec leur cystéine catalytique respectivement carboxyméthylée ou acylée par le G3P (hémithioacétal), semblent plutôt indiquer que cette conformation de la boucle  $\beta_2$ - $\alpha_2$  est liée à l'état de la cystéine. En effet, les structures des GAPDH de ces deux organismes ont, par ailleurs, été déterminées avec leur cystéine sous forme libre (Song *et al.*, 1998 : 1szj ; Duée *et al.*, 1996 : 1gad et 1gae ; Yun *et al.*, 2000 : 1dc3, 1dc5, 1dc6) et dans ces structures, la boucle 207-210 adopte une conformation comparable à celle de l'enzyme de *B. stearotherophilus*. Comparée à la conformation classique, la seconde conformation de cette boucle est déplacée vers la cystéine catalytique et délimite une nouvelle position du site Pi distante de plus de 3Å de la précédente et qui a été appelée "new Pi" ou nouveau site Pi (Figure I-10).

Chez les archaea, la position de l'anion sulfate assimilée au site Pi est différente du site Pi identifié dans la structure de *B. stearotherophilus*. Néanmoins, cette localisation est quasiment identique à celle du nouveau site Pi, l'atome de soufre des sulfates présent dans les structures de *S. solfataricus* (1b7g) et *M. fervidus* (1cf2) étant situé, après superposition des chaînes principales des trois structures, à environ 0,5Å de l'atome de phosphore du phosphate de la structure de *T. maritima* (1hdg). Ce site Pi est constitué, chez les archaea, des chaînes latérales des résidus Ser138 (équivalente à la Ser148 de *B. stearotherophilus*), Asn148 (qui contribue grâce à l'atome d'azote de sa chaîne principale), His192 et His193 qui sont strictement conservés dans les séquences des GAPDH d'archaea.

#### ***A.5.1.2.3. Apport de la mutagenèse-dirigée à l'étude des sites de reconnaissance anionique***

L'ensemble des résidus de l'enzyme de *B. stearotherophilus* dont la chaîne latérale contribue à la formation de l'un ou l'autre des sites de reconnaissance anionique ont été substitués par mutagenèse dirigée et les propriétés cinétiques des mutants comparées à celles de l'enzyme de type sauvage (Corbier *et al.*, 1989 ; Corbier *et al.*, 1994 ; Michels *et al.*, 1996). En cinétique classique, les résultats obtenus ne montrent que peu de changement pour les constantes michaeliennes, signe que la mutation d'un seul résidu ne perturbe que très légèrement la fixation des substrats. On peut noter toutefois que la mutation de la Ser148 en alanine, résidu participant à la formation du site Pi, affecte en premier lieu le  $K_M$  du phosphate inorganique (et dans une moindre mesure les  $K_m$  du 1,3-DPG et du G3P), laissant supposer que ce résidu interagit avec ce site, en accord avec les résultats de modélisation moléculaire (Corbier *et al.*, 1989). Si l'altération des sites Pi et Ps ne semble avoir qu'une influence limitée sur l'affinité des substrats, la mutation des résidus du site Ps conduisent néanmoins à des altérations très fortes des constantes catalytiques (diminution



**Figure I-10 : Comparaison des deux conformations de la boucle  $\beta_2$ - $\alpha_2$  (207-210) et relocalisation du site Pi.** Les structures de la GAPDH de *B. stearothermophilus* (*Igd1*) et de *T. maritima* (*Ihdg*) sont représentées respectivement en orange et en bleu, avec leur anion sulfate fixé respectivement dans les sites Pi et nouveau Pi. La distance séparant les deux anions est indiquée en Å.

d'un facteur 27, 4667, 467 et 2333 respectivement pour les mutants T179A, T179M, R231G et R231L par rapport à l'enzyme de type sauvage). Les constantes catalytiques des mutants du site Pi sont, elles, moins perturbées (facteur 9,3, 6,5, 2,4 et 11 pour les mutants S148, T208A, T150A et R195L, respectivement).

Les études, en stopped-flow, des étapes individuelles de la réaction montrent que du point de vue cinétique, le processus global conduisant à la formation de NADH n'est pas altéré pour les mutants du site Ps par rapport à l'enzyme sauvage alors que ce processus est ralenti de 3 à 24 fois pour les mutants du site Pi, suggérant que le groupement phosphate du substrat interagit avec les résidus du site Pi lors de l'étape d'acylation. En l'absence de méthode directe pour étudier l'étape de phosphorylation, les constantes de premier ordre associées à cette étape ont été prédites en se basant sur la détermination des constantes de second ordre. Pour la plupart des mutants des sites Pi et Ps (à l'exception du mutant T150A), ces constantes sont proches du  $k_{cat}$ , indiquant que l'étape de la phosphorylation, ou un processus qui lui est associé, devient limitante pour les mutants.

En conclusion, ces études tendraient donc à soutenir l'hypothèse d'une interaction entre le phosphate du substrat et le site Pi durant l'étape d'oxydo-réduction. Pourtant, le fait que l'étape de phosphorylation (ou une étape associée) devienne limitante pour les mutants des deux sites ne permet aucune conclusion quant à un éventuel changement de site du phosphate du substrat à partir de Pi vers Ps.

### ***A.5.2. Structures cristallographiques de complexes enzyme-substrat***

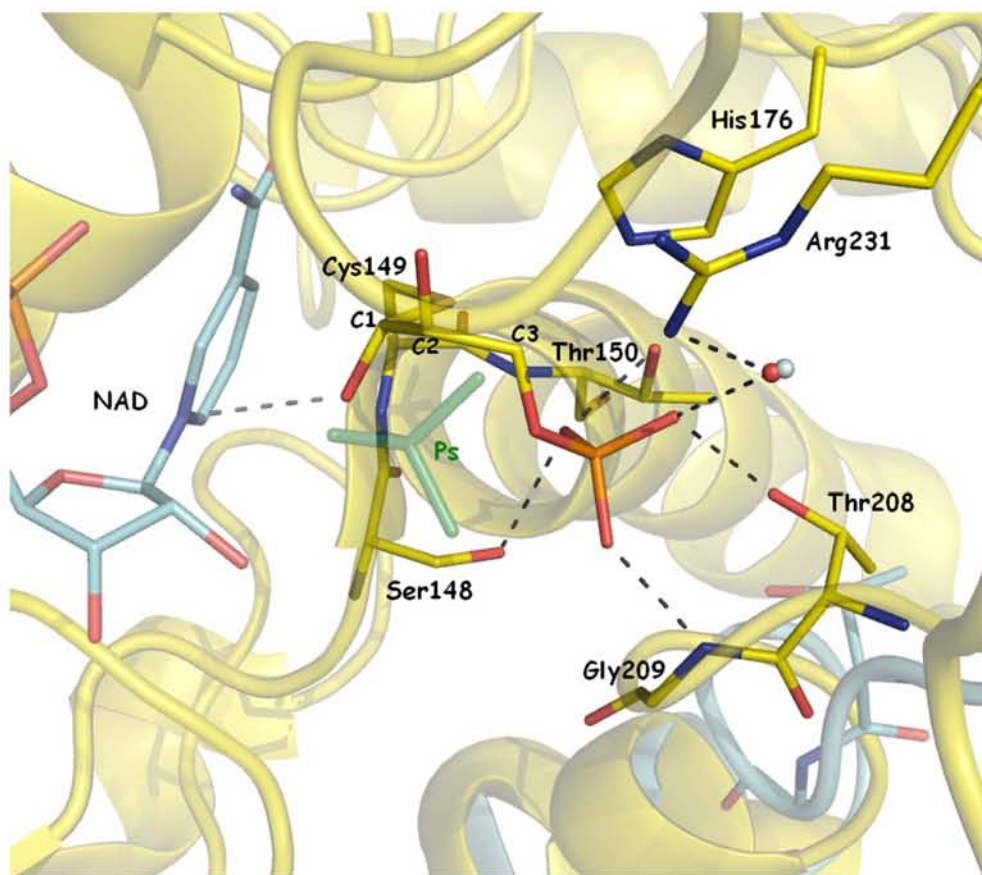
Comme nous l'avons vu dans le paragraphe précédent, malgré les nombreuses études qui y ont été consacrées, le rôle de chacun des deux sites de reconnaissance anionique au cours de la catalyse demeure controversé. S'il semble probable que le phosphate inorganique réalise son attaque nucléophile depuis le nouveau site Pi (attaque plus favorable que depuis le site Pi classique d'un point de vue géométrique), la localisation du phosphate en C3 du substrat reste incertaine et l'on ne peut exclure qu'il interagisse alternativement avec les deux sites en fonction de l'état d'avancement de la réaction. Ainsi, des études structurales ont eu pour but de résoudre la structure de la GAPDH en complexe avec des substrats ou inhibiteurs plus ou moins représentatifs des différents états physiologiques adoptés par cette enzyme au cours de la catalyse. Seuls les résultats obtenus pour quelques unes de ces structures, jugées comme étant les plus représentatives, seront présentés ici.

### *A.5.2.1. Structure de l'intermédiaire hémithioacétal de la GAPDH d'E. coli sous forme apoenzyme (Yun et al., 2000 : 1dc3-6)*

Park et collaborateurs, en 2000, ont résolu les structures de la GAPDH d'*E. coli* sous les formes apoenzyme, holoenzyme et sous la forme d'un complexe binaire covalent entre le G3P et l'apoenzyme représentant l'intermédiaire hémithioacétal (Figure I-11). Les observations tirées de la comparaison des structures apo et holoenzyme sont comparables à celles réalisées pour l'enzyme de *B. stearothermophilus* (Skarzynski and Wonacott, 1988), à ceci près que dans le cas de l'enzyme d'*E. coli*, la distance entre le thiol de la cystéine catalytique et l'atome N $\epsilon_2$  de l'His176 augmente très légèrement lors de la fixation du cofacteur.

La stratégie développée par les auteurs pour résoudre la structure de l'intermédiaire tétraédrique, qui possède une durée de vie extrêmement courte, repose sur l'absence du cofacteur. En effet, en l'absence de ce dernier, le transfert d'hydrure est impossible et la réaction ne peut se poursuivre au-delà de l'étape correspondant à l'attaque nucléophile du thiol et à la formation concomitante de l'hémithioacétal. L'analyse de cette structure montre un changement de conformation de la boucle 207-212 délimitant le site de reconnaissance anionique Pi. Ainsi, alors que dans les formes apo et holoenzyme, cette boucle présente une conformation comparable à celle de l'enzyme de *B. stearothermophilus*, dans la structure de l'hémithioacétal, la boucle s'est déplacée vers le centre du site actif (avec un déplacement maximum de plus de 3,5Å pour le Ca de la Gly209) pour former le nouveau site Pi initialement décrit par Korndörfer *et al.* (1995) pour la GAPDH de *T. maritima*. L'intermédiaire hémithioacétal, parfaitement défini dans la densité électronique, est fixé dans le site actif de l'enzyme avec son groupement phosphate en C3 localisé dans ce nouveau site Pi et stabilisé par des liaisons hydrogène avec les résidus Ser148 (O $\gamma$ ), Thr150 (N et O $\gamma_1$ ), Thr208 (O $\gamma_1$ ) et Gly209 (N). Bien que les auteurs en concluent que le substrat soit positionné dans ce site pour l'étape d'oxydo-réduction pour ensuite basculer vers le site Ps et ainsi libérer le site Pi pour la phosphorolyse, le point faible de cette structure repose sur l'absence du cofacteur. En effet, il est démontré que le NAD est un élément essentiel à la constitution d'un site Ps pleinement fonctionnel et donc, en son absence, la fixation du groupement phosphate en C3 dans le site Pi pourrait s'avérer être un artéfact.

Une autre observation importante découlant de l'étude de cette structure est le positionnement de la fonction hydroxyle portée par le C1 du substrat. Ainsi, malgré le nombre important des travaux impliquant l'His176 dans la stabilisation de l'intermédiaire hémithioacétal et le transfert d'hydrure, l'hydroxyle porté par le C1 pointe, dans cette structure, dans une direction qui



**Figure I-11 : Structure de l'intermédiaire hémithioacétal de la GAPDH d'*E. coli* (en jaune).** En bleu clair apparaissent la molécule de NAD et la boucle 207-210 (en conformation classique) issues de la structure de l'holoenzyme d'*E. coli* (1dc6). En vert est représenté l'anion sulfate du site Ps de l'holoenzyme de *B. stearothermophilus* (1gd1, Skarzynski et al., 1987). (D'après Yun et al., 2000)

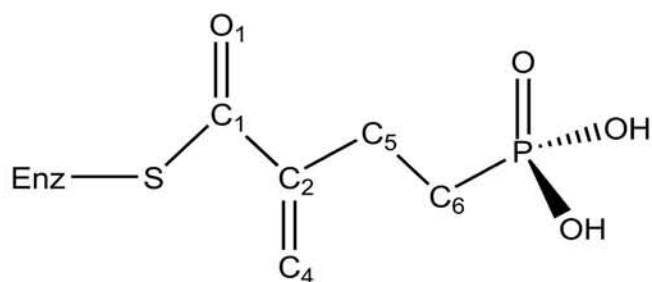
exclut totalement l'intervention de l'His176. En se basant sur une superposition de cette structure avec celle de l'holoenzyme, Yun *et al.* ont alors proposé que cet hydroxyle puisse être stabilisé, sous forme oxyanion, par une interaction avec l'atome d'azote du nicotinamide (situé à 3,2Å) remettant en cause le mécanisme réactionnel accepté jusqu'alors. L'absence du cofacteur dans la structure de l'hémithioacétal tend, une fois encore, à relativiser ces interprétations. De plus, la géométrie et la constitution du site oxyanion potentiel, qui ne serait composé que de l'atome d'azote du nicotinamide (néanmoins porteur d'une charge positive), associées aux preuves de l'implication du résidu His176 dans le mécanisme réactionnel des GAPDH, rendent cette hypothèse peu probable.

#### *A.5.2.2. Structure d'un complexe ternaire de la GAPDH de T. cruzi avec le NAD et un inhibiteur covalent de la cystéine catalytique (Castihlo et al., 2003 : 1ml3)*

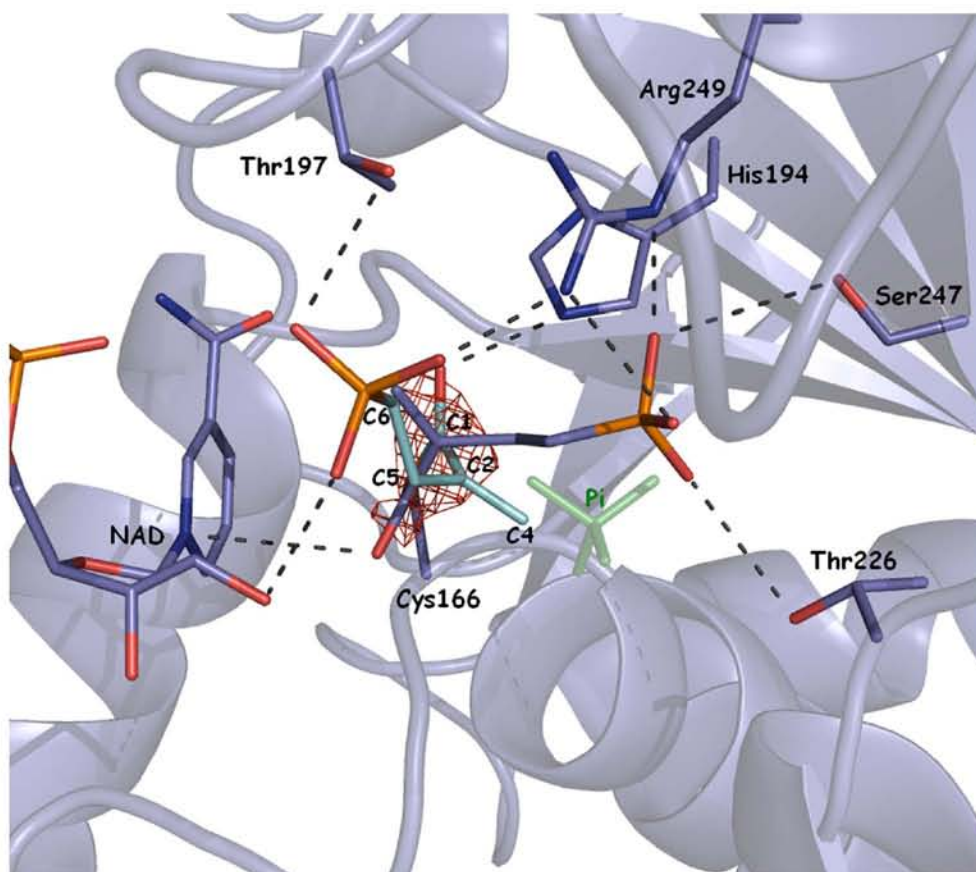
Une étude publiée en 2003, portant sur la GAPDH de *T. cruzi* a permis de résoudre la structure de cette enzyme en complexe ternaire avec le cofacteur et un inhibiteur lié covalamment par une liaison thioester à la cystéine catalytique (Cys166 dans la séquence de cet organisme). Cet inhibiteur est un analogue du substrat physiologique de l'enzyme, le G3P, de type phosphonate (liaison C-P) dans lequel la fonction hydroxyle en C3 a été remplacée par un groupement méthylène qui a pour effet de placer dans un même plan les atomes C1, C2, C4 et C5 (Figure I-12).

L'inhibiteur, présent dans trois des quatre sous-unités, est observé sous deux conformations alternatives, appelées conformation Pi ou conformation Ps selon que son groupement phosphate en C3 est fixé dans l'un ou l'autre des deux sites de reconnaissance anionique. La sous-unité B présente les deux conformations à la fois (Figure I-13), alors que les monomères C et D présentent les conformations Pi et Ps, respectivement. Il est important de noter que malgré la présence d'un groupement phosphate en C3 de l'inhibiteur, le changement conformationnel de la boucle  $\beta_2$ - $\alpha_2$  précédemment présenté dans le cas de l'intermédiaire hémithioacétal d'*E. coli* n'est pas retrouvé dans ce cas. De plus, la localisation du site Pi, caractérisé dans cette structure de *T. cruzi* par la position du groupement phosphonate, n'est pas la même que celle du nouveau site Pi, ni même celle du site Pi de la structure de *B. stearothermophilus*. Les distances entre l'atome de phosphore de l'inhibiteur et l'atome de soufre de la structure décrite par Korndörfer *et al.* (1hdg : nouveau site Pi) ou l'atome de soufre de la structure décrite par Skarzynski *et al.* (1gd1 : site Pi) après superposition des chaînes principales des trois structures sont respectivement de 3,4 et 4,6 Å. Ce site de reconnaissance anionique particulier fait intervenir les résidus Thr226 (O $\gamma_1$ ) et Arg249 (N $\epsilon$  et N $\eta_2$ ), respectivement équivalents aux résidus Thr208 et Arg231 de la GAPDH de *B. stearothermophilus*,





**Figure I-12 : Formule développée de l'inhibiteur (3-(p-nitrophenoxycarbonyl)-3-éthylène propyl dihydroxyphosphonate) et nomenclature des atomes.**



**Figure I-13 : Structure d'un complexe ternaire de la GADPH de *T. cruzi* avec le NAD et un inhibiteur covalent (bleu foncé). Les deux conformations alternatives de l'inhibiteur, dites Pi (bleu foncé) et Ps (bleu clair) sont représentées. Seuls les atomes de carbone de la conformation Ps ont été identifiés. En vert est représenté l'anion sulfate du nouveau site Pi de la structure de *T. maritima* (Korndörfer et al., 1995). La carte de densité électronique de type Fourier différence représentée autour des atomes S, C1 et O1 des deux conformations de l'inhibiteur présente un niveau de contour de  $-3\sigma$  (d'après Castihlo et al., 2003)**

mais aussi l'équivalent de l'His176, l'His194 (en interaction  $\pi$ ) et la Ser247 ( $O\gamma$ ) dont l'équivalent structural est l'Ala229 dans les autres GAPDH. Cette localisation particulière du site Pi pourrait être due aux contraintes géométriques inhérentes à la molécule inhibitrice du fait de l'existence, dans sa chaîne carbonée, du carbone C2 en hybridation  $sp^2$ . Le site Ps, en revanche, est quasiment équivalent à celui décrit pour la GAPDH de *B. stearothermophilus*, avec une distance entre les atomes centraux des anions d'environ 1 Å seulement et la contribution des résidus Thr197 ( $O\gamma_1$ ) (équivalente à Thr179), Arg249 ( $N\epsilon$  et  $N\eta_2$ ) et de la fonction hydroxyle en 2' du ribose du nicotinamide.

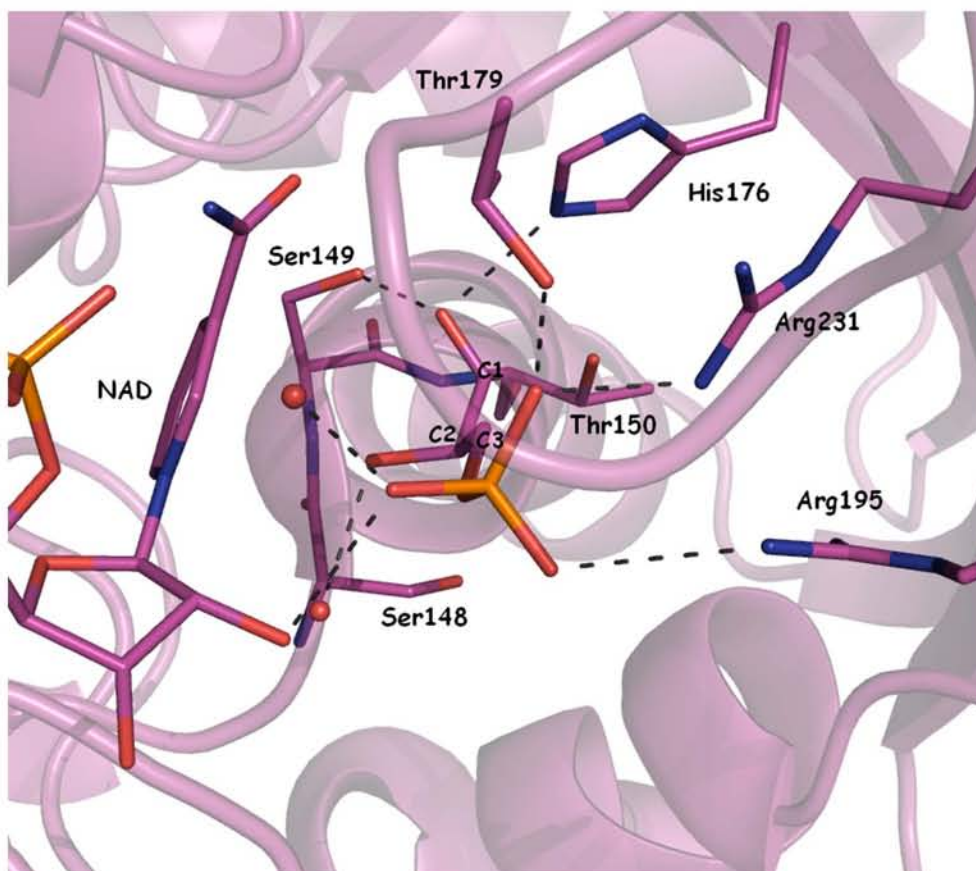
L'information principale apportée par cette structure repose sur le fait que l'inhibiteur se fixe dans les deux sites de reconnaissance anionique au sein des différentes sous-unités de l'enzyme. Cette observation tendrait ainsi à soutenir la possibilité d'un changement de site du groupement phosphate en C3 du substrat au cours de la catalyse sans toutefois apporter plus de précision. La fonction carbonyle portée par l'atome de carbone C1 de l'inhibiteur est représentative de celle portée par le G3P et peut donc apporter des informations sur le mécanisme réactionnel de l'enzyme. Dans la conformation Pi, la position de cette fonction est très similaire à celle de l'hydroxyle de l'intermédiaire hémithioacétal de la GAPDH d'*E. coli*. En s'appuyant sur la similarité avec la structure de l'hémithioacétal, les auteurs relèvent une interaction possible entre l'atome O1 et l'atome d'azote du nicotinamide distant de 3,6 Å, mais soulignent que cette distance est plus longue que celle prédite par Yun *et al.* (2000). Dans la conformation Ps, ce carbonyle en C1 de l'inhibiteur pointe dans la direction opposée et forme une liaison hydrogène avec l'His194. Une telle interaction pourrait très vraisemblablement être présente lors de la formation d'un intermédiaire tétraédrique fixé en Ps (Castilho *et al.*, 2003). Toutefois, et malgré le fait que les auteurs n'en fassent pas mention dans la publication, l'observation des cartes de densité électronique disponibles dans la banque de données structurales PDB montre que, contrairement au reste de la molécule inhibitrice, cette fonction carbonyle est très mal définie dans la densité. Les cartes Fourier différence présentent également de forts pics de densité négative à ce niveau ce qui relativise les interprétations faites à partir des positions de cette fonction carbonyle.

#### *A.5.2.3. Structure du complexe ternaire de type Michaelien de la BstGAPDH (GAPDH - NAD - G3P) (Didierjean et al., 2003)*

La structure de la GAPDH de *B. stearothermophilus* en complexe ternaire avec le NAD et son substrat physiologique, le D-G3P, a été publiée en 2003. Les auteurs ont utilisé deux mutants de

l'enzyme dans lesquels la cystéine catalytique avait été remplacée par une sérine (faible activité résiduelle) ou une alanine (mutant inactif). Ces structures de complexes ternaires de la GAPDH ont été les premières à mimer le complexe michaelien de l'enzyme, c'est-à-dire, à être représentatives de l'état fondamental de l'enzyme juste avant l'acte catalytique.

Dans le cas des deux mutants, le trempage des cristaux de l'holoenzyme en présence de G3P fait apparaître une densité électronique au niveau du site actif, densité qui a été attribuée par les auteurs à la présence d'une molécule de D-G3P liée de façon non covalente à l'enzyme. Les deux complexes ternaires obtenus (mutant C149S ou C149A) montrent que le groupement phosphate en C3 du substrat est localisé dans le site de reconnaissance anionique Ps et confirment ainsi l'hypothèse première émise par Moras *et al.* (1975), d'une fixation initiale du substrat dans ce site. La position du groupement phosphate tout comme le réseau de liaisons hydrogène qu'il établit avec l'enzyme sont comparables à ceux décrits par Skarzynski *et al.* pour l'holoenzyme de type sauvage et impliquent la contribution des chaînes latérales des résidus Thr179, Arg231 et de la fonction 2'OH du ribose du nicotinamide (Figure I-14). On note également que l'Arg195 (N $\eta$ 1) à une distance de 3,14 Å pourrait contribuer directement à la stabilisation du phosphate en C3. Dans les deux structures, la boucle 207-210 délimitant le site Pi, adopte une conformation classique, identique à celle précédemment observée dans les formes apo et holoenzyme de la GAPDH de cet organisme. Le site Pi correspondant ne présente aucune trace d'un quelconque anion (Didierjean *et al.*, 2003). Malgré l'absence de la cystéine catalytique qui pourrait avoir des conséquences sur le positionnement de l'extrémité aldéhydique du substrat, on peut noter que dans les deux structures du complexe ternaire, le carbonyle porté par l'atome C1 du G3P forme une liaison hydrogène avec l'atome N $\epsilon_2$  de l'His176, ce qui semble en accord avec les rôles potentiels attribués à ce résidu au cours de la catalyse (assistance acide et basique lors de la formation des intermédiaires hémithioacétal et du transfert d'hydrure, respectivement). La fonction hydroxyle portée par le carbone C2 du substrat est, quant à elle, stabilisée par une liaison hydrogène avec une molécule d'eau, elle-même en interaction avec la fonction hydroxyle en 2' du ribose du nicotinamide. Contrairement à la conformation du G3P dans l'intermédiaire hémithioacétal, les fonctions carbonyle en C1 et hydroxyle en C2 n'adoptent pas une conformation *trans*, l'angle de torsion O1-C1-C2-O2, égal à 46,6°, étant plus proche d'une conformation éclipsée. Bien que les auteurs ne le mentionnent pas dans la publication, la fonction hydroxyle en C2 pointe en direction du cycle nicotinamide à une distance correcte, quoique un peu longue (3,8Å), de l'atome d'azote du nicotinamide pour envisager une possible interaction entre ces deux partenaires dans le véritable complexe michaelien.



**Figure I-14 : Structure du complexe ternaire de type michaelien de la GAPDH de *B. stearothermophilus* ( $Enz-NAD^+-G3P$ ) avec le phosphate en C3 du substrat fixé dans le site de reconnaissance anionique Ps (d'après Didierjean et al., 2003).**

# **PARTIE B : L'ÉRYTHROSE-4-PHOSPHATE DÉSHYDROGÉNASE (E4PDH)**

## **B.1. Introduction**

L'érythrose-4-phosphate déshydrogénase (E4PDH) (EC 1.2.1.72) est une oxydoréductase de la famille des aldéhydes déshydrogénases. Cette enzyme, caractérisée chez trois organismes uniquement (*Escherichia coli*, *Vibrio cholerae* et *Giardia lamblia*), est codée par les gènes initialement nommés *gapB* (Alefounder and Perham, 1989), puis renommés *epd* (Zhao *et al.*, 1995 ; Carroll *et al.*, 1997). Le nom initial du gène provient de la forte identité de séquence existant entre son produit d'expression et les GAPDH glycolytiques, produits des gènes *gapA* (41,6 % d'identité entre les deux séquences protéiques chez *E. coli*) (Alefounder and Perham, 1989). Cependant, des souches d'*E. coli* dans lesquelles une mutation non sens a été introduite dans le gène *gapA* sont incapables de se développer sur glucose et il est clair que seul le gène *gapA* code pour la forte activité GAPDH détectée chez *E. coli* (Hillman and Fraenkel, 1975). Le produit du gène *gapB* a été caractérisé plus tard par Zhao *et al.* (1995) comme possédant une activité de type érythrose-4-phosphate déshydrogénase (E4PDH). Ces auteurs proposèrent que cette enzyme soit impliquée dans la voie de biosynthèse du pyridoxal 5'-phosphate (PLP), cofacteur très important dans la biosynthèse des acides aminés (Lam and Winkler, 1990 ; Zhao *et al.*, 1995 ; Yang *et al.* 1998).

Néanmoins, si cette hypothèse semble aujourd'hui couramment acceptée, le fait que des souches mutantes d'*E. coli*, délétées du gène *epd*, puissent présenter ou non une auxotrophie vis-à-vis du PLP (Zhao *et al.*, 1995 ; Seta *et al.*, 1997) suscite des questions quant au rôle physiologique de cette enzyme.

## ***B.2. Réaction catalysée par l'E4PDH et mécanisme réactionnel***

L'E4PDH catalyse l'oxydation irréversible du *D*-érythrose-4-phosphate (*D*-E4P) en 4-phosphoérythronate (4PE) en présence de NAD et d'une molécule d'eau. L'alignement de la séquence de l'E4PDH d'*E. coli* (*EcoE4PDH*) avec la séquence consensus des GAPDH glycolytiques (Boschi-Muller *et al.*, 1997 ; Figure I-15) met clairement en évidence la conservation, dans les E4PDH, des principaux résidus impliqués dans la catalyse chez les GAPDH. Ainsi, la présence d'une cystéine en position 149 et d'une histidine en position 176, tout comme la proximité structurale existant entre les substrats des deux enzymes (l'E4P et le G3P ne différant que par un unique groupement CHOH supplémentaire dans le cas de l'E4P), a permis de proposer pour l'E4PDH un mécanisme réactionnel à deux étapes, similaire à celui des GAPDH (Boschi-Muller *et al.*, 1997).

La première étape, dite d'acylation, est une oxydo-réduction mettant tour à tour en jeu l'attaque nucléophile de la cystéine catalytique sur le carbone aldéhydique du substrat, le passage par un intermédiaire tétraédrique covalent, puis un transfert d'hydrure vers le NAD conduisant à la réduction du cofacteur et à la formation d'un intermédiaire covalent de type acylenzyme entre l'enzyme et le substrat. La seconde étape, dite de désacylation, est, elle, différente de celle catalysée par les GAPDH. En effet, le produit de la réaction catalysée par l'E4PDH, le 4-phosphoérythronate (4PE), résulte d'une oxydation non phosphorylante et irréversible du substrat, au cours de laquelle une molécule d'eau (et non une molécule de phosphate inorganique) réalise l'attaque nucléophile sur le thioacylenzyme. Après le passage par un nouvel intermédiaire tétraédrique, le produit de la réaction est libéré. Cette enzyme est donc capable d'une part, d'utiliser efficacement un substrat à 4 atomes de carbone (E4P) au lieu de 3 (G3P) dans le cas des GAPDH et d'autre part, d'hydrolyser efficacement l'intermédiaire thioacylenzyme, contrairement aux GAPDH pour lesquelles l'efficacité de l'hydrolyse est très faible par rapport à celle de la phosphorylation (Michels *et al.*, 1996).

Les études en stopped-flow réalisées par Boschi-Muller *et al.* (1997) sur l'*EcoE4PDH* ont,

```

GapB      1      10      20
GAPDHs    TVRVAINGGFGRIGRNVVRALYESGRRRAEIT
          < βA > < αB >

GapB      30      40      50
GAPDHs    VVAINELADAAGMAHLLKYDTSHGRFAWEV
          < βB > < αC > <

GapB      60      70      80
GAPDHs    RQERDQLFVGDDAIRVLHERSLQSLPWREL
          β' > < β'' > < βC > < αD >

GapB      90      100     110
GAPDHs    GVDVVLDCTGVYGSREHGEAHIAAGAKKVL
          < βD > < αE > < βE >

GapB      120     130     140
GAPDHs    FSHPGSNDLDATVVYGVNQDQLRAEHRIVS
          > < β''' > <

GapB      150     160     170
GAPDHs    NASCTTNCIIPVIKLLDDAYGIESGTVTTI
          βF > < α1 > > < β1 >

GapB      180     190     200
GAPDHs    HSAMHDQQVIDAYHPDLRRTRAASQSIIPV
          > < S loop > < β2 >

GapB      210     220     230
GAPDHs    DTKLAAGITRFPQFNDRFEAIAVRVPTI
          > < α2 > < β3 > >

GapB      240     250     260
GAPDHs    VTAIDLSVTVKKPVKANEVNLLLQKAAQGA
          < β4 > < α3 > >

GapB      270     280     290
GAPDHs    FHGIVDYTELPLVSVDFNHDPHSAIVDGTQ
          < β5 > < β6 >

GapB      300     310     320
GAPDHs    TRVSGAHLIKTLVWCDNEWGFANRMLDTTL
          > < β7 > >< α4 > >

GapB      330
GAPDHs    AMATVAFR

```

**Figure I-15 : Alignement de la séquence de l'EcoE4PDH avec une séquence consensus des GAPDH. La séquence consensus des GAPDH est issue de l'alignement de 95 structures de GAPDH et contient les résidus conservés dans 95 % des séquences. D'après Boschi-Muller et al. (1997).**

par ailleurs, permis de montrer que l'étape limitante était située à un stade ultérieur à la réduction du cofacteur. L'absence d'effet isotopique de solvant suggère en outre que cette étape limitante est associée, non pas à l'attaque de la molécule d'eau, mais plutôt à la libération d'un des produits de la réaction (4PE ou NADH). Le fait que l'hydrolyse elle-même ne soit pas limitante semble indiquer qu'il existe, dans le site actif de l'enzyme, un acide aminé capable de positionner et/ou d'activer cette molécule d'eau, et donc de jouer un rôle équivalent au résidu Glu250 dans les GAPDH non phosphorylantes (Marchal *et al.*, 2000).

De façon intéressante, Boschi-Muller *et al.* (1997) ont également pu détecter, pour l'E4PDH d'*E. coli*, une faible activité phosphorylante avec le G3P ( $k_{cat}$  de  $0,12 \text{ s}^{-1}$  contre  $20 \text{ s}^{-1}$  avec l'E4P). De même, la GAPDH de *B. stearothermophilus* peut utiliser l'E4P en tant que substrat avec une faible efficacité et dans une réaction de type phosphorylante ( $k_{cat}$  de  $0,10 \text{ s}^{-1}$  comparé à  $76 \text{ s}^{-1}$  avec le G3P). Ces activités croisées entre les deux enzymes et leur substrat, associées à la possibilité pour l'E4PDH de catalyser, malgré une faible efficacité, une réaction de type phosphorylante indiquent qu'il existe probablement de très fines différences structurales entre les sites actifs de ces deux enzymes apparentées qui leur permettent d'utiliser de façon très spécifique et efficace leur propre substrat naturel.

Sur la base de la ressemblance existant à la fois entre les deux enzymes et leur substrat, il semble évident que le site actif de l'E4PDH présente (au moins) un site de reconnaissance anionique. En effet, la réaction catalysée par cette enzyme avec l'E4P n'est pas de type phosphorylante et seul le groupement phosphate en C4 du substrat a besoin d'être reconnu. Néanmoins, la faible activité de type phosphorylante mise en évidence pour l'E4PDH envers le G3P pourrait laisser supposer l'existence d'un second site.

## ***B.3. Structures de l'E4PDH***

### ***B.3.1. Structure primaire***

58 séquences protéiques d'E4PDH sont actuellement disponibles dans la banque de données UniProt (The UniProt Consortium). La plupart des séquences présentent environ 340 acides aminés, mais quelques organismes pourraient avoir jusqu'à 371 résidus dans leur séquence d'E4PDH. L'alignement global des 58 séquences montre une insertion de 5 à 10 résidus au niveau de la position 71 (numérotation de la séquence d'*E. coli*) qui semble être un fait particulier de l'espèce



*Vibrio* (5 organismes) mais aussi de l'organisme *Photobacterium profundum*. Les variations principales au niveau des séquences se retrouvent au niveau du C-terminal qui présente ou non des extensions. Un alignement de 6 séquences d'E4PDH représentatives des espèces les plus souvent représentées dans la banque de données avec la séquence consensus des GAPDH glycolytiques (Boschi-Muller *et al.*, 1997) est présenté en Figure I-16. Cet alignement met clairement en évidence la forte conservation des séquences protéiques de l'E4PDH. Ainsi, la séquence de l'E4PDH d'*E. coli* présente 91,6, 64,2, 100, 67,9 et 82,2% d'identité de séquence respectivement avec les séquences de *Salmonella typhi*, *Shewanella oneidensis*, *Shigella flexneri*, *Vibrio cholerae* et *Yersinia pestis*. Les deux résidus impliqués dans la catalyse, la Cys149 et l'His176 (numérotation de la GAPDH) sont strictement conservés. Les séquences de l'E4PDH et de la GAPDH (*gapA*) d'*E. coli* présentent 40,6% d'identité. Comparée à la séquence consensus des GAPDH glycolytiques établie par Boschi-Muller *et al.*, 1997, les séquences d'E4PDH présentent une conservation quasi systématique des résidus appartenant à cette séquence consensus. On peut toutefois noter quelques différences potentiellement informatives pour le mécanisme réactionnel de cette enzyme, comme par exemple la substitution des résidus Thr179 et Gly209 (invariants dans les séquences de GAPDH et qui interviennent dans la formation des sites de reconnaissance anionique) par des résidus méthionine et lysine, qui pourrait avoir des conséquences en termes d'existence et/ou de fonctionnalité des sites de reconnaissance anionique. En ce qui concerne le site de fixation du cofacteur, les résidus dont la chaîne latérale est directement impliquée dans la fixation du NAD dans les GAPDH (Asp32, Asn180, Ser119 et Asn313) sont conservés dans les séquences d'E4PDH ou remplacés par des résidus pouvant jouer un rôle équivalent (Glu32, Asn/His180, Ser119, Asn313).

### ***B.3.2. Structure tridimensionnelle***

Comparativement à la GAPDH pour laquelle plus 50 structures cristallographiques ont été résolues et sont disponibles dans la Protein Data Bank, seule une structure d'E4PDH était connue au début de cette étude, celle d'*Escherichia coli*, résolue au LCM3B (Kauffmann, 2000). C'est cette structure sous forme apoenzyme à 2,7 Å de résolution qui sera décrite dans la suite de cette présentation.

#### ***B.3.2.1. Structure quaternaire***

Tout comme la GAPDH, l'E4PDH est un homotétramère dont la masse moléculaire varie autour de 145 kDa et dont les sous-unités sont arrangées selon une symétrie 222. La symétrie

```

      1      10      20      30      40      50      60      70
ECOLI TVRVAINGFGRIGRNVVRLALYESGRRAEITVVAINELADAAGMAHLLKYDTSHGRTFAWEVRQERDQLFVGD-----DA
SALTI TVRIAINGFGRIGRNVVRLALYESGRRAEITVVAINELADAAGMAHLLKYDTSHGRTFAWEVRHEREQLFVGD-----DV
SHEON -IRVAINGYGRIGRSILRALYESGKRQIQIVAINELAKPEAIIHLTQYDTHGRTFQPRVKLVLDGQMIGD-----DT
SHIFL TVRVAINGFGRIGRNVVRLALYESGRRAEITVVAINELADAAGMAHLLKYDTSHGRTFAWEVRQERDQLFVGD-----DA
VIBCH -LRVAINGFGRIGRNVLRVAVYESGKRDRIQVVAVNELAKPDAMAHLLQYDTSHGRTFGKKISHDQQHIYVHHQNGEYDS
YERPE AIRIAINGFGRIGRSVLRALYESGRRAEISVVAINELASAEGBMAHLLKYDSSHGRTFAWDVRQECDSLIVGD-----DI
Cons  :*:****:*****.:**:*:**:*. * :*:*****. .: ** :*:***:**** : : : . *
GAPDHs -----GFGRIGR---R-----V--ND-----KYDS--G-----G-----

      80      90      100     110     120     130     140
ECOLI IRVLHEKSLQSLPWRELGVDDVLDCTGVYGSREHGEAIIAAGAKKVLFSHPGSNDLDATVVYGVNQDLRAEHRIVSN
SALTI IRILHEKTLADLPWRELGVDDVLDCTGVYGNREHGEAIIAAGAKKVLFSHPGSNDLDATVVYGVNQDLRAEHRIVSN
SHEON IKIFHEKDPAKLPWRELIDIVYEATGAILDRQSCAEIIHAGAKQVLISHPSSADVDGTIVYGVNHDLLRAEHTVVSN
SHIFL IRVLHEKSLQSLPWRELGVDDVLDCTGVYGSREHGEAIIAAGAKKVLFSHPGSNDLDATVVYGVNQDLRAEHRIVSN
VIBCH IRILHEKSEIPLPWRLGVDDVLDCTGVYGCQEDGQQHIDAGAKLVLFSHPGASDLNTIYGVNHETLTAEHKIVSN
YERPE IRLIHQSEIEQLPWSELGIDVLDLDCSGVYGSREDGEAHVASGAKKVLFAHPGGHDLATVVYGVNHQDLRAEHRIVSN
Cons  *:::* ***:**:* **:* **:* **:* **:* **:* **:* **:* **:* **:* **:* **:* **:* **:* **:* **:*
GAPDHs -----P---W---G---E-TG-F---A--H---GAKKV---P-----V-GVN---Y-----SN

     150     160     170     180     190     200     210     220
ECOLI ASRTTNCIIPVIKLLDDAYGIESGTVTTISSAMHDQQVIDAYHPDLRRTAAASQSIIPVDTKLAAGITRFFPQFNDRF
SALTI ASRTTNCIIPVIKLLDDAYGIESGTVTTISSAMNDQQVIDAYHSDLRRTAAASQSIIPVDTKLAAGITRIFPQFNDRF
SHEON ASRTTNCIIPVVIDVLDKHFVGVKSGAITTISSAMNDQQVIDAYHDDLRTAAAGQSIIPVDTKLARGIERILPHMKDKF
SHIFL ASRTTNCIIPVIKLLDDAYGIESGTVTTISSAMHDQQVIDAYHPDLRRTAAASQSIIPVDTKLAAGITRFFPQFNDRF
VIBCH GSRTTNCIIPVPIKVLDDAFGIDSGTITTISSMNDQQVIDAYHNDLRTAAASQSIIPVDTKLHKGIERIFPKFSNKF
YERPE ASRTTNCIIPVPIKLLDIAGIESGTVTTISSMNDQPVIDAYHQDLRRTAAASQSIIPVDTKLAAGITRIFPKFCDRF
Cons  .*****:*:*:** **:* **:* **:* **:* **:* **:* **:* **:* **:* **:* **:* **:* **:* **:*
GAPDHs ASRTTNCIIPV---KV---FGI---G-MTT---T---Q---D---D-R---R-A---NI-P-STGAAKAV---V-P-L-GKL

     230     240     250     260     270     280     290     300
ECOLI EAIAVRVPTINVTAIDLSVTVKPKVKANEVNLQKAAQGFHGIYDYTELPLVSVDFNHDPHSAIVDGTQTRVSGAH
SALTI EAIAVRVPTINVTAIDLSVTVKPKVKAQEVNQLQKAAQGFHGIYDYTEPLVSDFNHDPHSAIVDGTQTRVSGAH
SHEON EAIASRVPTINVTAIDLSVTLDKTVDIATVNVQVLELAANGRFNGILGYTDEPLVSCDFNHDPRSSIVDGTQTRVSAQ
SHIFL EAIASRVPTINVTAIDLSVTVKPKVKANEVNLQKAAQGFHGIYDYTELPLVSVDFNHDPHSAIVDGTQTRVSGAH
VIBCH EAIASRVPTINVTAIDLSVTVKPKVNDVNVQIVNASQCTLRGIVDYTEAPLVSIDFNHDPHSAIVDGTQTRVSNQ
YERPE EAIASRVPTINVTAIDLSVSVTHPVGVAEVNQLQKAAQGFHGIYDYTELPLVSDFNHDPHSAIVDGTQTRVSGQH
Cons  ***:*****:****:****: * ** : * . : ** . ** : **** *****:*****:***** :
GAPDHs -G---RVPT---VSVVDL-----K-----KG-L-----VS-D-----S---D-----

     310     320     330
ECOLI LIKTLVWCDNEWGFANRMLDITLAMATVAFR-----
SALTI LIKTLVWCDNEWGFANRMLDITLAMAAGVFRDLASASTKL
SHEON LVKLLLVWCDNEWGFANRMLDITLAMIAAKQS-----
SHIFL LIKTLVWCDNEWGFANRMLDITLAMATVAFR-----
VIBCH LVKMLVWCDNEWGFANRMLDITLAMQATQ-----
YERPE LIKTLVWCDNEWGFANRMLDITLAMAKSGF-----
Cons  *:* * : *****:*****:***
GAPDHs -VK---WYDNE-GY-----

```

Figure I-16 : Alignement des séquences d'E4PDH de six microorganismes et comparaison à la séquence consensus des GAPDH (Boschi-Muller et al., 1997). Les séquences d'E4PDH, déduites à partir de leur séquence nucléique, ont été extraites de la banque de données UniProt. Les deux résidus catalytiques invariants sont surlignés en rouge, les résidus impliqués dans la fixation du cofacteur via leur chaîne principale ou leur chaîne latérale dans la BstGAPDH sont surlignés respectivement en bleu et jaune. Les noms complets des organismes sont les suivants : Escherichia coli, Salmonella typhi, Shewanella oneidensis, Shigella flexneri, Vibrio cholerae et Yersinia pestis.

interne du tétramère d'E4PDH est définie par 3 axes binaires, notés *P*, *Q* et *R* par analogie avec les GAPDH, qui font correspondre à une sous-unité A choisie arbitrairement les monomères B, C et D, respectivement (Figure I-17, Moras *et al.*, 1975). La structure tridimensionnelle de l'E4PDH est très semblable à celle de la GAPDH puisqu'une superposition des carbones  $\alpha$  des tétramères de ces deux enzymes d'*E. coli* (1dc3 : apo GAPDH ; Yun *et al.*, 2000) conduit à des valeurs de rmsd de 2,8 Å (pour comparaison la superposition des structures apo des GAPDH d'*E. coli* (1dc3) et de *B. stearothermophilus* (2gd1) conduit à un rmsd de 1,2 Å).

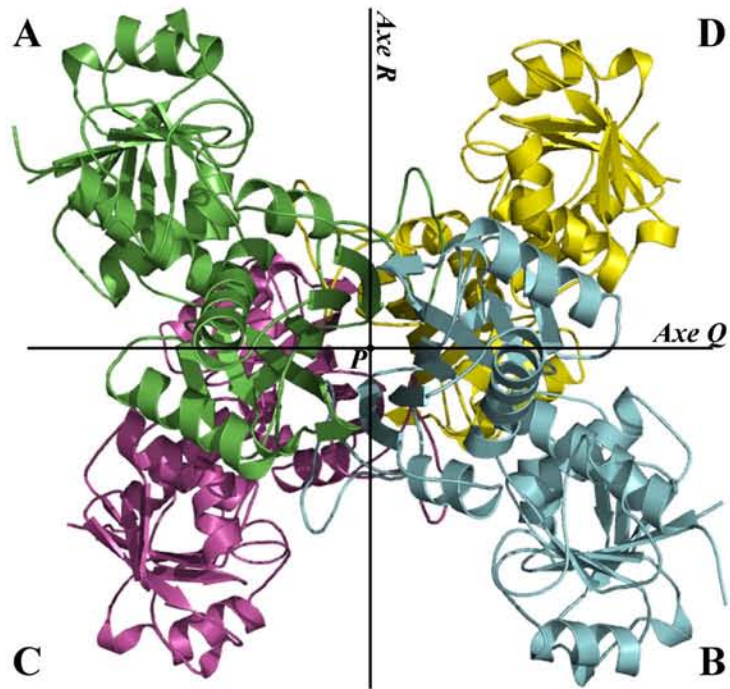
### **B.3.2.1.1. Structure des monomères**

Les monomères de l'E4PDH se composent de 338 résidus se repliant pour former deux domaines distincts, dénommés domaine de fixation du cofacteur et domaine catalytique. La superposition d'un monomère de l'E4PDH avec un monomère de la GAPDH d'*E. coli* (1dc3) est présentée en Figure I-18.

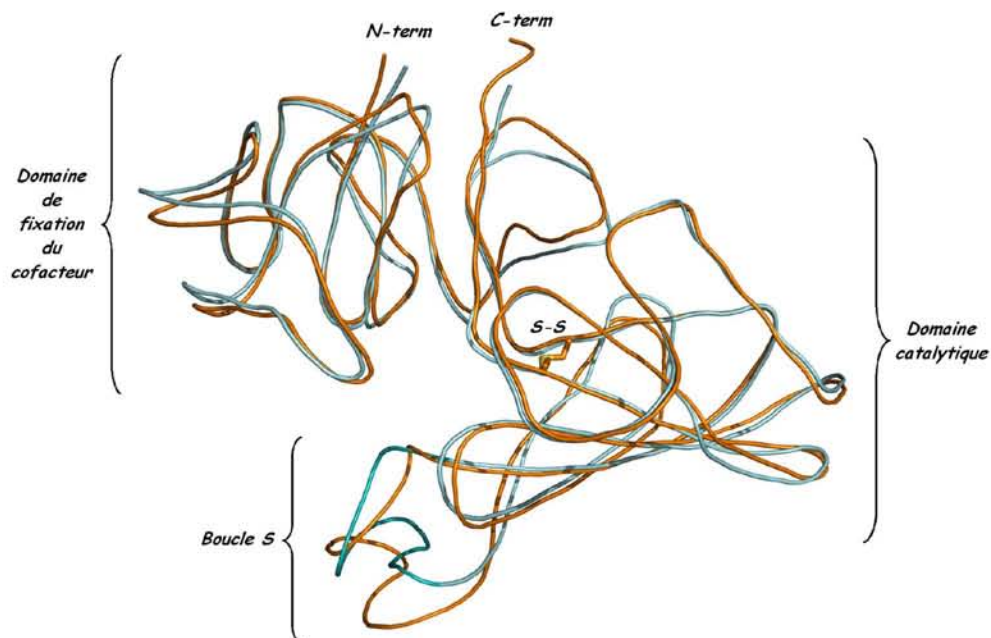
#### **Le domaine de fixation du cofacteur**

Le domaine de fixation du cofacteur est composé des 153 premiers acides aminés N-terminaux, auxquels s'associe l'hélice  $\alpha$  312-331 en C-terminal de la chaîne polypeptidique. Les 4 hélices  $\alpha$  et 9 brins  $\beta$  formés s'organisent selon un repliement caractéristique de fixation de cofacteur dit repliement de Rossmann (Rossmann *et al.*, 1974). La valeur de rmsd résultant de la superposition des carbones  $\alpha$  de ce domaine de l'E4PDH avec celui de la GAPDH d'*E. coli*, toutes deux sous forme apoenzyme, est de seulement 1,2 Å, ce qui illustre la très forte analogie structurale entre les domaines de fixation du cofacteur de ces deux enzymes. Le domaine peut être divisé en deux sous-sites dédiés à la fixation des parties adénosine monophosphate (AMP) et nicotinamide monophosphate (NMP) du cofacteur NAD.

En l'absence de données structurales, seule une modélisation du cofacteur dans la structure apoenzyme de l'E4PDH a pu être réalisée sur la base d'une superposition avec la structure holo de la GAPDH de *B. stearothermophilus* (Kauffmann, 2000). Celle-ci semble indiquer que six des huit interactions protéine-cofacteur identifiées dans la GAPDH existent potentiellement dans l'E4PDH (Tableau I-1). En outre, deux nouvelles interactions pourraient exister dans l'E4PDH, impliquant les résidus Tyr99 et His120 (respectivement Phe99 et Ala120 dans la GAPDH) (Kauffmann, 2000). Néanmoins, en l'absence d'une structure du complexe protéine-NAD, aucune conclusion ne peut être tirée concernant le mode de fixation exact du cofacteur par l'E4PDH.



**Figure I-17 : Présentation du tétramère de l'EcoE4PDH.** Par analogie avec la nomenclature des GAPDH, les sous-unités A, B, C et D (respectivement colorées en vert, bleu, mauve et jaune) sont reliées par des axes de symétrie d'ordre 2. Ces axes P, Q et R font correspondre à une sous-unité A choisie arbitrairement les sous-unités B, C et D, respectivement. (D'après Kauffmann, 2000)



**Figure I-18 : Superposition des monomères de l'E4PDH (orange ; Kauffmann, 2000) et de la GAPDH (bleu ; code pdb 1dc3 ; Yun et al., 2000) d'E. coli.**

**Tableau I-1 : Comparaison des sites de fixation du cofacteur de la GAPDH de *B. stearothermophilus* et de l'E4PDH d'*E. coli*.**

Liaisons hydrogène NAD-GAPDH dans la structure de <i>B. stearothermophilus</i> (1gd1) <sup>a</sup>		Existence potentielle de l'interaction dans l'E4PDH d' <i>E. coli</i> <sup>b</sup>	
Interactions	Distances (Å)	Interactions	Distances (Å)
OP2A-NAD --- N-Arg10	3,25	Oui	3,19
OP2N-NAD --- N-Ile11	2,79	Oui	3,03
O2'A-NAD --- O $\delta_2$ -Asp32	2,51	<sup>c</sup>	-
O3'A-NAD --- O $\delta_1$ -Asp32	2,61	<sup>c</sup>	-
N6-NAD --- O-Arg77	3,19	Oui	5,34
O4'N-NAD --- O $\gamma$ -Ser119	3,11	Oui	2,64
OP2A-NAD --- N $\delta_2$ -Asn180	2,87	<sup>d</sup>	-
O7N-NAD --- N $\delta_2$ -Asn313	2,89	Oui	2,87
<b>Liaisons hydrogène potentielles avec le NAD dans l'E4PDH d'<i>E. coli</i> <sup>b</sup></b>		O2'A-NAD --- O $\gamma_2$ -Glu32	3,17
		O2'A-NAD --- O $\gamma_2$ -Glu32	2,32
		N6-NAD --- O $\eta$ -Tyr99	3,01
		O3'N-NAD --- N $\delta_1$ -His120	2,71

<sup>a</sup> d'après Skarzynski et al., 1987. <sup>b</sup> d'après la superposition de la structure de l'apoE4PDH d'*E. coli* sur celle de l'holoGAPDH de *B. stearothermophilus*. (Kauffmann, 2000). <sup>c</sup> interactions potentiellement équivalentes (le résidu Asp32 de la GAPDH est équivalent au Glu32 de l'E4PDH. <sup>d</sup> résidus non définis dans la structure de l'E4PDH (modèle en polyalanine 178-191).

Malgré ces similitudes, l'E4PDH possède une faible affinité pour son cofacteur ( $K_D = 0,8\text{mM}$ ) comparée à celle de la GAPDH de *B. stearothermophilus* ( $K_D = 1\mu\text{M}$ ). Cette différence d'affinité reste aujourd'hui inexplicée puisque les expériences de mutagenèse dirigée, consistant à introduire dans l'E4PDH des résidus présents dans la GAPDH, n'ont pu améliorer l'affinité des mutants pour le NAD (Boschi-Muller *et al.*, 1997).

## **Le domaine catalytique**

Les acides aminés 154 à 311, constituant ce domaine, sont repliés en un feuillet  $\beta$  central à 6 brins associé à trois hélices  $\alpha$ . Le domaine catalytique de l'E4PDH est très similaire à celui de la GAPDH, la superposition des domaines des deux enzymes sous forme apo et issue toutes deux d'*E. coli* conduisant à des valeurs de rmsd de 1,4 Å (Kauffmann, 2000). Chez les GAPDH, la boucle reliant les feuillets  $\beta_1$  et  $\beta_2$  (résidus 178-201), appelée "S-loop", est localisée à l'interface des monomères et participe, en partie, au maintien de la structure tétramérique. Elle est d'autre part impliquée dans la fixation et dans la spécificité pour le NAD notamment *via* les résidus Asn180, Leu187 et Pro188 (Skarzynski *et al.*, 1987 ; Corbier *et al.*, 1990a ; Didierjean *et al.*, 1997). Dans la structure de l'E4PDH, aucune densité électronique n'a pu être observée dans cette région et la boucle S a donc, en partie, été modélisée par une chaîne de polyalanine (résidus 178-191) (Kauffmann, 2000). L'absence de densité électronique pour cette chaîne de 14 acides aminés suggère qu'elle puisse être particulièrement mobile, tout au moins dans la forme apoenzyme. Une différence notable existant par rapport à la structure de la GAPDH est la présence d'un pont disulfure entre les résidus 153 et 311 (Figure I-18). La position 311 dans les séquences de GAPDH est occupée par un résidu tyrosine strictement conservé, alors qu'elle correspond à une cystéine, elle aussi invariante, dans l'E4PDH (Figure I-16).

### ***B.3.2.1.2. Le site actif de l'enzyme***

## **Les résidus catalytiques**

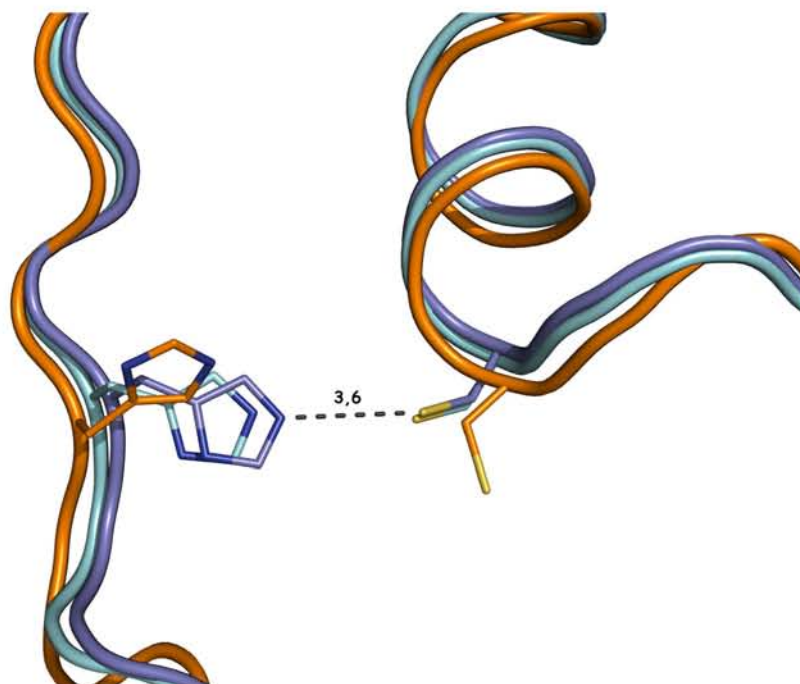
Le site actif de l'enzyme forme une large poche au fond de laquelle sont situés les acides aminés essentiels à la catalyse, Cys149 et His176 (numérotation GAPDH) ainsi que les résidus impliqués dans la reconnaissance et la fixation du substrat. Boschi-Muller *et al.* (1997) ont montré que l'His176, invariante dans les séquences d'E4PDH, jouait, comme dans le cas des GAPDH, un rôle essentiel dans le mécanisme catalytique de cette enzyme. En effet, la substitution de ce résidu

par une asparagine diminue d'un facteur 100 le  $k_{cat}$  de la réaction de l'E4PDH avec l'E4P. Pourtant, alors que dans l'ensemble des structures de GAPDH, les deux résidus catalytiques pointent l'un vers l'autre et forment une pair d'ions ((S $\gamma$ )Cys149-(N $\epsilon_2$ )His176), dans l'unique structure d'E4PDH connue à ce jour, ces deux résidus adoptent des conformations différentes, incompatibles avec la formation d'une telle interaction (les atomes S $\gamma$  et N $\epsilon_2$  sont situés à plus de 8 Å de distance) (Figure I-19).

Ainsi, dans le cas de la GAPDH de *B. stearothermophilus*, si la transition apo-holo décrite par Skarzynski and Wonacott (1988) entraîne un rapprochement de 0,4 Å entre les deux résidus catalytiques, ceux-ci sont néanmoins déjà bien orientés l'un par rapport à l'autre dans la forme apo (Figure I-19). Il semblerait donc que la formation d'un site actif compétent pour la catalyse nécessite pour l'E4PDH des réarrangements conformationnels plus importants, vraisemblablement induits par la fixation du cofacteur et/ou du substrat.

### **Le(s) site(s) de reconnaissance anionique**

Comme précédemment évoqué (*cf.* en A.3.2.2.2.), deux sites de reconnaissance anionique existent dans le site actif des GAPDH et sont nécessaires à la fixation du groupement phosphate de son substrat, le G3P, et du phosphate inorganique. Or, l'E4PDH catalyse une réaction de type non phosphorylante et ne nécessite ainsi vraisemblablement qu'un seul des deux sites pour fixer le groupement phosphate de son substrat, l'E4P. La grande majorité des résidus impliqués dans la constitution des sites anioniques chez les GAPDH est conservée dans la séquence de l'E4PDH (Ser148, Thr150, Thr208 contribuant à la formation du site Pi ainsi que le résidu Arg231 contribuant au site Ps dans les GAPDH). Deux différences notables existent cependant au niveau des résidus Thr179 et Gly209, contribuant respectivement à la formation des sites Ps et Pi dans les GAPDH, qui correspondent respectivement, dans les séquences d'E4PDH, à un résidu méthionine et un résidu lysine. Ainsi, les deux sites pouvant être affectés, il est impossible de prédire lequel des deux est susceptible d'intervenir dans le mécanisme réactionnel de l'E4PDH. De plus, étant donné que l'unique structure d'E4PDH disponible ne présente aucun anion susceptible de mimer la présence d'un phosphate au niveau du site actif de l'enzyme, les sites anioniques n'ont pas pu être clairement identifiés. Toutefois, le fait qu'une faible activité phosphorylante ait pu être détectée pour le G3P ( $k_{cat}$  600 fois inférieur à celui de la *Bst*GAPDH) a conduit Boschi-Muller *et al.* (1997) à suggérer que le site Pi, accueillant le phosphate inorganique chez les GAPDH, devait être fonctionnel.



**Figure I-19 : Superposition des résidus catalytiques de la GAPDH de *B. stearothermophilus* et de l'E4PDH d'*E. coli*.** Les structures apo et holo de la GAPDH sont représentées en bleu clair et bleu foncé respectivement (Skarzynski et al., 1987, 1gd1 ; Skarzynski and Wonacott, 1988, 2gd1) et la structure de l'E4PDH en orange (Kauffmann, 2000). La distance séparant les atomes S $\gamma$ -Cys149 et N $\epsilon_2$ -His176 de la structure holo de la GAPDH est indiquée en Å.



# *Objectifs des travaux de thèse*

Ce travail de thèse avait pour objectif l'étude structurale de deux enzymes de la famille des aldéhyde déshydrogénases, la glycéraldéhyde-3-phosphate déshydrogénase de *Bacillus stearothermophilus* (*Bst*GAPDH) et l'érythrose-4-phosphate déshydrogénase d'*Escherichia coli* (*Eco*E4PDH), dans le but d'apporter des informations concernant les relations existant entre leur structure tridimensionnelle et leur mécanisme catalytique.

Comme nous l'avons mentionné dans l'introduction, la contribution individuelle de chacun des deux sites de reconnaissance anionique présents dans le site actif des GAPDH au cours de l'acte catalytique était toujours indéterminée, et ce, malgré de nombreuses études cinétiques et structurales. La résolution de la structure de l'intermédiaire thioacylenzyme pourrait permettre de lever une partie des ambiguïtés. Pour ce faire, nous avons développé une stratégie consistant tout d'abord à cristalliser la GAPDH sous la forme holoenzyme, puis à faire diffuser le substrat naturel dans ces cristaux de façon à initier la réaction. Afin de déterminer les conditions idéales de trempage, le suivi du déroulement de la réaction dans le cristal par microspectrophotométrie paraissait indispensable. Une collaboration a donc été engagée avec l'équipe du Professeur Mozzarelli de l'Université de Parme (Italie). D'autre part, en utilisant une démarche identique à celle décrite par Didierjean *et al.* (2003), à savoir la diffusion de substrat dans des cristaux d'un mutant du site actif, le second objectif était d'obtenir la première structure d'un complexe de type michaelien entre la *Bst*GAPDH et le 1,3-DPG (produit de la réaction).

Découverte plus récemment, l'E4PDH, qui catalyse l'oxydation non phosphorylante de l'Erythrose-4-Phosphate (substrat structuralement très proche du G3P) en 4-Phosphoérythronate (4PE), est une enzyme structuralement apparentée aux GAPDH. L'unique structure disponible au début de cette étude a permis de confirmer la forte ressemblance entre les deux enzymes mais a également mis en évidence quelques différences notables (Kauffmann, 2000). Cependant, aucun ligand (anion ou cofacteur) n'étant présent dans cette première structure, aucune information n'a pu être obtenue concernant la faible affinité de cette enzyme envers son cofacteur, sa spécificité de

substrat, ou encore la façon dont le site de désacylation a évolué par rapport à celui des GAPDH. Concernant cette enzyme, les objectifs initiaux de ces travaux de thèse étaient donc :

- d'une part, de rechercher de nouvelles conditions de cristallisation pour l'E4PDH, si possible en présence d'un ligand anionique (phosphate ou analogue de phosphate) de façon à pouvoir caractériser les potentiels sites de reconnaissance anionique de l'E4PDH.
- d'autre part, de mettre au point des conditions de cristallisation de l'enzyme en présence de  $\text{NAD}^+$  et ou de substrat, dans le but d'obtenir la structure de l'enzyme sous forme de complexes binaires de type holoenzyme (enzyme- $\text{NAD}^+$ ) ou enzyme-substrat, et à l'aide d'un mutant inactif, un complexe ternaire de type enzyme/substrat/cofacteur.

# **CHAPITRE II : MATÉRIEL ET MÉTHODES**

# ***PARTIE A : PRINCIPE DE L'ÉTUDE DES PROTÉINES PAR CRISTALLOGRAPHIE DES RAYONS X***

## ***A.1. Préambule***

La cristallographie des rayons X est une méthode physique permettant de déterminer la structure des macromolécules biologiques (protéines, ADN, ARN), c'est-à-dire leur architecture tridimensionnelle. La structure de ces biomolécules étant toujours en relation étroite avec leur fonction, sa détermination permet ainsi de mieux comprendre leur fonctionnement moléculaire et leur rôle dans les êtres vivants. Cette méthode utilise le signal de diffraction des rayons X d'un monocristal pour obtenir une enveloppe détaillée de la macromolécule, sa densité électronique.

## ***A.2. Principes généraux de la cristallisation des macromolécules biologiques***

### ***A.2.1. Qu'est ce qu'un cristal ?***

Dans certaines conditions, de nombreuses molécules (dont les protéines) sont capables de passer à

l'état solide en formant des cristaux. En entrant dans l'état cristallin à partir d'une solution, les molécules individuelles adoptent toutes une ou un nombre restreint d'orientations identiques. Le cristal résultant de l'association de ces molécules est donc un empilement moléculaire ordonné tridimensionnel dans lequel les molécules sont maintenues par des interactions non covalentes. La Figure II-1 schématise une portion d'un réseau cristallin. Dans cette figure, les lignes divisent le cristal en mailles élémentaires identiques. Les points aux sommets des mailles, appelés noeuds, définissent le réseau ponctuel. La maille élémentaire est un élément de volume qui est totalement représentatif du cristal. Le cristal est caractérisé par les paramètres géométriques de sa maille élémentaire ( $a$ ,  $b$  et  $c$ , les vecteurs unités de la maille et  $\alpha$ ,  $\beta$  et  $\gamma$ , leurs angles associés) qui contient un ou plusieurs motifs équivalents selon la nature du réseau. Lorsqu'il y a plusieurs motifs, composés chacun d'une ou plusieurs macromolécules, ceux-ci sont reliés par des éléments de symétrie qui dépendent du groupe d'espace. La cristallographie des rayons X, permet l'obtention de la densité électronique moyenne de l'ensemble des  $10^{12}$  à  $10^{14}$  motifs du cristal.

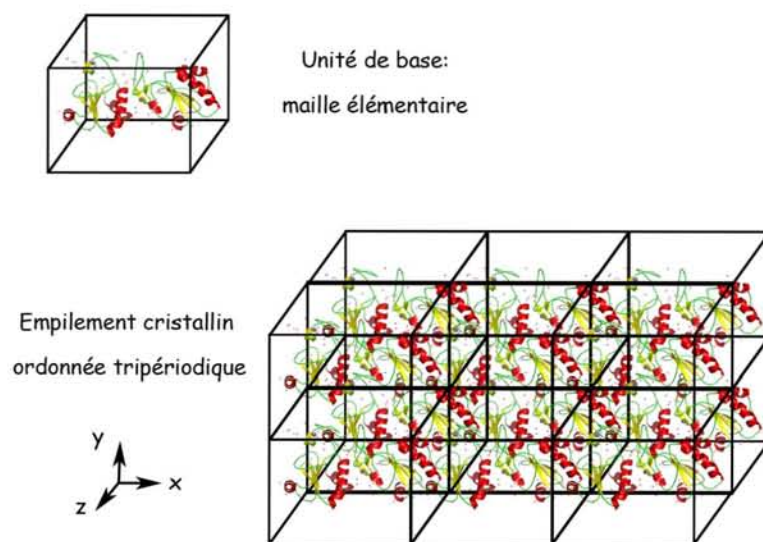
### ***A.2.2. Préparation de la protéine***

A la suite de l'étape de purification, les échantillons protéiques doivent impérativement être placés dans des conditions les plus propices possibles à leur cristallisation. De ce fait, une connaissance précise des propriétés physico-chimiques de la protéine (masse, séquence, point isoélectrique, degré de pureté, taux d'oligomérisation en solution, ...) est primordiale. La nature du tampon et des ions, le pH, ou la présence d'agent réducteur ou de ligand de la protéine sont autant de paramètres à prendre en compte dans la préparation de l'échantillon protéique.

Typiquement, la protéine est dialysée contre un tampon adéquat avant d'être concentrée par centrifugation sur membrane pour atteindre des gammes de concentrations allant de 5 à plusieurs dizaines de milligrammes de protéine par millilitre de solution.

### ***A.2.3. Cristallogenèse***

La cristallisation d'une molécule initialement en solution correspond à son passage dans une phase solide ordonnée tri périodique. Classiquement, les méthodes de cristallisation des protéines reposent sur leur précipitation contrôlée à partir d'une solution aqueuse, obtenue en les plaçant dans des conditions spécifiques dites de "sursaturation". Ces conditions ne doivent cependant pas entraîner leur dénaturation. De nombreuses substances conduisent les protéines à



**Figure II-1 : Représentation schématique d'un cristal de protéine.** La maille élémentaire, unité de base de l'empilement cristallin, est occupée dans ce cas par deux molécules de thiorédoxine dans des orientations différentes qui constituent le motif du cristal.

précipiter. Les composés ioniques, ou sels, entraînent la précipitation des protéines par un processus appelé " salting out ". Les polymères de polyéthylène glycol (PEG) sont également souvent utilisés en raison de leur pouvoir précipitant très important et de leur faible pouvoir dénaturant envers les protéines. Ces polymères sont disponibles dans une variété de masses moléculaires, tels que les PEG 4000 ou PEG 8000 par exemple, qui ont respectivement des masses moléculaires moyennes de 4000 et 8000 Da. Le principe de cristallisation le plus courant consiste à ajouter à une solution protéique une quantité d'agent précipitant juste inférieure à la concentration nécessaire pour précipiter la protéine, puis à faire lentement évaporer l'eau. Le fait que la protéine forme ensuite des cristaux ou précipite tout simplement de manière amorphe repose sur de nombreux paramètres tels que la concentration protéique, la température, le pH ou encore la force ionique. La recherche des conditions de cristallisation d'une protéine et la production de cristaux de qualité demeurent encore aujourd'hui un travail long et minutieux qui nécessite une démarche de type essais-erreurs.

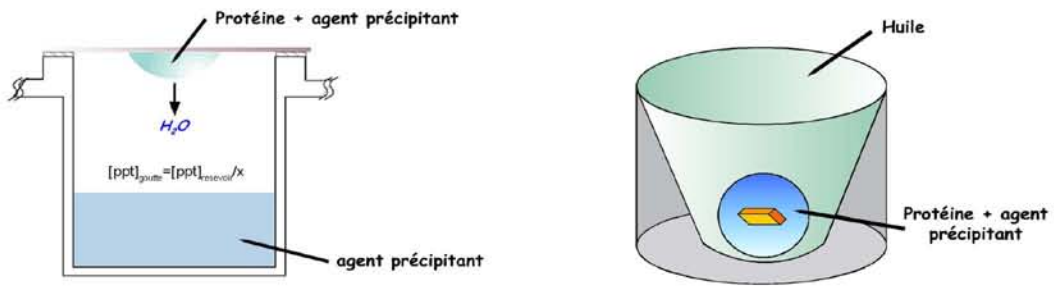
Diverses méthodes sont employées pour obtenir la cristallisation d'une protéine et seules les deux techniques les plus couramment utilisées seront présentées ci-dessous.

#### *A.2.3.1. La diffusion de vapeur*

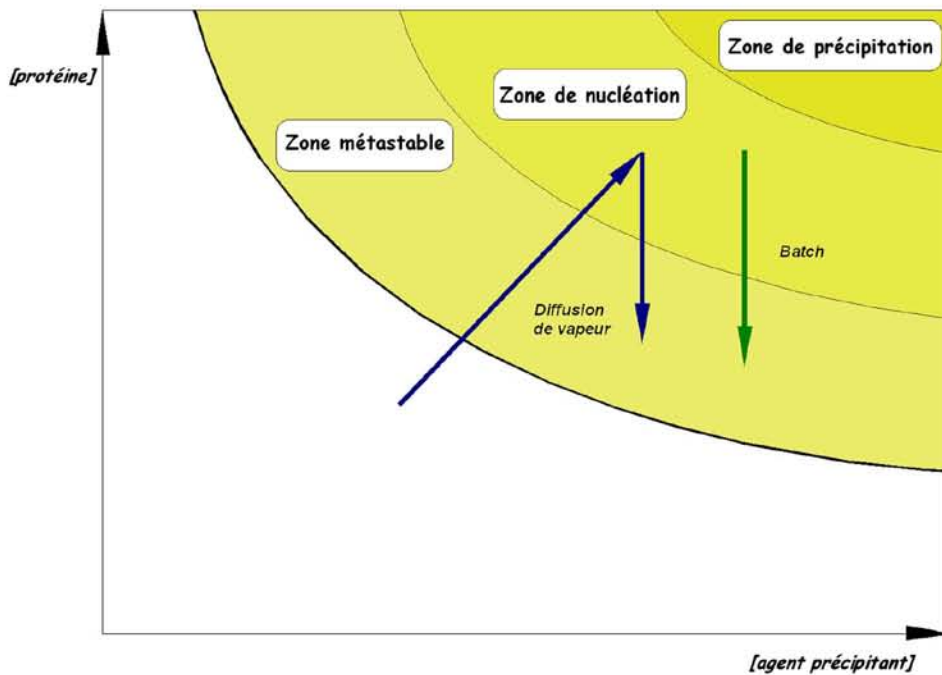
Cette méthode très répandue (McPherson *et al.*, 1995) repose sur l'équilibration, dans une enceinte close, d'une goutte de quelques microlitres composée d'un mélange protéine/agent précipitant contre un volume important (quelques centaines de microlitres) de solution aqueuse contenant l'agent précipitant à une concentration (C) supérieure à celle de la goutte de mélange (Figure II-2). Une fois le système clos, la diffusion de vapeur à partir de la goutte de mélange conduit à l'augmentation de la concentration de la protéine et de l'agent précipitant dont la concentration tend vers C. Dans des conditions favorables, la solution composant la goutte atteint la zone de nucléation du diagramme de phase (Figure II-3) où des germes cristallins peuvent se former. La formation de ces germes et la croissance des cristaux entraînent alors une diminution de la concentration en protéine (trajet bleu de la Figure II-3). La croissance des cristaux continue tant que les conditions présentes dans la goutte se situent dans la zone métastable du diagramme de phase.

#### *A.2.3.2. La méthode "batch"*

Cette méthode est la plus ancienne et la plus simple des méthodes employées pour la cristallisation des protéines (Drenth, 1994). Son principe repose sur l'ajout direct d'agent précipitant



**Figure II-2 : Représentation schématique de deux méthodes de cristallisation des protéines. (à gauche) La méthode de diffusion de vapeur en goutte suspendue avec  $[ppt]$ , la concentration de l'agent précipitant et  $X$  le facteur de dilution de l'agent précipitant dans la goutte au départ. (à droite) La méthode " batch " en goutte à l'huile.**



**Figure II-3 : Diagramme de phase théorique d'un mélange protéine-agent précipitant. La limite de solubilité de la protéine correspond au tracé en gras. Dans la partie colorée du diagramme, la protéine est dite en " sursaturation ". Les trajets bleu et vert représentent l'évolution des conditions d'une expérience de cristallisation selon la méthode de diffusion de vapeur et de batch, respectivement.**



à une solution protéique pour la placer instantanément en condition de sursaturation. Dans des conditions favorables, des cristaux vont croître à partir de cette solution dont les conditions évoluent selon le trajet vert dans le diagramme de phase (Figure II-3). La technique la plus répandue est celle de la goutte à l'huile dans laquelle le mélange protéine/agent précipitant est placé dans un puits recouvert d'une couche d'huile de type paraffine qui isole le système du milieu extérieur pour prévenir l'évaporation, la porosité de l'huile n'éliminant cependant pas totalement ce phénomène.

### *A.2.3.3. Un processus multi-paramétrique*

Les deux éléments essentiels à tout essai de cristallisation d'une protéine sont l'obtention d'un échantillon protéique d'une extrême pureté et en quantité suffisante. En effet, des préparations protéiques contenant des impuretés ne conduisent que très rarement à l'obtention de cristaux convenables. Toutefois, même le plus pur des échantillons nécessitera le plus souvent un nombre important d'essais avant d'obtenir des cristaux d'une qualité suffisante.

La cristallisation d'une protéine est un phénomène multi-paramétrique au cours duquel plusieurs paramètres sont explorés dans le but d'obtenir des conditions optimales de cristallisation. Les facteurs les plus couramment optimisés sont la concentration protéique, la nature et la concentration de l'agent précipitant, le pH et la température. Des additifs spécifiques affectant la cristallisation sont également fréquemment employés à faible concentration (cations, détergents ...).

La plupart du temps, la recherche de conditions initiales de cristallisation pour une protéine donnée repose sur le criblage des paramètres par matrice d'échantillonnage comptant plusieurs centaines de conditions de cristallisation. Ces matrices sont composées de conditions ayant déjà conduit à l'obtention de cristaux de diverses protéines et ont pour avantage d'explorer une grande variété d'agents précipitants (sels, polymères, solvants organiques), de pH, et d'additifs (Jancarik and Kim, 1991). Une fois les conditions de cristallisation initiales déterminées, celles-ci sont optimisées dans le but d'obtenir des cristaux d'une qualité et d'une taille compatibles avec les expériences de diffraction des rayons X. Diverses matrices de criblage sont disponibles commercialement dont les "Crystal screen<sup>TM</sup>" 1 et 2 (Hampton Research), "Wizard<sup>TM</sup>" 1 et 2 (Emerald Biosystem), et "Clear strategy" 1 et 2 (Molecular Dimension Limited) qui ont été utilisées au cours de cette étude. La matrice d'échantillonnage "Clear strategy" offre l'avantage de pouvoir choisir le tampon et donc le pH désiré en fonction des connaissances déjà disponibles concernant la protéine à cristalliser ou en fonction des contraintes inhérentes aux expériences que l'on désire réaliser.

Dans la plupart des cas, la détermination des conditions initiales de cristallisation ne conduit pas directement à l'obtention de cristaux d'une qualité et d'une taille convenable et un protocole d'optimisation des conditions est nécessaire. Classiquement, la concentration de l'agent précipitant, de la protéine ou encore le rapport de volume entre la solution protéique et l'agent précipitant sont optimisés par des essais dans lesquels on fait varier ces paramètres autour des valeurs déterminées par le criblage initial. Toutefois, il est parfois nécessaire d'ajouter des additifs, de changer la nature des ions présents dans la solution ou encore d'utiliser des techniques d'ensemencement pour parvenir à l'obtention de cristaux de qualité suffisante. Différentes techniques d'ensemencement (ou "seeding") existent telles que les "macroseeding", "microseeding" et "streakseeding" mais seule la technique de microseeding a été utilisée au cours de cette étude. Ces méthodes reposent toutes sur un principe commun qui consiste à apporter à un essai de cristallisation des germes cristallins, ou *nuclei*, qui pourront ensuite croître à partir de la nouvelle solution. Le "macroseeding" consiste à apporter dans un essai de cristallisation un cristal déjà formé pour le faire grossir plus encore. Le microseeding consiste, lui, à produire de tous petits germes cristallins grâce au broyage de cristaux de qualité insuffisante. Le "broyat" obtenu est ensuite dilué en série de manière à obtenir le nombre désiré de cristaux dans l'essai de cristallisation. Enfin, le "streakseeding" consiste à effleurer un cristal à l'aide d'une fibre, traditionnellement une moustache de chat, pour en arracher des germes cristallins puis à ensemercer un nouvel essai de cristallisation en réalisant une strie le long de celui-ci. Ces méthodes sont généralement très performantes pour améliorer la qualité de cristaux polycristallins et obtenir des monocristaux, pour contrôler la quantité de cristaux et donc leur taille, ou pour induire par exemple la cristallisation d'un mutant ou variant protéique « récalcitrant » à la cristallisation à partir de cristaux de la protéine native (comme par exemple les formes sélénées).

## ***A.3. Collecte et traitement des données de diffraction***

### ***A.3.1. Acquisition des données***

#### ***A.3.1.1. Le rayonnement X***

Le rayonnement X est le rayonnement électromagnétique situé entre l'ultraviolet et les rayons gamma, dont la gamme de longueurs d'onde est comprise entre 0,01 et 10 nm. L'étude par

radiocristallographie des macromolécules biologiques utilise habituellement des longueurs d'onde allant de 0.5 à 3 Å. Deux types de source de rayonnement X sont utilisées en biocristallographie, les sources de laboratoire (tube à rayons X ou anode tournante) ou les synchrotrons (Figure II-4) qui sont, eux, de grands instruments. Les études présentées dans ce manuscrit ayant été réalisées grâce "au rayonnement" synchrotron, cette présentation se limitera à ce type de source.

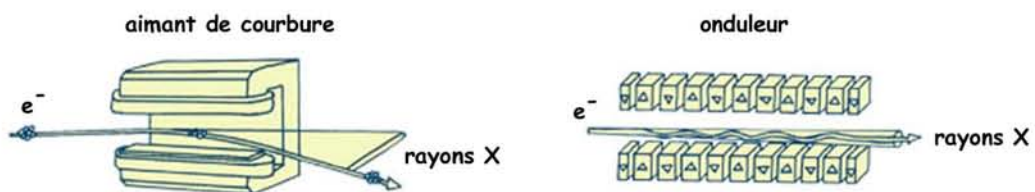
### ***A.3.1.1.1. Les Synchrotrons***

Les synchrotrons sont de grands instruments capables de conserver dans un anneau un faisceau de particules chargées à des vitesses proches de celle de la lumière. Ce faisceau produit un rayonnement électromagnétique intense utilisé pour faire les expériences. L'ESRF (European Synchrotron Radiation Facility) est un synchrotron de troisième génération construit à Grenoble au début des années 90. C'est l'un des plus grands et des plus puissants jamais construits, au même titre que l'APS (Advanced Photon Source) aux Etats-Unis ou Spring-8 au Japon. Les synchrotrons fonctionnent selon le principe suivant :

- Des électrons, émis par un canon à électrons, sont tout d'abord accélérés brusquement dans un accélérateur linéaire, ou "linac", puis transférés dans un accélérateur circulaire (booster synchrotron) où les électrons atteignent une énergie de plusieurs milliards d'électronvolts (6, 7 et 8 GeV respectivement pour l'ESRF, l'APS et Spring-8) (Figure II-4).
- Les électrons sont alors envoyés dans l'anneau de stockage, qui est une chambre à vide d'une circonférence de plusieurs centaines de mètres, où ils tournent à énergie constante pendant plusieurs heures. Dans les instruments de troisième génération, l'anneau de stockage est constitué de sections courbes et de sections droites (31 à l'ESRF). La trajectoire des électrons est définie par les champs magnétiques qu'ils traversent. On distingue plusieurs types d'aimants (Figure II-5) :
  - ✓ Les aimants de courbure, ou "bending magnets" qui dévient la trajectoire des électrons de plusieurs degrés et courbent ainsi la trajectoire des électrons. Cette déviation conduit les électrons à émettre le rayonnement électromagnétique dit synchrotron, constitué en grande partie de rayons X.
  - ✓ Situés dans les parties droites de l'anneau, les ondulateurs sont, eux, constitués d'une juxtaposition de plusieurs aimants de polarité inverse qui forcent les électrons à suivre une trajectoire ondulée. Les cônes de lumière émise à chaque courbure se superposent pour produire un faisceau de lumière beaucoup plus intense et concentré



*Figure II-4 : Vue aérienne de l'ESRF (European Synchrotron Radiation Facility) de Grenoble. L'anneau de stockage est symbolisé par un cercle mauve, les accélérateurs linéaire et circulaire, symbolisés par une flèche jaune et un anneau rouge respectivement, sont invisibles en surface.*



*Figure II-5 : Les différents aimants utilisés pour la production de lumière sur synchrotron.*

qu'à la sortie d'un aimant de courbure.

- ✓ Les aimants de focalisation, enfin, sont disposés sur les sections droites de l'anneau de stockage pour concentrer le faisceau d'électrons et le rendre le plus fin possible. De la finesse du faisceau d'électrons dépend la qualité des rayons X produits, et plus celui-ci est fin, plus les rayonnements X seront brillants.
- Les faisceaux de rayonnement synchrotron émis par les électrons se dirigent enfin vers les lignes de lumière ("beamline"), disposées tangentiellement tout autour de l'anneau de stockage, qui sont le lieu où les expériences se déroulent. Les applications utilisant "le rayonnement" synchrotron sont très variées et près de 50 lignes de lumière existent autour de l'anneau de stockage de l'ESRF. Aussi seules les lignes de lumière dédiées à la cristallographie des macromolécules biologiques (une douzaine) seront présentées ici.

Une ligne de lumière est généralement composée de trois salles successives et indépendantes nommées cabines, qui sont constituées de parois protectrices vis-à-vis des rayons X (Figure II-6).

- Juste à la sortie de l'anneau de stockage, se trouve la cabine optique qui comporte les instruments d'optique qui préparent la lumière synchrotron aux expériences à réaliser en aval. On y trouve notamment des miroirs qui permettent de focaliser les rayons X et un monochromateur qui permet de déterminer la longueur d'onde des rayons X (cet élément peut dans certains cas être réglable pour modifier la longueur d'onde en fonction des besoins de l'expérience).
- La cabine expérimentale contient le dispositif sur lequel est monté l'échantillon à analyser. En biocristallographie, il s'agit d'un diffractomètre qui autorise le centrage du cristal selon l'axe correspondant à l'axe de rotation utilisé pour la collecte des données ( $\varphi$ ). Le dispositif comprend également une canne de froid permettant de maintenir le cristal à 100K dans un flux d'azote et un détecteur (caméra CCD) pour enregistrer les clichés de diffraction.
- Une troisième cabine, dite de contrôle, permet aux expérimentateurs de piloter l'expérience et de recueillir les données.

### *A.3.1.2. Préparation des cristaux : la cryocristallographie*

L'absorption du rayonnement X par les cristaux de protéines conduit à des dégradations

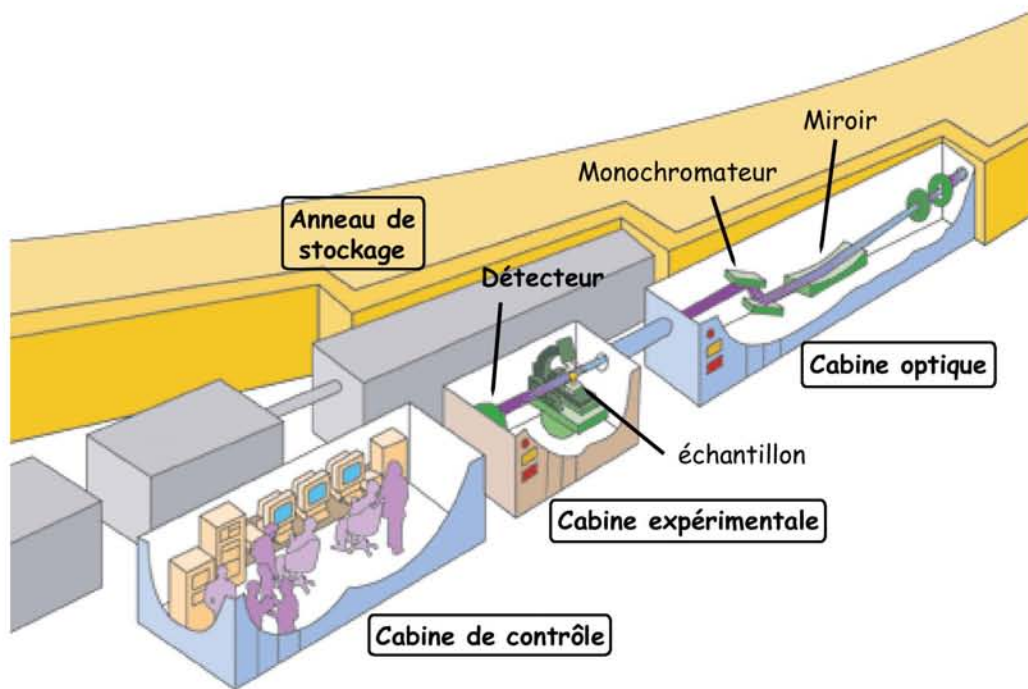
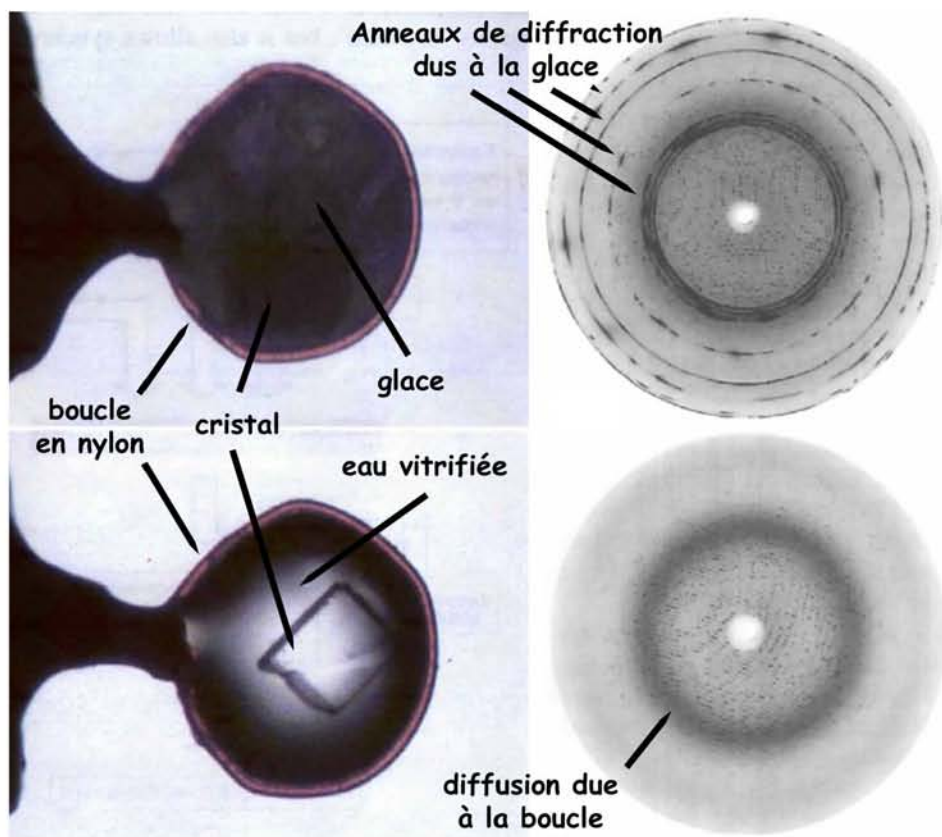


Figure II-6 : Représentation schématique d'une ligne de lumière sur synchrotron. (d'après [www.esrf.eu](http://www.esrf.eu))

principalement par des réactions radicalaires et plus marginalement par effet thermique (Burmeister, 2000 ; Ravelli and Mc Sweeney, 2000). Pour limiter la dégradation du cristal mais aussi pour améliorer la qualité des données de diffraction, les collectes s'effectuent aujourd'hui à une température de 100 K, le cristal étant maintenu dans un flux constant d'azote tout au long de l'irradiation. La cryocristallographie présente divers avantages. En effet, en plus de la protection contre les rayonnements ionisants et l'allongement de la durée de vie des cristaux, l'abaissement de la température peut augmenter l'ordre cristallin et améliorer ainsi le pouvoir de diffraction des cristaux. Ceci nécessite cependant une bonne maîtrise des conditions de cryogénie. Ainsi, les premiers essais de congélation des cristaux de protéines se révélèrent néfastes pour les cristaux, en raison de la formation de glace. En effet, les cristaux de protéines contiennent généralement entre 30 et 70 % de solvant dont le volume augmente lors de la congélation et peut endommager les cristaux. Dans les cas les moins dramatiques, la formation de multiples cristaux de glace, qui par nature diffractent aussi, apportent du bruit dans les données de diffraction. Aussi depuis une vingtaine d'années, les cristallographes ont développé des techniques de congélation rapide, ou "flash freezing", en présence d'agents cryoprotecteurs tels que le glycérol qui empêchent la formation de glace (Figure II-7). Lors de ce traitement, l'eau vitrifie, c'est-à-dire congèle dans un état amorphe. D'autres composés peuvent être utilisés en tant qu'agents cryoprotecteurs, tels que le glucose, les polymères de polyéthylène glycol de faible poids moléculaire ou le 2-méthyl-2,4-pentanediol (MPD). Typiquement, les cristaux sont placés pendant quelques secondes dans une solution dérivant de la condition de cristallisation à laquelle a été ajouté un agent cryoprotecteur, ou encore dans laquelle la concentration d'un agent précipitant (présentant un pouvoir cryoprotecteur (PEG, MPD...)) et déjà présent lors de la cristallisation) a été augmentée. Dans certains cas, l'ajout brutal d'une forte quantité d'agent cryoprotecteur peut modifier l'équilibre et altérer le cristal et il faut donc augmenter sa concentration graduellement de manière à préserver l'intégrité du cristal. Une fois protégés, les cristaux sont montés sur une boucle de nylon ou de Mylar™ puis sont plongés dans l'azote liquide (77 K) ou congelés directement sous le flux d'azote (100 K) sur le diffractomètre. Lorsque la vitrification est réussie, la solution entourant le cristal apparaît transparente (Figure II-7). Néanmoins, seul le cliché de diffraction atteste de la qualité de la vitrification.

### *A.3.1.3. Protocole de collecte*

Pour la collecte des données de diffraction, le cristal est monté sur un support constitué d'une tige métallique maintenant la boucle. Ce support est placé sur la tête goniométrique dans le



**Figure II-7: Influence de l'utilisation d'agent cryoprotecteur sur la congélation de cristaux de protéine. (à gauche) Cristaux de glycogène phosphorylase congelés dans des conditions non cryoprotectrices (haut) et cryoprotectrices (bas). (à droite) Clichés de diffraction typiques, illustrant la présence de glace (haut) ou une vitrification réussie (bas). (d'après Garman and Schneider, 1997).**



flux de rayons X. Classiquement, la démarche consiste à faire tourner le cristal selon un axe perpendiculaire au faisceau de rayons X avec un pas d'oscillation défini par l'expérimentateur (de 0,2 à 1°) en fonction des caractéristiques propres au cristal et au faisceau. A chacune des oscillations, un cliché de diffraction est enregistré par le détecteur jusqu'à la constitution d'un jeu de données composé d'une à plusieurs centaines de ces clichés, selon le pas d'oscillation choisi et les paramètres cristallins. Dans une expérience classique, la collecte de 180° d'oscillation du cristal est en général suffisante pour obtenir un jeu de données complet. Néanmoins, selon le groupe d'espace des cristaux et l'existence d'ordre supérieur de symétrie, un jeu de données complet peut être obtenu plus rapidement.

### ***A.3.2. Le traitement des données***

De nos jours, le traitement des données est réalisé à l'aide d'une suite de programmes informatiques. Trois programmes principaux existent pour l'analyse des clichés de diffraction : Denzo (Otwinowski and Minor, 1997), Mosflm (Leslie, 1999) et XDS (Kabsch, 1993), mais alors que le premier nécessite l'achat d'une licence, les deux autres sont gratuits pour les utilisateurs académiques. C'est le programme XDS (Kabsch, 1993) qui a principalement été utilisé au cours de cette étude. Néanmoins, la démarche utilisée est globalement la même pour les trois suites de programmes, avec quelques spécificités propres à chacune d'elles. Le traitement des données se décompose en deux étapes distinctes. La première étape, ou indexation, consiste, tout d'abord, à déterminer les paramètres de la maille et à choisir le système cristallin qui vont permettre au logiciel de prédire la position des taches de diffraction. Ainsi il est possible de vérifier la bonne adéquation entre les positions des taches de diffraction prédites et observées. La seconde étape, ou intégration, consiste alors à mesurer l'intensité de ces taches de diffraction, intensités qui dépendent directement du motif du cristal étudié.

#### ***A.3.2.1. Indexation : détermination des paramètres de maille***

Lors de cette étape, le programme XDS utilise un algorithme d'autoindexation développé par Kabsch (1993). Sur une série de clichés de diffraction représentant quelques degrés d'oscillation (5 à 10° dans l'idéal) les pics les plus forts, c'est-à-dire les taches de diffraction les plus intenses, sont tout d'abord localisés. Ensuite, le programme établit la liste des vecteurs qui existent entre ces pics parmi lesquels il recherche trois vecteurs réciproques non coplanaires, linéairement indépendants et de plus petits modules qui indexent, avec des valeurs entières, les pics de

diffraction. A l'issue de cette étape, le programme obtient les paramètres de la maille réduite du cristal. Les paramètres géométriques du système tels la position du faisceau direct sur le détecteur, la distance cristal-détecteur ou encore les paramètres de mailles sont alors affinés, puis le programme propose une liste des paramètres de maille calculés pour chacun des réseaux types (44 au total) associés à un facteur (indice de distorsion) illustrant l'écart entre les paramètres de maille observés et les paramètres idéaux pour un réseau donné. Le choix se porte alors sur le facteur le plus bas (qui indique la distorsion la plus faible) et le réseau ponctuel de plus haute symétrie. Pour autoriser le passage à l'étape suivante, le programme XDS exige qu'au moins 70 % des taches de diffraction aient une localisation géométriquement conforme avec les paramètres de maille.

### *A.3.2.2. Intégration des réflexions*

Le processus d'intégration consiste tout d'abord à mesurer l'intensité de chaque réflexion, puis à la corriger à l'aide de l'établissement d'un profil moyen. Cette méthode suppose que des réflexions voisines (de faible comme de forte intensité) possèdent le même profil d'intensité normalisée (Diamond, 1969). Ainsi le profil moyen établi par le logiciel au cours de la première étape à partir des profils de réflexions fortes voisines est-il appliqué à chaque réflexion pour en évaluer au mieux l'intensité. Lors d'une dernière étape, XDS applique aux intensités une série de facteurs de correction tels que les facteurs de Lorentz et de polarisation.

### *A.3.2.3. Mise à l'échelle et réduction du jeu de données*

La suite de programmes XDS comprend un programme dédié à la mise à l'échelle des données d'une part, et à leur réduction en un jeu de réflexions indépendantes d'autre part. La première étape est très importante puisqu'elle vise à corriger l'ensemble des données provenant des dizaines ou centaines de clichés de diffraction. En effet, l'absorption du rayonnement X par les cristaux tout au long de la collecte conduit à leur dégradation progressive causée par des réactions radicalaires et par un effet thermique (Burmeister, 2000 ; Ravelli and McSweeney, 2000). Alors que l'effet thermique est quasiment négligeable sur des sources de rayons X de laboratoire, il peut être plus important sur les lignes de lumière synchrotron, notamment lorsque le faisceau utilisé est particulièrement intense et focalisé. La cause majeure de dégradation des cristaux provient cependant des dommages dus aux radiations qui peuvent être décomposé en dommages dits "primaires" et "secondaires". Les dommages primaires sont inhérents à la cristallographie des rayons X et découlent de l'interaction entre le faisceau de rayon X et les molécules du cristal qui

conduit à la rupture de liaisons covalentes et à la production de radicaux libres. Les dommages secondaires sont, eux, dus aux radicaux libres produits qui diffusent à l'intérieur du cristal et réagissent aléatoirement sur les macromolécules en altérant l'ordre cristallin. La réalisation de collectes à des températures cryogéniques (100 K en général) permet de limiter les dommages dits secondaires en prévenant la diffusion des radicaux libres, néanmoins les dommages primaires causés par les rayons X sont quasi indépendants de la température et sont directement reliés à la dose de rayonnement reçu par le cristal. Par ailleurs, la collecte de données peut, en fonction du pouvoir de diffraction du cristal ou des paramètres du faisceau de rayons X, prendre de quelques minutes à plusieurs heures. Or, durant ce temps, les paramètres du rayonnement synchrotron varient en intensité. L'intensité du signal connaît, en effet, une baisse constante au fur et à mesure du temps de vie du faisceau due à des pertes d'électrons tout au long de l'anneau de stockage (phénomène de "decay"). Dans certains cas, une réinjection d'électrons dans l'anneau de stockage a lieu au cours de la collecte et provoque une augmentation brutale de l'intensité du signal, et donc de l'intensité des taches de diffraction. La forme du cristal peut également entraîner des variations de l'intensité des taches de diffraction au cours de la collecte en fonction de l'orientation du cristal, selon qu'un volume plus ou moins important est traversé par le faisceau de rayons X. La mise à l'échelle des données vise donc à corriger l'ensemble de ces phénomènes inhérents à la collecte des données de diffraction.

Une fois cette correction effectuée, le nombre de réflexions est réduit en moyennant les intensités redondantes (observation multiple d'une réflexion donnée) et équivalentes par symétrie (en tenant compte du groupe de Laue du cristal). Dans une expérience classique, les réflexions d'indices  $h, k, l$  sont considérées, de par la loi de Friedel, comme étant équivalentes aux réflexions d'indices  $-h, -k, -l$ , faisant du même coup augmenter la redondance. Cependant, dans le cas de certaines expériences, MAD et SAD notamment ("multiple-" ou "single-wavelength anomalous dispersion"), on cherche à extraire le signal anomal qui correspond à la différence existant entre les réflexions d'indices  $h, k, l$  et  $-h, -k, -l$  et il convient alors de ne pas moyennner ces réflexions.

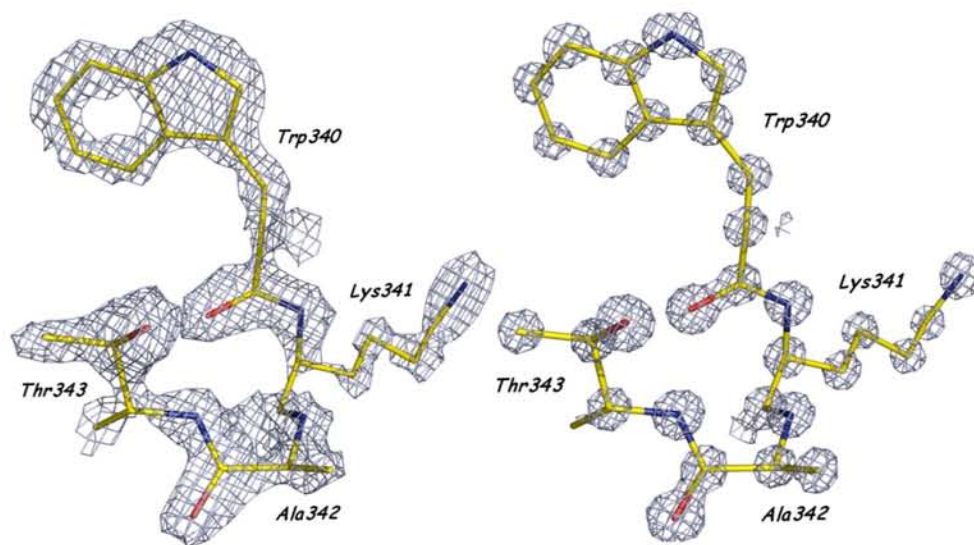
#### *A.3.2.4. Évaluation de la qualité du jeu de données*

Plusieurs facteurs, obtenus par une analyse statistique du jeu de données, permettent d'évaluer la qualité de ce dernier. Ainsi, chaque publication faisant référence à une structure cristallographique contient un tableau de résultats décrivant ces facteurs, pour l'ensemble du jeu de données, mais aussi pour la tranche de plus haute résolution.

- **La résolution limite** (en Å) est l'un des facteurs importants pour décrire un jeu de données. Elle est liée au niveau des détails visibles dans les cartes de densité électronique. Par exemple, un jeu de données avec une résolution inférieure à 2,5 Å ne permet de définir que la position de groupes d'atomes alors qu'un jeu de données avec une résolution de l'ordre de l'angstroem, permet de définir la position des atomes de manière individuelle et même de révéler la position de certains atomes d'hydrogène dans les régions non désordonnées de la structure (Figure II-8).
- **La redondance** (N) donne le nombre moyen de mesures de chaque réflexion indépendante du jeu de données. La symétrie interne du cristal influence de manière primordiale ce facteur car pour une même couverture du réseau réciproque (oscillation de 180° du cristal, par exemple), ce facteur peut passer de 2, dans le cas du groupe ponctuel 1, à 48 dans le cas du groupe ponctuel 432.
- **La complétude** (en %) représente le taux de réflexions mesurées par rapport au nombre théorique de réflexions à mesurer pour une résolution donnée.
- **Le rapport I/σ(I)** est le rapport existant entre l'intensité mesurée et l'écart type estimé de cette mesure. Couramment appelé rapport signal sur bruit, celui-ci diminue parallèlement à l'augmentation de la résolution. La limite haute de la résolution qui peut être conservée pour l'affinement est généralement définie pour une valeur de I/σ(I) supérieure ou égale à 2.
- **Le R<sub>merge</sub>** (en %) est fréquemment utilisé pour estimer la qualité d'un jeu de données de diffraction. C'est un facteur de désaccord entre les mesures de réflexions équivalentes ou symétriquement équivalentes du jeu de données. Le R<sub>merge</sub> correspond à la somme des différences entre l'intensité de la i<sup>e</sup> mesure de la réflexion indépendante  $hkl(I_i(hkl))$  et l'intensité moyenne de cette réflexion  $(\overline{I(hkl)})$  divisée par la somme de l'ensemble des intensités mesurées.

$$R_{merge} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \overline{I(hkl)}|}{\sum_{hkl} \sum_i I_i(hkl)}$$

Néanmoins, il convient d'être conscient que ce facteur est intrinsèquement dépendant de la redondance N du jeu de données. Ainsi, la valeur du R<sub>merge</sub> augmente avec le nombre de fois qu'une réflexion donnée est mesurée, même si d'un point de vue statistique la valeur



**Figure II-8 : Illustration de la notion de résolution en cristallographie des rayons X.** Une portion du modèle de la protéine DING de *Pseudomonas fluorescens* (résidus 340-343) est représentée dans sa carte de densité électronique de type 2Fo-Fc. A gauche, les données ont été volontairement coupées à une résolution limite de 2 Å alors qu'à droite, les données de ultra-haute résolution ont été conservées (0,88 Å). (cf. en Annexe 2)

moyenne est de plus en plus précise (Weiss, 2001). Weiss (2001) présente deux facteurs R indépendants de la redondance du jeu de données, le  $R_{r.i.m.}$  (pour "redundancy-independent merging factor") qui correspond au facteur  $R_{meas}$  donné par le programme SCALA (Diederichs and Karplus, 1997), et le  $R_{p.i.m.}$  (pour "precision-indicating merging R-factor"). Ce dernier facteur indique également la précision avec laquelle la mesure moyenne d'une réflexion a été réalisée. La dernière version du programme SCALA donne la valeur de ce  $R_{p.i.m.}$  alors que le programme XSCALE, associé à XDS (Kabsch, 1993), donne une valeur de  $R_{meas}$ , correspondant au  $R_{r.i.m.}$ . Les valeurs de  $R_{r.i.m.}$  et de  $R_{p.i.m.}$  sont données par les équations suivantes, dans lesquelles N désigne la redondance :

$$R_{r.i.m.} = \frac{\sum_{hkl} \sqrt{\frac{1}{(N-1)}} \sum_i |I_i(hkl) - \overline{I(hkl)}|}{\sum_{hkl} \sum_i I_i(hkl)} \quad R_{p.i.m.} = \frac{\sum_{hkl} \sqrt{\frac{N}{(N-1)}} \sum_i |I_i(hkl) - \overline{I(hkl)}|}{\sum_{hkl} \sum_i I_i(hkl)}$$

## A.4. Résolution de structures par la cristallographie

### A.4.1. Le problème de la phase

A partir de la diffraction du cristal par les rayons X, l'image moléculaire que les cristallographes cherchent à obtenir est une carte de densité électronique de l'unité asymétrique du cristal. Chaque pic de diffraction d'indice  $hkl$  est la somme de toutes les ondes diffusées par l'ensemble des atomes du cristal. Le facteur de structure  $F_{hkl}$  représente la somme de toutes les ondes diffusées par l'ensemble des atomes  $j$  de la maille et s'écrit sous la forme :

$$F(hkl) = \sum_{j=1}^N q_j \cdot f_j \left( \frac{\sin \Theta}{\lambda} \right) \cdot e^{2\pi i (hx_j + ky_j + lz_j)} \cdot e^{-B_j \frac{\sin^2 \Theta}{\lambda^2}}$$

où  $q_j$  est le taux d'occupation de l'atome  $j$ ,  $f_j$  son facteur de diffusion dépendant de l'angle de diffusion  $\theta$  et de la longueur d'onde  $\lambda$ ,  $x$ ,  $y$  et  $z$  ses coordonnées atomiques et  $B$  son facteur d'agitation thermique dérivé de  $U$ , le déplacement anisotropique moyen, selon la formule  $U = B/8\pi^2$ .

Ainsi, en un point de l'espace  $(x,y,z)$  de la maille, la densité électronique, notée  $\rho(x,y,z)$ , est reliée aux données de diffraction par la relation suivante :

$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l F_{hkl} e^{-2\pi i (hx + ky + lz)}$$

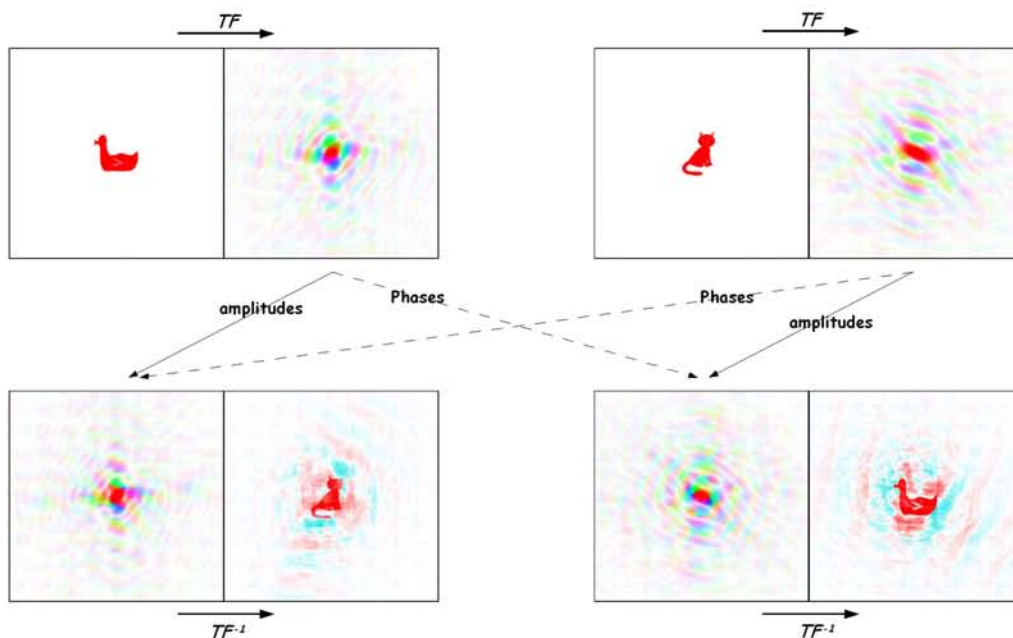
avec  $V$ , le volume de la maille. Toute onde  $F_{hkl}$  est décrite par une amplitude et une phase. L'amplitude est proportionnelle à  $(I_{hkl})^{1/2}$ , soit à la racine carrée de l'intensité mesurée pour la réflexion  $hkl$ . La phase, en revanche, n'est pas contenue dans l'information provenant des clichés de diffraction. Ainsi, la phase est la seule donnée non mesurable nécessaire au calcul de la densité électronique et donc à l'obtention d'un modèle structural de la protéine étudiée. Différentes méthodes existent pour obtenir ces phases, dont les principales seront brièvement décrites dans la suite de cette présentation.

#### *A.4.1.1. Les différentes méthodes de phasage*

Comme précédemment indiqué, l'information concernant les phases n'est pas contenue dans les clichés de diffraction et la mise au point de méthodologies visant à les obtenir a été l'un des défis les plus importants pour les cristallographes. Il faut noter également que chaque facteur de structure possède sa propre phase et donc que le problème de la phase doit être résolu pour chacune des milliers de réflexions du jeu de données. Par ailleurs, comparée à l'information apportée par les intensités, l'information contenue dans la phase est primordiale et absolument déterminante dans l'obtention d'un modèle structural "juste" de la molécule (Figure II-9).

#### **Méthode du remplacement isomorphe**

Chacun des atomes de la molécule contribue à chacune des taches de diffraction. L'ajout dans le cristal d'un faible nombre d'atomes lourds (de type Hg, Pt, Pb, Xe ou Au) dans des sites identiques dans l'ensemble des molécules de protéines du cristal a une influence seulement sur les intensités diffractées s'il n'y a pas de modification dans le réseau. La contribution d'un atome étant d'autant plus importante qu'il possède un numéro atomique élevé, l'utilisation d'atomes lourds, de haut pouvoir diffractant, facilite la détection des perturbations des intensités du cliché de diffraction. Grâce à la comparaison des jeux de données obtenus sur les cristaux de protéine native et sur les cristaux de protéine dans lesquels des métaux lourds ont été introduits par diffusion, il est possible d'extraire l'information nécessaire pour obtenir des phases initiales pour le modèle natif. Il est important de noter à nouveau que cette méthode de phasage n'est possible que si les cristaux de la protéine native et du dérivé lourd sont isomorphes. En d'autres termes, le trempage des cristaux de protéine native en présence de l'atome lourd ne doit en aucun cas modifier le réseau cristallin. L'obtention de cristaux dérivés isomorphes est une étape expérimentale souvent longue et coûteuse en cristaux.



**Figure II-9 : Illustration de l'importance de l'information contenue dans les phases.** Les transformées de Fourier (TF) de deux dessins basiques représentant respectivement un canard ou un chat sont générées. Ensuite, on croise les informations et, aux amplitudes provenant du canard, on associe les phases provenant du chat, et inversement pour obtenir des TF composites. A ces images est alors appliquée une fonction transformée de Fourier inverse ( $TF^{-1}$ ). Le résultat ne ressemble plus du tout au sujet initial mais a pris l'apparence du sujet dont proviennent les phases. Le chat apparaît comme un canard et inversement. (d'après la page Web personnelle de K. Cowtan : <http://www.yybl.york.ac.uk/~cowtan/>)



## Diffusion anormale

Une autre méthode pour obtenir une estimation initiale des phases à associer aux intensités des réflexions s'appuie sur la capacité des atomes à absorber les rayons X de longueur d'onde donnée avant réémission. Du fait de ces phénomènes d'absorption/réémission, la loi de Friedel, selon laquelle l'intensité des réflexions d'indices  $hkl$  est identique à l'intensité des réflexions  $-h-k-l$ , n'est plus respectée. La différence d'intensité entre les réflexions symétriquement liées résulte du processus de diffusion anormale.

Un élément présente des propriétés de diffuseur anormal uniquement lorsque la longueur d'onde du faisceau de rayons X est proche d'une bande d'absorption qui lui est spécifique. Les atomes de carbone, d'azote ou d'oxygène ne peuvent donner lieu à la diffusion anormale aux longueurs d'onde utilisées en biocristallographie. En revanche, certains atomes plus lourds ont des bandes d'absorption situées dans une gamme de longueurs d'onde compatibles, allant de 0.5 à 3 Å. Les collectes sont effectuées sur une ligne de lumière synchrotron de longueur d'onde variable, à une longueur d'onde spécifique de l'atome lourd présent dans la protéine. Ce ou ces atomes peuvent être les métaux endogènes (Zn, Cu, Fe, ...) naturellement présents dans les protéines, ou des atomes incorporés artificiellement à des positions fixes. Le remplacement des atomes de soufre des résidus méthionine par des atomes de sélénium, ayant un pic d'absorption théorique à 0,97974 Å, est aujourd'hui couramment utilisé du fait de la relative facilité d'obtention de dérivés sélénométhionylés d'une protéine recombinante donnée, par culture de l'organisme hôte auxotrophe pour la méthionine sur un milieu minimum supplémenté en sélénométhionine.

Deux méthodes principales existent, le MAD ou le SAD pour "Multiple-" ou "Single-wavelength Anomalous Dispersion". Dans le premier cas, 2 à 3 collectes sont réalisées à des longueurs d'onde spécifiques correspondant notamment à la bande d'absorption de l'atome lourd d'intérêt, et à une longueur d'onde éloignée. Aujourd'hui, la puissance des programmes informatiques dédiés à la résolution du problème de la phase autorise l'utilisation du seul jeu collecté au pic d'absorption de l'atome selon la méthode SAD. Comme dans la méthode précédente, la perturbation apportée par le comportement de ces quelques atomes permet de déterminer un jeu de phases à associer aux facteurs de structure.

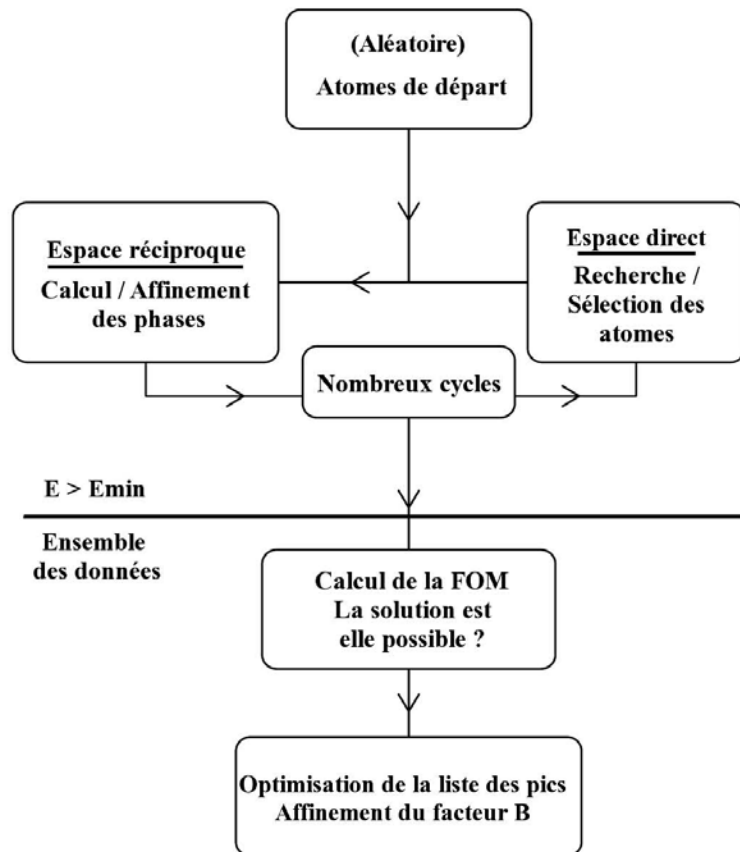
## Méthode de phasage direct

Les structures cristallographiques de petites molécules contenant jusqu'à 100 atomes sont

communément résolues en utilisant des méthodes de phasage dites directes qui ne requièrent que les seules intensités d'un jeu de données de diffraction natif. Utilisant une théorie probabiliste complexe, ces méthodes sont basées sur l'existence de relations mathématiques entre certaines combinaisons de phases. A partir de ces relations, un nombre suffisant d'estimations initiales de phase peut être obtenu pour générer un jeu complet de phases. Cette méthode n'est, à priori, pas adaptée au cas des macromolécules biologiques classiques qui contiennent généralement plus de 1000 atomes. Cependant, le développement d'une méthode d'itération utilisant à la fois l'espace réel et l'espace réciproque, appelée "Shake-and-Bake", a complètement modifié les limites d'application des méthodes de phasage direct (Weeks *et al.*, 1993 ; Miller *et al.*, 1993). Schématisée dans la Figure II-10, cette méthode débute par la génération d'un arrangement aléatoire d'atomes qui permet de générer des phases. Celles-ci entrent ensuite dans un processus itératif qui consiste alternativement à affiner ces phases dans l'espace réciproque (procédure de "shaking") puis à retourner dans l'espace direct pour "affiner" l'arrangement des atomes (procédure de "baking") qui sera alors utilisé dans un nouveau cycle. Bien que les arrangements aléatoires d'atomes générés soient limités à des cas physiquement plausibles (par exemple, ne pas avoir deux atomes à une distance inférieure à leur distance de liaison), leur nombre peut rapidement devenir trop important lorsque le nombre d'atomes augmente. Il est aujourd'hui possible grâce à ces algorithmes et à la puissance de calcul des ordinateurs de résoudre *ab initio* la structure de protéines de plus de 1000 résidus par cette méthode directe. Néanmoins, il est primordial de disposer pour cela de données de diffraction d'une qualité exceptionnelle, d'une résolution supérieure ou égale de 1,1 Å. Une telle résolution n'est toutefois atteinte que pour moins de 1% des macromolécules biologiques de structure connue. Cependant, la présence d'atomes lourds tels que S ou Cl autorise l'application de cette méthode directe à des données de diffraction allant jusqu'à 1,4 Å de résolution et la combinaison de cette méthode directe avec d'autres méthodes, comme la dispersion anormale, ou le remplacement isomorphe autorise un phasage à des résolutions de 3 Å.

## **Le remplacement moléculaire**

Grâce notamment au développement des projets de génomique structurale qui enrichissent la banque de données des structures de protéines (plus de 50 000 structures sont actuellement disponibles dans la Protein DataBank), il existe de plus en plus souvent un modèle structurel proche de la protéine dont on veut résoudre la structure. Il est alors possible d'utiliser ce modèle pour générer un jeu de phases initiales à associer aux amplitudes mesurées sur un cristal de la protéine étudiée. Initiée dès le début des années 60 (Rossmann and Blow, 1962), cette méthode de

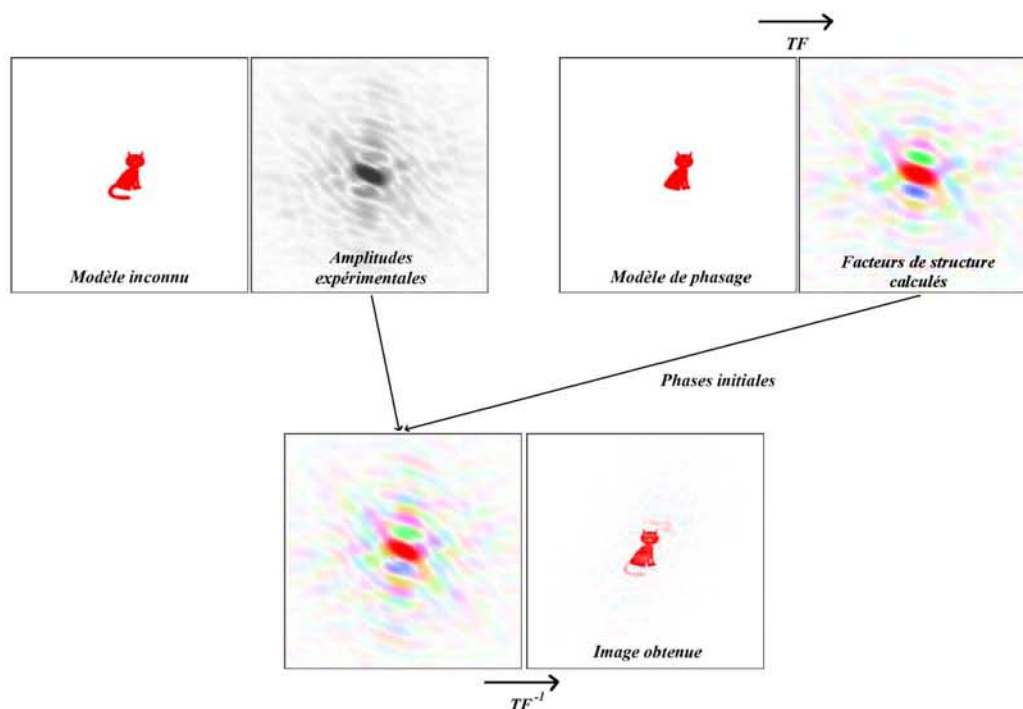


*Figure II-10 : Schéma général de la méthode de phasage direct tel que développée par le programme Shake-and-Bake. Emin est le facteur de structure normalisé minimum E utilisé pour les calculs et FOM la Figure de mérite. (D'après Uson and Sheldrick, 1999).*

phasage, appelée remplacement moléculaire, a été considérablement améliorée depuis, notamment grâce au développement des techniques de calcul numérique (programmes tels que Amore (Navazza, 1994) ou Molrep (Vagin and Teplyakov, 1997) par exemple). Le modèle supposé proche de la protéine étudiée est positionné dans l'unité asymétrique de la nouvelle protéine en deux étapes successives, qui consistent tout d'abord à orienter le modèle par la recherche d'une solution de rotation puis à trouver sa position dans l'unité asymétrique (étape de translation). Certains logiciels de remplacement moléculaire se basent sur l'utilisation de la fonction de Patterson qui correspond à la transformée de Fourier des intensités diffractées. Cette fonction présente des maxima aux coordonnées correspondant aux vecteurs interatomiques. En ces positions, la valeur de la fonction est égale au produit des densités électroniques des deux atomes. Elle est ainsi représentative de la distribution des vecteurs entre atomes d'une même molécule (vecteurs intramoléculaires) et entre atomes de molécules distinctes reliées par symétrie (vecteurs intermoléculaires). La fonction de Patterson du modèle est calculée directement à partir de sa structure atomique alors que la fonction de Patterson de la structure inconnue est calculée à partir des facteurs de structure mesurés. La recherche des solutions de rotation et de translation consiste ainsi à rechercher la corrélation maximale entre les deux fonctions de Patterson.

L'importance de l'information portée par les phases comparées à l'information provenant des intensités a été soulignée dans l'introduction de cette partie. Ainsi, lorsque le modèle est assez proche structurellement, les phases approximatives associées aux intensités donnent une image suffisamment précise de la nouvelle molécule pour résoudre la structure et même faire apparaître les différences (Figure II-11). Le modèle de phasage peut être soit une protéine appartenant à la même famille de protéines et ayant un repliement tridimensionnel similaire (à partir de 25-30 % d'identité de séquence), soit la même protéine sous un état différent structurellement très proche (présence ou absence d'un ligand, état d'oxydation différent, ...). Afin d'améliorer artificiellement la corrélation entre le modèle et la protéine d'intérêt, il est parfois nécessaire de modifier le modèle en supprimant, par exemple, les chaînes latérales pour les remplacer par des résidus alanine (modèle en polyalanine), ou en supprimant des parties plus importantes telles que des boucles ou des domaines lorsque ceux-ci sont supposés trop éloignés de la structure de la protéine inconnue.

La méthode du remplacement moléculaire ne nécessite qu'un seul jeu de données des cristaux de la protéine native, collecté à une longueur d'onde fixe.



**Figure II-11 : Illustration du concept de phasage par la méthode du remplacement moléculaire.** Il est possible de reconstruire une image à partir des seules amplitudes de Fourier dès lors qu'une image similaire est disponible pour fournir une information sur les phases. Par exemple, supposons que l'image d'un chat doive être reconstruite et que ses amplitudes soient connues. Comme l'expérience ne nous apporte que les amplitudes de la transformée, celle-ci est représentée en noir et blanc. En plus de ces amplitudes, une image que l'on sait être proche de l'image du chat à reconstruire est à notre disposition. Cette image est celle d'un chat sans queue à partir de laquelle les amplitudes et phases de Fourier peuvent être calculées. La méthode la plus simple pour reconstruire l'image du chat inconnu est de combiner les amplitudes de Fourier obtenues par l'expérience avec les phases calculées à partir de l'image du chat sans queue. Ainsi, et en dépit du fait que les phases contiennent plus d'informations que les amplitudes, l'image du chat inconnu obtenue est de bonne qualité et fait même apparaître la présence de la queue. (d'après la page Web personnelle de K. Cowtan : <http://www.ysbl.york.ac.uk/~cowtan/>)

## A.4.2. Affinement du modèle

Lorsqu'une première estimation des phases a été obtenue par l'une ou l'autre des méthodes de phasage, une carte de densité électronique initiale de la structure de la protéine peut être obtenue pour laquelle les éléments de structures secondaires (hélices et feuilletts), au minimum, sont identifiables. Un modèle préliminaire lui est associé. Les facteurs de structure calculés à partir de ce modèle initial ne sont généralement que partiellement en accord avec les facteurs de structure observés. Une mesure du désaccord existant entre les facteurs de structure calculés ( $F_{calc}$ ) à partir du modèle et les facteurs de structure observés ( $F_{obs}$ ) issus des données expérimentales est représentée par le facteur R :

$$R = \frac{\sum_{hkl} \left| |F_{obs}| - k |F_{calc}| \right|}{\sum_{hkl} |F_{obs}|}$$

A l'issue de l'étape de phasage, un facteur R de l'ordre de 50% n'est pas inhabituel. Le processus d'affinement consiste alors à améliorer le modèle de manière à trouver un meilleur accord entre les facteurs de structure calculés et les facteurs de structure observés pour atteindre des valeurs de R allant de 10 à 20 %, en fonction de la qualité des données de diffraction. L'ajustement du modèle consiste à affiner à la fois la position (les coordonnées atomiques, x, y et z), le facteur de température (facteur B) et parfois le facteur d'occupation de chacun des atomes de la molécule. En raison de leur faible interaction avec les rayons X, les atomes d'hydrogène, qui ne possèdent de fait qu'un seul électron, ne sont généralement pas modélisés dans les structures cristallographiques. Néanmoins, le nombre d'atomes autres que les atomes d'hydrogène d'une protéine reste très important et donc, le nombre de paramètres à affiner (4 au minimum par atomes) l'est plus encore. Comparé au nombre de réflexions indépendantes du jeu de données de diffraction, il n'est pas rare d'observer des rapports nombre d'observations / nombre de paramètres aussi faibles que 3, n'offrant qu'une faible surdétermination. C'est la raison pour laquelle est inclus dans le processus d'affinement un ensemble de contraintes géométriques, telles que la longueur des liaisons interatomiques, les angles de valence, issues de la connaissance de la structure de petites molécules ou de protéines à haute résolution.

En pratique, le processus d'affinement cristallographique fait intervenir, tour à tour, des phases de reconstruction manuelle du modèle à l'aide de logiciels de visualisation et des cycles d'affinements numériques des paramètres de celui-ci (coordonnées et facteur d'agitation thermique).

La reconstruction manuelle permet de modifier le modèle, de manière importante parfois, en le positionnant correctement dans les cartes de densité électronique et l'affinement numérique permet ensuite de compenser les erreurs introduites lors de la reconstruction manuelle. Au cours de cette procédure d'affinement, la "pseudo-énergie potentielle" du système ( $E_{tot}$ ) est minimisée. Cette énergie est composée de l'énergie moléculaire de la protéine ( $E_{mol}$ ) et d'une pseudo-énergie cristallographique ( $E_{cryst}$ ). Le logiciel utilisé au cours de cette étude pour l'affinement numérique du modèle est le programme Refmac5 (Murshudov *et al.*, 1997) de la suite CCP4. Ce programme utilise la méthode statistique du maximum de vraisemblance en tant que critère de convergence, dont le concept central repose sur la définition de la vraisemblance du modèle (des facteurs de structure calculés) comme étant la probabilité de réaliser la mesure des facteurs de structure observés, si l'hypothèse du modèle est juste. L'affinement se fait alors par la recherche du maximum du logarithme de la probabilité.

### ***A.4.3. Les cartes de densité électronique***

La densité électronique obtenue dessine l'enveloppe de la molécule. On peut en effet considérer que le nuage électronique entourant les atomes matérialise leur surface de van der Waals, le noyau des atomes étant modélisé au pic de densité électronique. Lors de la reconstruction graphique de la molécule, tout comme pour l'interprétation finale, le modèle est visualisé dans des cartes de densité électronique. Le contour de la molécule est matérialisé par la carte de type  $2F_{obs} - F_{calc}$ , le modèle est construit à l'intérieur de cette enveloppe en respectant des contraintes géométriques. Sont également utilisées des cartes de densité de type  $F_{obs} - F_{calc}$ , ou cartes Fourier différence. Celles-ci mettent en évidence les erreurs du modèle : les pics positifs indiquent les zones où il manque des électrons, et les pics négatifs les zones où les électrons sont en excès. Il est important de noter que les cartes de densité électronique, notamment de type  $2F_{obs} - F_{calc}$ , peuvent être en partie biaisées par l'inexactitude du modèle. Le plus souvent, on utilise donc des cartes de type  $2mF_{obs} - DF_{calc}$  (ou  $mF_{obs} - DF_{calc}$ ), où des coefficients  $m$  (figure de mérite) et  $D$  (facteur de pondération basé sur une estimation des erreurs sur les coordonnées atomique) sont attribués à chacun des facteurs de structure. Ils prennent en compte l'incomplétude et/ou l'inexactitude du modèle et permettent d'améliorer significativement la qualité des cartes de densité électronique (Read, 1986).

Afin d'étudier une région particulière de la structure (ligand, boucle mobile) en s'affranchissant totalement du biais introduit par le modèle, il est également possible de calculer des cartes de densité électronique de type "omit". En pratique, les atomes de cette région du modèle

sont éliminés, puis après quelques cycles d'affinement pour retirer un éventuel biais, des cartes sont recalculées, le reste du modèle (soit la majeure partie) apportant l'information de phase pour cette région d'intérêt. Les cartes ainsi obtenues ne présentent plus de biais dû au modèle dans cette région particulière.

#### ***A.4.4. Validation du modèle***

Afin d'évaluer et de valider la bonne progression de l'affinement ainsi que la qualité du modèle final, plusieurs critères sont à la disposition de l'expérimentateur. Le facteur R tout d'abord, déjà présenté en A.4.2., permet de juger de l'accord entre le modèle (facteurs de structure calculés) et les facteurs de structure observés. Néanmoins, les méthodes d'affinement numérique ayant pour but de minimiser au mieux ce facteur, il peut arriver que les erreurs du modèle compensent les erreurs des mesures. De plus, au cours des étapes de reconstruction manuelle du modèle, l'interprétation des cartes de densité présente une certaine part de subjectivité qui peut conduire l'expérimentateur à introduire des erreurs. C'est pourquoi l'utilisation d'un autre facteur permettant une estimation moins biaisée de la qualité du modèle a été introduite, au début des années 90, en tant que critère d'évaluation croisée de la qualité du modèle (Brünger, 1992 et 1993). Ce facteur, appelé facteur  $R_{\text{free}}$ , est calculé de la même manière que le facteur R mais à partir d'un ensemble de réflexions choisies aléatoirement (5 à 10% des données) et non incluses dans l'affinement, appelé jeu de données test. La valeur de ce  $R_{\text{free}}$  mesure ainsi la justesse avec laquelle le modèle permet de prédire les données de diffraction exclues de l'affinement. En effet, les propriétés des données de diffraction sont telles que chaque réflexion contient des informations concernant l'ensemble des atomes de la structure. C'est pourquoi des modifications apportées au modèle qui ne permettent pas d'améliorer l'accord entre ce dernier et les données de diffraction, n'améliorent pas non plus la capacité du modèle à être en accord avec les réflexions omises de l'affinement. Dans ce cas, la valeur du  $R_{\text{free}}$  reste constante, ou augmente même. Ainsi, par exemple, dans le cas d'un suraffinement du modèle lorsque le nombre d'observations n'est pas suffisant, le  $R_{\text{free}}$  aura tendance à augmenter tandis que le facteur R diminue (Kleywegt and Brünger, 1996). L'étude statistique, pour un ensemble de 357 structures cristallographiques déposées dans la PDB, de la relation existant entre ces deux facteurs, montre que l'écart moyen entre ces deux facteurs est de 7 % mais qu'il peut varier de 3 à plus de 15 % (Kleywegt and Brünger, 1996). De nombreux facteurs tels que la qualité des données, la procédure d'affinement (utilisation de contraintes géométriques plus ou moins strictes ou de contraintes de symétrie non cristallographique), la complétude du modèle ou la présence d'erreurs influent directement sur la différence entre ces deux facteurs qui dans tous les cas



doit être la plus faible possible.

D'autres critères, basés cette fois sur la géométrie, permettent également de valider un modèle. Ainsi, les déviations moyennes des longueurs et angles de liaisons entre atomes du modèle par rapport aux valeurs standards permettent d'estimer la qualité stéréochimique de celui-ci. La qualité du modèle est jugée correcte lorsque ces déviations ne sont pas supérieures aux variations physiques déjà observées dans les structures de petites molécules (environ 0.02 Å pour les longueurs et 2° pour les angles) (Engh and Huber, 1991). Le diagramme de Ramachandran, lui aussi, permet de juger de la qualité stéréochimique du modèle (Ramakrishnan and Ramachandran, 1965). En effet, en raison des contraintes d'encombrement stérique dues aux chaînes latérales des résidus d'acides aminés, les liaisons peptidiques ne peuvent adopter qu'un nombre restreint de conformations énergétiquement favorables. Ces conformations, caractérisées par les angles de torsion  $\varphi$  et  $\psi$  le long de la chaîne principale, se concentrent dans des zones préférentielles du diagramme de Ramachandran ( $\varphi$  en fonction de  $\psi$ ). Ces zones sont directement reliées aux structures secondaires (hélices  $\alpha$ , feuillettes  $\beta$ ) rencontrées dans les structures de protéines. Les chaînes latérales, elles aussi, adoptent un nombre limité de conformations préférentielles regroupées sous le terme de rotamères qui permettent de valider leur conformation dans le modèle. Enfin, les contacts atomiques comme les interactions hydrophobes ou la géométrie des liaisons hydrogène potentielles sont encore autant d'informations pour juger de la qualité du modèle cristallographique.

# ***PARTIE B : PROCÉDURES EXPÉRIMENTALES***

## ***B.1. Généralités***

### ***B.1.1. Les échantillons protéiques***

Les différentes protéines étudiées, la GAPDH de *Bacillus stearothermophilus*, l'E4PDH d'*E. coli* ainsi que leurs mutants (C149A/S de la GAPDH et C149A ou double mutant C149A/C153S de l'E4PDH), ont été produits et purifiés par l'équipe du Professeur Guy Branlant du Laboratoire de Maturation des ARN et Enzymologie Moléculaire (MAEM ; UMR 7567 CNRS-UHP). Les conditions de production et de purification de la GAPDH sont identiques à celles décrites par Boschi-Muller and Branlant (1999). Les protocoles de production et de purification de l'E4PDH d'*E. coli* sont également publiés (Boschi-Muller *et al.*, 1997). Les échantillons protéiques sont fournis sous la forme d'un précipité en présence de sulfate d'ammonium.

Préalablement aux expériences de cristallisation, les échantillons protéiques sont dialysés contre 1 L de tampon A (100 mM Tris pH 8,0, 2 mM d'EDTA, 2 mM DTT). La solution protéique est alors concentrée par centrifugation sur membrane (centricon<sup>®</sup>, 30 kDa MWCO, Millipore Corporation) et amenée à une concentration variant de 40 à 60 mg.mL<sup>-1</sup> de manière à constituer une

solution stock de protéine qui peut être diluée dans du tampon A lors des essais de cristallisation. La concentration de la protéine est déterminée par spectrophotométrie à 280 nm (spectrophotomètre Uvikon 922, Kontron Instruments) en utilisant des coefficients d'extinction molaires ( $\epsilon_{280}$ ) de  $1,31 \cdot 10^5$  et  $1,5 \cdot 10^5 \text{ M}^{-1} \cdot \text{cm}^{-1}$  et des masses molaires de 36075 et 37169 Da/monomère, respectivement pour la GAPDH et l'E4PDH. Afin d'éliminer d'éventuelles traces d'enzyme précipitée, les solutions protéiques sont systématiquement centrifugées pendant 10 min à 15 000 g à 4°C avant chaque essai de cristallisation.

Les substrats (G3P, 1,3-DPG et E4P) ont également été préparés par l'équipe du Professeur Guy Branlant au MAEM. Le *D,L*-glycéraldéhyde-3-phosphate a été préparé à partir de diéthylacétal-G3P (*D,L*-G3P; Sigma, St. Louis, MO) en se référant au protocole fourni. Sa concentration a été déterminée par dosage enzymatique avec la GAPDH. Le 1,3-DPG a été préparé à partir de G3P, grâce à un système couplé GAPDH/Lactate déshydrogénase selon la méthode décrite par Furfine et Velick (1965). Le dosage du 1,3-DPG est réalisé enzymatiquement en présence de GAPDH et de NADH en fort excès. Enfin, l'érythrose-4-phosphate, le cofacteur NAD, l'acide iodoacétique (IAA), les produits chimiques utilisés pour la préparation des différentes solutions tampons ainsi que les polymères de polyéthylène glycol utilisés pour les expériences de cristallogenèse proviennent de Sigma ou de Fluka.

### ***B.1.2. Les essais de cristallisation***

L'ensemble des travaux présentés dans ce manuscrit a débuté par la recherche de conditions de cristallisation des protéines par criblage de matrices d'échantillonnage commerciales (Crystal screen™ 1 et 2 (96 conditions), Wizard 1 et 2 (96 conditions), JBS screen (240 conditions) et Clear Strategy 1 et 2 (48 conditions)). Les criblages ont été réalisés pour différentes concentrations protéiques (10, 15, 20 mg.mL<sup>-1</sup>) en utilisant la méthode du "batch" dans laquelle 1,2 µL de solution protéique et 1,2 µL de solution de cristallisation sont déposés dans le fond des 72 puits remplis d'huile de paraffine d'une microplaque de cristallisation (Microbatch, Hampton Research). Des criblages ont également été réalisés en utilisant le robot de cristallisation du Biopôle Biologie Santé (Mosquito, TTP Labtech). Dans ce cas, la méthode de cristallisation choisie est la diffusion de vapeur en goutte assise. Les volumes de solution mis en oeuvre lors de ce type d'expériences sont considérablement réduits (de 200 à 400 nL pour la goutte de cristallisation). Les plaques de cristallisation utilisées présentent 96 puits, chacun des puits pouvant accueillir trois gouttes de cristallisation qui sont équilibrées contre la solution de cristallisation. Cet avantage a été mis à profit

pour le criblage de 3 concentrations protéiques différentes ou pour le criblage de différents rapports de volume du mélange solution protéique / solution de cristallisation composant la goutte. Une autre application a été la recherche de conditions de cocrystallisation en présence de différents ligands (cofacteur, substrat, cofacteur + substrat) ou additifs. Les essais de cristallisation ont été réalisés à une température de 20°C, et plus rarement à 4°C, et les gouttes de cristallisation régulièrement observées à la loupe binoculaire pour évaluer les résultats et détecter la présence éventuelle de cristaux. En l'absence de cristaux, l'aspect des gouttes (précipité, limpide) a été notamment pris en compte et les conditions modifiées (concentration protéique augmentée ou diminuée, concentration d'un ligand ajustée) avant de recommencer une série de criblage. Dans la majorité des cas, un processus d'optimisation a été conduit en utilisant la méthode de la diffusion de vapeur en goutte suspendue avec des boîtes Linbro classiques de 24 réservoirs. En pratique, la concentration de l'agent précipitant, le rapport du mélange composant la goutte de cristallisation, et la concentration protéique ont été optimisés sur la base des valeurs déterminées lors du criblage pour obtenir des cristaux compatibles avec la mise en oeuvre d'expériences de diffraction. Certaines conditions de cristallisation s'avérant impossibles à transposer à cette technique, leurs processus d'optimisation ont été réalisés par la méthode batch.

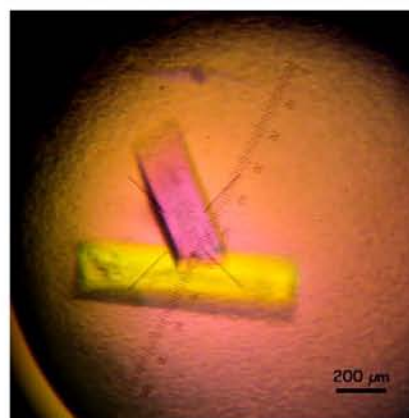
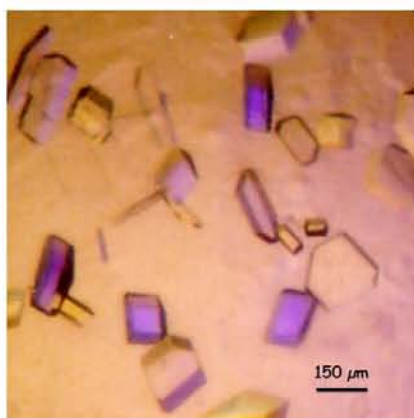
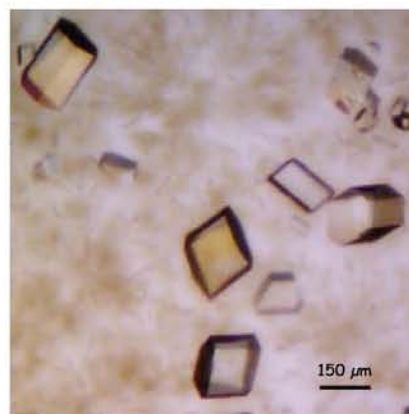
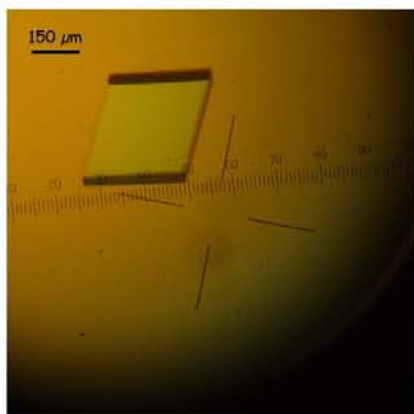
## ***B.2. Cristallisation de la GAPDH***

Les conditions de cristallisation de la GAPDH (forme holoenzyme) de *B. stearothermophilus* optimisées au cours de ces travaux sont présentées dans le Tableau II-1. Les conditions de cristallisation en présence de PEG (A1 à 3) sont dérivées de conditions de la matrice de criblage Clear Strategy (Molecular Dimension Limited) qui autorise le choix de la nature et du pH du tampon à ajouter aux 48 conditions préétablies. En effet, la recherche des conditions de cristallisation de l'enzyme a été guidée par une stratégie particulière (développée dans le Chapitre III - A.2.1.), visant à se placer dans les conditions les plus favorables, de pH notamment (réactivité suffisante de la cystéine catalytique et hydrolyse du thioacylenzyme lente), pour accumuler l'intermédiaire réactionnel thioacylenzyme dans les cristaux de GAPDH. En revanche, la condition de cristallisation de la GAPDH à base de sulfate d'ammonium avait déjà été décrite (Skarzynski et al., 1987 ; Didierjean et al., 1997) et a simplement fait l'objet d'une fine optimisation pour remplacer le tampon phosphate précédemment employé par un tampon Tris compatible avec l'isolement de l'intermédiaire thioacylenzyme (*cf.* Chapitre III - A.2.1.).

**Tableau II-1 : Cristallisation de la GAPDH de *B. stearothermophilus*. La méthode utilisée est la diffusion de vapeur en goutte suspendue, à 20°C.**

Conditions	A1	A2	A3	B1
<b>A</b> Concentration de la protéine	15 mg.mL <sup>-1</sup>	15 mg.mL <sup>-1</sup>	15 mg.mL <sup>-1</sup>	20 mg.mL <sup>-1</sup>
Concentration du ligand	2 mM NAD	2 mM NAD	2 mM NAD	2 mM NAD
<b>B</b> Conditions de cristallisation	16 % PEG 4000, 300 mM Na acétate, 100 mM Tris-HCl pH 6,9	18 % PEG 3350, 150 mM KSCN, 100 mM Imidazole pH 6,9	18 % PEG 5000 MME, 150 mM KSCN, 100 mM Imidazole pH 6,9	2,5 à 3 M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> , 100 mM Tris-HCl pH 6,9
Proportions de la goutte (A/B) en µL	1/2 ; 2/2 ; 2/1			
Temps de croissance	de 4 à 6 jours			1 à 2 semaines
Dimension des cristaux	de 0,1*0,1*0,1 à 0,4*0,4*0,3 mm			0,5*0,2*0,2 mm
Groupe d'espace	C2, parfois P21 à la suite du trempage en présence de substrat			P21

**Photographies des cristaux :**  
de haut en bas et de gauche à droite : conditions A1, A3, A2, B1



### *B.2.1.1. Trempage des cristaux*

Afin d'obtenir la structure de l'intermédiaire réactionnel thioacylenzyme de la GAPDH, les cristaux de GAPDH sous forme holoenzyme ont été trempés en présence du substrat (G3P) et la réaction (l'apparition du NADH à 340 nm) suivie par spectrophotométrie. Ces expériences de microspectrophotométrie ont été réalisées en collaboration avec messieurs les Dr. Stephano Bruno et Pr. Andrea Mozzarelli du Laboratoire de Biochimie Moléculaire et Structurale de l'Université de Parme, Italie, sur un microspectrophotomètre Zeiss MPM03 équipé d'une cellule thermostatée et d'un objectif X10 pour l'observation. Les cristaux obtenus à partir des conditions A2 et A3 (Tableau II-1) s'étant révélés plus fragiles, seuls les cristaux issus des conditions A1 et B1 ont finalement été utilisés. Pour ces expériences, les cristaux sont montés dans une cellule en quartz contenant une solution stabilisatrice dérivée de la solution de cristallisation et contenant les composés avec lesquels l'enzyme réagit. Les cristaux obtenus en présence de PEG (condition A1) ou en présence de sulfate d'ammonium (condition B1) sont tout d'abord prélevés dans la goutte de cristallisation puis lavés au moins 6 fois dans une solution stabilisatrice contenant soit 20 % de PEG 4000, 300 mM de Na acétate et 100 mM de Tris-HCl pH 6,9 ou 2,7 M de sulfate d'ammonium et 100 mM de Tris-HCl pH 8,2. Une fois le cristal monté sur le spectrophotomètre, un spectre d'absorption est enregistré entre 300 et 500 nm. La solution entourant le cristal est alors remplacée (au moins 3 rinçages) par une solution contenant le substrat et de nouveaux spectres sont enregistrés à différents temps d'incubation. Différentes concentrations du substrat ont été utilisées : 1,5 ou 10 mM de D,L-G3P pour les cristaux obtenus en présence de PEG et de 1,5 à 60 mM pour les cristaux obtenus en présence de sulfate d'ammonium. Dans le cas de l'étude des cristaux obtenus en présence de sulfate d'ammonium, un cristal unique a été utilisé pour enregistrer les spectres d'absorption en présence de concentrations croissantes du substrat. Dans tous les cas, un spectre final a été enregistré après avoir incubé le cristal en présence du substrat et d'arséniate. En effet, la réaction d'arsénolyse sur le thioacylenzyme conduit à un produit instable qui aboutit au déplacement total de la réaction vers la droite, permettant de consommer l'intégralité du NAD présent dans les cristaux et de déterminer ainsi la quantité maximale de NADH pouvant être formée. Préalablement à ces expériences, la réactivité de la cystéine a été vérifiée de façon à exclure toute possibilité d'une oxydation irréversible de celle-ci. Ces expériences d'alkylation par l'acide iodoacétique (IAA) ont été réalisées en présence de 10 mM d'IAA ajoutés à la solution stabilisatrice.

Les structures des complexes non covalents de la GAPDH avec le 1,3-DPG ont été obtenues à partir de cristaux de GAPDH mutante (C149A et C149S) cocrystallisée avec du NAD en présence

de PEG (condition A1) qui ont été trempés pendant 10 à 20 minutes en présence de 1 mM de 1,3-DPG dans une solution stabilisatrice et cryoprotectrice à 30 % de PEG, identique à celle utilisée lors de l'incubation des cristaux PEG en présence de G3P.

### ***B.3. Cristallisation de l'E4PDH***

L'enzyme qui a été utilisée au cours de cette étude est l'E4PDH d'*E. coli* de type sauvage (WT) ainsi qu'un double mutant (C149A-C153S). Des conditions de cristallisation et de cocrystallisation en présence de différents ligands ont été recherchées pour chacune de ces deux formes afin d'obtenir la structure de l'enzyme seule (WT, double mutant) ou en complexe avec son cofacteur (WT, mutant) et ou avec son substrat (double mutant). Les conditions ayant conduit à l'obtention de cristaux de l'E4PDH sous forme apoenzyme, en cocrystallisation avec le cofacteur ou le substrat ou en complexe covalent avec l'inhibiteur 3-CAPAD (3-(chloroacétyl)pyridine adénine dinucléotide), analogue du NAD, sont listées dans le Tableau II-2.

Les expériences de cristallisation de l'E4PDH se sont généralement avérées difficiles non pas à cause de l'absence de résultats positifs, mais au contraire à cause de la propension toute particulière de cette enzyme à cristalliser dans un grand nombre de conditions. Par exemple, le criblage initial de l'enzyme de type sauvage contre le screen Wizard (Emerald Bioscience) regroupant 96 conditions a permis d'identifier 6 conditions donnant des premiers cristaux (Moniot, 2004). Malheureusement, les cristaux obtenus, tous avec une morphologie similaire en faisceaux de fines baguettes, étaient d'une qualité insuffisante pour les expériences de diffraction, et l'optimisation des conditions relativement peu efficace pour améliorer leur qualité. L'optimisation de la condition II-8 contenant des ions phosphate potentiellement intéressants pour l'étude des sites de reconnaissance anioniques de l'enzyme à néanmoins permis d'obtenir de manière reproductible des cristaux d'une taille de 0,5\*0,05\*0,05 mm, compatible avec la collecte de données de diffraction.

Le criblage des conditions de cocrystallisation de l'enzyme (WT, double mutant) en présence de cofacteur (oxydé/réduit) et/ou de substrat (E4P) n'a pu permettre d'identifier des conditions de cristallisation d'un complexe entre l'enzyme et un ligand. En effet, quelques conditions, comme la condition 4D4 (Tableau II-2), ont conduit à l'obtention de cristaux à la morphologie type de fines baguettes appartenant au même groupe d'espace ( $P2_12_12_1$ ) que ceux obtenus précédemment en présence de phosphate. Dans tous les cas, les structures ne présentaient aucun signal de densité pouvant correspondre à la présence d'un ligand. Le criblage, effectué en parallèle sur la présence

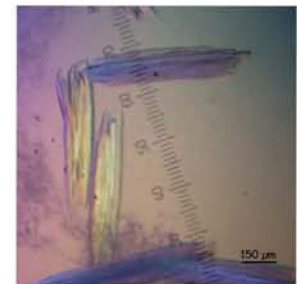
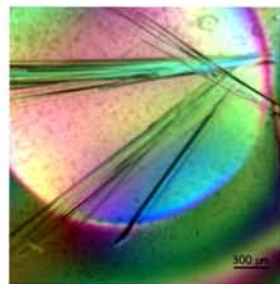
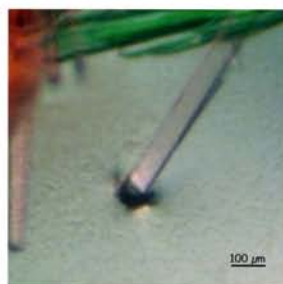
**Tableau II-2 : Cristallisation de l'E4PDH d'E. coli. Les expériences ont été réalisées à une température de 20°C.**

Conditions	II-8 (Wizard)	4D4 (JBS)	II-32 (Wizard)	CS1-1 (Clear strategy)
<b>Méthode de cristallisation</b>	diffusion de vapeur, goutte suspendue	batch	diffusion de vapeur, goutte suspendue	diffusion de vapeur, goutte suspendue, micro ensemencement
<b>Enzyme</b>	type sauvage	C149A	C149A/C153S	type sauvage
<b>Concentration de la protéine</b>	60 mg.ml <sup>-1</sup>	15 mg.mL <sup>-1</sup>	20 mg.mL <sup>-1</sup>	11,3 mg.mL <sup>-1</sup>
<b>A Concentration du ligand</b>	/	2 mM NAD	3 mM E4P	3-CAPAD*
<b>B Conditions de cristallisation</b>	10 % PEG 8000, 0,1 M Na/K phosphate pH 6,2, 0,2 M NaCl	10 % PEG 8000, 0,2 M Acétate de Mg	24 % PEG 1000, 0,1 M Tris-HCl pH 8,5	20 % PEG 2000 MME, 0,1 M Tris-HCl pH 8,5, 0,2 M Li <sub>2</sub> SO <sub>4</sub>
<b>Proportions de la goutte (A/B) en µL</b>	1/2 ; 2/2 ; 2/1	1/1	1/2 ; 2/2 ; 2/1	1/2 ; 2/2 ; 2/1
<b>Temps de croissance</b>	1 semaine	2 semaines	1 semaine	1 à 2 semaines
<b>Dimension des cristaux (mm)</b>	0,5*0,05*0,05	0,4*0,03*0,03	0,2*0,2*0,2	0,2*0,02*0,02
<b>Cryoprotection</b>	+ 20 % glycérol	**	+ 20 % glycérol (3 heures)	**
<b>Groupe d'espace</b>	<i>P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub></i>	<i>P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub></i>	<i>P4<sub>1</sub>2<sub>1</sub>2</i>	<i>P1</i>

\* Le 3-CAPAD est lié à l'enzyme par une liaison covalente préalablement à la cristallisation (cf. partie B.3.)

\*\* condition de la goutte cryoprotectrice

**Photographies des cristaux**  
de haut en bas et de gauche à droite :  
conditions II-8, 4D4 et CS1-1





des différents ligands (E4P ; NAD ; E4P/NAD) grâce au robot de cristallisation (Mosquito, Labtech), a toutefois conduit à l'obtention de cristaux d'une nouvelle morphologie, pyramidaux, se développant dans la condition II-32 (Wizard, Emerald Bioscience) uniquement en présence d'E4P. Cette condition a alors été optimisée jusqu'à obtenir des cristaux d'une dimension moyenne de 0,2\*0,2\*0,2 mm (Tableau II-2).

Enfin, la recherche des conditions de cristallisation de l'E4PDH en complexe covalent avec l'inhibiteur 3-CAPAD a été réalisée avec la matrice Clear Strategy 1 (Molecular Dimension Limited). Préalablement, l'enzyme a été incubée en présence de 610  $\mu$ M d'inhibiteur 3-CAPAD en tampon Tris-HCl 50 mM, EDTA 2mM, pH 8,0 de façon à conduire à une inactivation totale. A nouveau, le nombre de conditions entraînant la cristallisation de l'enzyme est important puisque des cristaux sont apparus dans 7 des 24 conditions composant le "screen". Si les conditions 1 et 2 ont toutes les deux fait l'objet d'une démarche d'optimisation, la condition 1 contenant des ions sulfate, a finalement été préférée pour résoudre cette structure. La morphologie des cristaux obtenus correspond à nouveau à des faisceaux de baguettes voire même à des polycristaux (*cf.* Tableau II-2). Aussi des techniques de micro ensemencement ont-elles été utilisées pour améliorer la qualité des cristaux. Pour ce faire, des polycristaux ont été collectés puis broyés dans une solution de cristallisation contenant une concentration accrue de PEG (30 % m/v). La solution ainsi obtenue a été diluée en cascade jusqu'à un facteur de dilution de  $10^5$ . Les différentes solutions de germes cristallins ont enfin été utilisées pour ensemercer (0,2  $\mu$ L) des gouttes de cristallisation prééquilibrées pendant 24h contre le réservoir. Des monocristaux de qualité acceptable ont ainsi été obtenus et soumis aux rayons X.

## ***B.4. Collecte des données de diffraction***

Les collectes des données de diffraction effectuées au cours de cette étude ont toutes été réalisées sur les lignes de lumière synchrotron X11 et X12 à Hambourg (EMBL, DESY) ou BM30A (FIP) et ID14eh1-4 à Grenoble (ESRF). Les mesures ont été réalisées à 100 K et ont donc nécessité la mise au point de conditions cryoprotectrices pour la congélation des cristaux. Lorsque la condition de cristallisation n'était pas elle-même cryoprotectrice, du glycérol a été ajouté à la condition jusqu'à obtenir une congélation correcte. Les conditions cryoprotectrices utilisées pour les cristaux de GAPDH seront présentées en partie Résultats. La condition cryoprotectrice utilisée pour chacun des cristaux d'E4PDH est indiquée dans le Tableau II-2. La mise au point de la condition cryoprotectrice des cristaux d'E4PDH cocrystallisés en présence d'E4P a été particulièrement

délicate. En effet, malgré la taille et l'aspect encourageant des cristaux de forme pyramidale, les premiers clichés de diffraction obtenus à 100 K présentaient une très faible résolution allant de 6 à 9 Å. Des tests, réalisés à température ambiante, ont ensuite permis de confirmer que les cristaux avaient un réel pouvoir de diffraction mais que le protocole de congélation utilisé était néfaste pour l'ordre cristallin. A la suite de différents essais, la solution protectrice initiale, composée de 24 % de PEG 1000, 100 mM de tampon Tris-HCl pH 8,5 et de 20 % de glycérol a été conservée mais le protocole de congélation modifié. En effet, un trempage de 3 heures minimum dans cette solution protectrice avant congélation est nécessaire pour que ces cristaux conservent leur pouvoir de diffraction lors de la congélation à 100 K.

Les données de diffraction collectées ont été indexées et intégrées en utilisant le programme XDS (Kabsch, 1993). La mise à l'échelle des jeux de données a été réalisée avec les programmes XSCALE (Kabsch, 1993) ou Scala (Evans, 1993) de la suite de programme CCP4 (COLLABORATIVE COMPUTATIONAL PROJECT, Number 4). Les tableaux statistiques des différents jeux de données sont présentés dans la partie Résultats.

## ***B.5. Phasage et affinement des structures cristallographiques.***

Les structures présentées dans ce manuscrit ont été résolues en utilisant la méthode du remplacement moléculaire (programme Molrep ; Vagin and Teplyakov, 1997). Les modèles choisis pour le phasage des structures de GAPDH sont la structure holoenzyme de la GAPDH de *B. stearothermophilus* (Skarzynski *et al.*, 1987) dans le cas des structures de forme monoclinique *P21* ou la structure holoenzyme de la GAPDH de *B. stearothermophilus* résolue par Didierjean *et al.* (2003) dans le cas de la structure de forme monoclinique *C2* (code pdb 1gd1 et 1nqo, respectivement) (Tableau II-1). Le modèle de phasage utilisé pour les structures d'E4PDH est la forme apoenzyme de l'enzyme d'*E. coli* résolue au laboratoire par Kauffmann (2000). Du fait de l'utilisation de modèles de phasage très proches des structures à résoudre, l'utilisation des données allant de la basse résolution jusqu'à 3,5 Å et des paramètres choisis par défaut par le programme Molrep a généralement permis d'obtenir facilement une première estimation des phases et des modèles structuraux de bonne qualité. Dans tous les cas, l'existence d'une symétrie interne 222 au sein des protéines étudiées conduit à l'obtention de 4 solutions possibles pour le remplacement moléculaire. Lorsque plusieurs molécules étaient présentes dans l'unité asymétrique (cas des structures de la GAPDH en complexe avec le G3P ou le 1,3-DPG et de l'E4PDH en complexe avec

le 3-CAPAD, qui présentent 4 tétramères indépendants dans leur unité asymétrique), deux étapes de remplacement moléculaire ont parfois été nécessaires : la première étape a conduit au positionnement de 2 à 3 des tétramères de l'unité asymétrique qui ont ensuite été repris dans une seconde étape en tant que molécules fixes lors du positionnement du ou des tétramères manquants. Les détails du remplacement moléculaire seront repris pour chacune des structures dans la partie Résultats.

Les modèles obtenus au cours de ces travaux ont été affinés contre les données cristallographiques avec le programme Refmac5 (Murshudov *et al.*, 1997). Plusieurs cycles d'affinement en corps rigide ont été tout d'abord réalisés à une résolution moyenne (3 à 3,5 Å) avec les tétramères puis avec les monomères indépendants considérés comme éléments rigides. La position et le facteur d'agitation thermique de chacun des atomes ont ensuite été affinés individuellement en incluant peu à peu les couronnes de haute résolution. Entre chaque procédure d'affinement, des cartes de densité 2Fo-Fc et Fo-Fc (cartes Fourier différences) de type sigma-A ont été calculées et visualisées à l'aide du logiciel Coot (Emsley and Cowtan, 2004). Ces cartes de densité ont été utilisées pour reconstruire et corriger manuellement le modèle. En fin d'affinement, les molécules d'eau ont été ajoutées au modèle en vérifiant l'existence d'accepteurs/donneurs de liaison hydrogène et les éventuels ligands présents modélisés. Le processus d'affinement a été conduit jusqu'à convergence des facteurs R et R<sub>free</sub> dont les valeurs seront reprises pour chaque structure dans la partie Résultats. La structure de l'intermédiaire thioacylenzyme de la GAPDH de *B. stearothermophilus* a été soumise à un affinement avec le logiciel Buster-TNT (Bricogne and Irwin, 1996), affinement réalisé par Clemens Vonrhein (Global Phasing Limited). Cet affinement a permis d'améliorer nettement la qualité du modèle et a surtout permis l'affinement du taux d'occupation des ligands présents en occupation partielle dans le modèle (*cf.* Résultats). De la même manière, l'affinement par le logiciel Buster-TNT des structures d'un complexe entre la GAPDH et le 1,3-DPG est actuellement en cours.

Préalablement au travail d'interprétation, la qualité stéréochimique des modèles a été vérifiée en utilisant le serveur internet Molprobity (Lovell *et al.*, 2003) et le logiciel Procheck (Laskowski *et al.*, 1993).

La structure de l'intermédiaire thioacylenzyme de la GAPDH de *B. stearothermophilus* a été déposée dans la Protein Data Bank, avec les facteurs de structure, sous le code 3cmc, et a été publiée dans J. Biol. Chem (*cf.* Annexe 1).

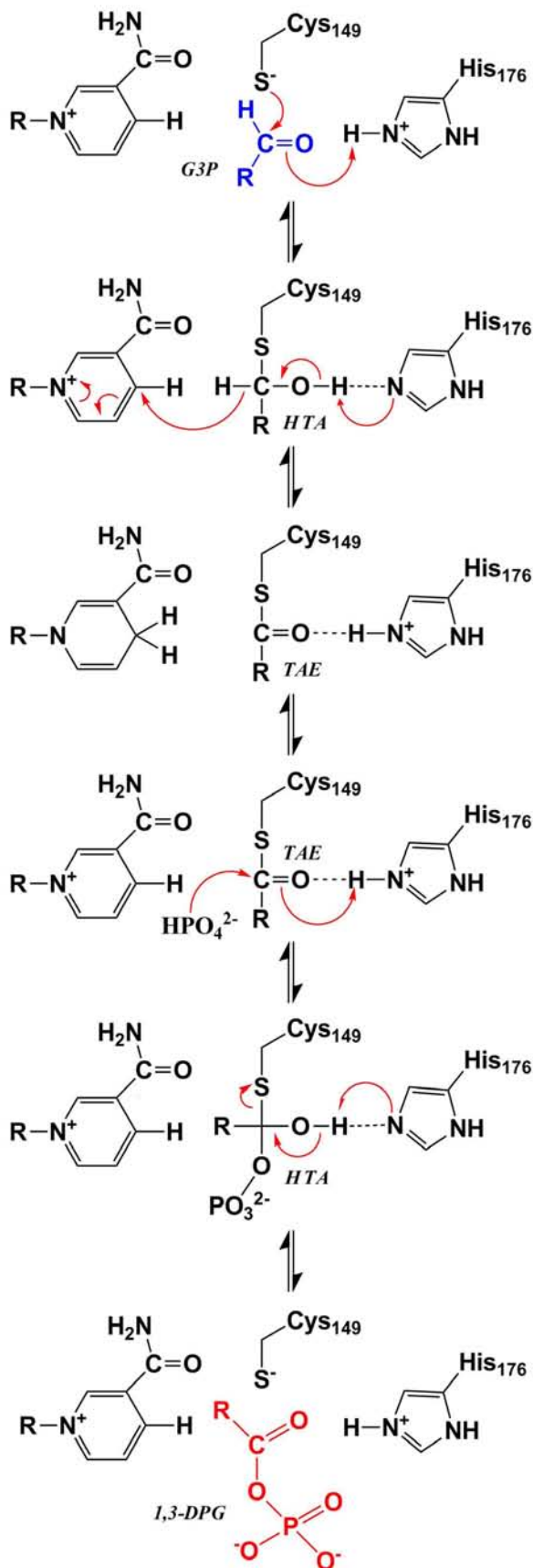


CHAPITRE III : CONTRIBUTION À  
L'ÉTUDE DU MÉCANISME RÉACTIONNEL  
DE LA GLYCÉRALDÉHYDE-3-  
PHOSPHATE DÉSHYDROGÉNASE

# *PARTIE A : RÉSULTATS*

## *A.1. Contexte et principaux résultats*

La glycéraldéhyde-3-phosphate déshydrogénase, de par son rôle clé dans la voie de production d'énergie, est une enzyme très bien caractérisée. Son mécanisme catalytique (Figure III-1), étudié en détail chez les bactéries et eucaryotes (Segal and Boyer, 1953 ; Trentham, 1971 ; Moras *et al.*, 1975 ; Harris and Waters, 1976 ; Soukri *et al.*, 1989 ; Michels *et al.*, 1996 ; Talfournier *et al.*, 1998 ; Boschi-Muller and Branlant, 1999), procède en deux étapes. La première étape consiste en une réaction d'oxydoréduction qui conduit *via* un transfert d'hydrure à la formation d'un intermédiaire réactionnel covalent appelé thioacylenzyme. Ensuite intervient la seconde étape de la réaction qui, au travers de la phosphorylation de l'intermédiaire, mène à la libération du produit de la réaction (1,3-DPG). La phosphorylation est précédée d'une étape d'échange du cofacteur (Trentham, 1968 ; Harrigan and Trentham, 1974). Deux sites de reconnaissance anionique, Ps et Pi, sont nécessaires pour accueillir respectivement le groupement phosphate du substrat et la molécule de phosphate inorganique dans le site actif de l'enzyme. Cependant, bien qu'un rôle propre leur ait été attribué, définissant le site Ps et le site Pi comme étant respectivement dédié à la reconnaissance du phosphate du substrat (C3P) et à la fixation du phosphate inorganique, leur véritable implication dans le mécanisme catalytique semble plus complexe et reste encore aujourd'hui matière à débat (*cf.*



La Cys149, stabilisée sous forme thiolate grâce à la paire d'ion formée avec l'His176, attaque le centre électrophile du G3P : la formation de l'hémithioacétal est assistée par l'His176, protonée, jouant le rôle de catalyseur acide.

L'hémithioacétal est stabilisé par liaison hydrogène avec l'His176. Jouant ensuite le rôle d'une base, l'His176 assiste la réaction d'oxydoréduction, c'est-à-dire le transfert d'hydrure de l'hémithioacétal au pyridinium qui est réduit en NADH

L'acylenzyme est stabilisé par liaison hydrogène avec l'His176. Etape d'isomérisation de l'acylenzyme : remplacement du NADH par le NAD<sup>+</sup>

Attaque nucléophile d'un phosphate inorganique (phosphorolyse), assistée par l'His176, jouant le rôle d'un catalyseur acide.

Relargage du 1,3-DPG assisté par l'His176 (catalyseur basique)

**Figure III-1 : Représentation schématique du mécanisme réactionnel postulé des GAPDH phosphorylantes. (G3P : glycéraldéhyde-3-phosphate ; HTA : hémithioacétal ; TAE : thioacylzyme ; 1,3-DPG : 1,3-diphosphoglycérate)**

Chapitre I-A). C'est dans ce cadre que l'équipe de Biocristallographie du LCM3B a débuté, en collaboration avec l'équipe du Professeur Branlant (MAEM, UMR CNRS-UHP 7567), des travaux visant à résoudre les structures de la GAPDH en complexe avec son substrat à différents stades de la réaction, dans le but d'obtenir des réponses claires quant à la localisation du substrat et en particulier de son groupement phosphate au cours de l'acte catalytique. La première structure, représentative d'un complexe de type michaelien de la GAPDH de *B. stearothermophilus* (*Bst*GAPDH) a été publiée en 2003 (Didierjean *et al.*, 2003) et confirme le positionnement initial du phosphate du substrat au niveau du site Ps de l'enzyme. Cependant, cette structure ne permet pas de prévoir la localisation de ce groupement phosphate (C3P) au cours des étapes consécutives. C'est ainsi qu'à mon arrivée au sein de l'équipe en 2004, il m'a été confié l'obtention de la structure de l'intermédiaire thioacylenzyme et de la structure d'un complexe enzyme / 1,3-DPG dans le but de préciser les rôles respectifs des deux sites de reconnaissance anionique de l'enzyme au cours de la catalyse.

Obtenue par trempage des cristaux de la forme holoenzyme de la *Bst*GAPDH de type sauvage en présence de G3P, la structure de l'intermédiaire thioacylenzyme résolue à 1,8 Å a récemment été publiée (Moniot *et al.*, 2008). Nos travaux se sont notamment appuyés sur une démarche d'optimisation des conditions d'accumulation de l'intermédiaire réactionnel dans les cristaux grâce à des expériences de microspectrophotométrie. La structure montre une image claire du positionnement du groupement phosphate du substrat (C3P) au niveau du site Pi de l'enzyme et suggère donc fortement, par comparaison au complexe michaelien (Didierjean *et al.*, 2003), un repositionnement du substrat concomitant ou postérieur à l'étape d'oxydoréduction. La conformation du substrat, notamment au niveau de son atome O1, soulève par ailleurs certaines interrogations par rapport au mécanisme et aux conséquences structurales de l'étape d'échange de la molécule de cofacteur.

Deux structures de complexes enzyme-1,3DPG, à 1,66 et 2,14 Å de résolution ont, par ailleurs, été obtenues en utilisant des mutants inactifs ou peu actifs de l'enzyme (C149A et C149S) et révèlent, quant à elles, que le groupement phosphate C3P du substrat est localisé dans le site Ps tandis que le second groupement phosphate, noté C1P, est, lui, positionné dans le site Pi. Ces résultats confirment que l'attaque nucléophile du phosphate inorganique sur le thioacylenzyme est réalisée à partir du site Pi, mais aussi que le groupement C3P opère un mouvement complet de "va-et-vient" entre les deux sites de reconnaissance anionique au cours de la catalyse.

Enfin, la structure à 1,77 Å de résolution d'un intermédiaire covalent de type acylenzyme a été obtenue, *via* le sens inverse de la réaction, grâce la diffusion de 1,3-DPG dans des cristaux de la



forme holoenzyme du mutant C149S de la *Bst*GAPDH (qui présente une faible activité résiduelle). Cet intermédiaire présente la particularité de faire coexister un intermédiaire covalent avec le cofacteur sous forme oxydée. En effet, dans le cas de la première structure, l'intermédiaire thioacylenzyme avait été piégé dans le sens « normal » de la réaction (à partir du G3P) à l'issue de l'étape d'oxydoréduction (formation du NADH). Les informations issues de la comparaison des structures de ces deux intermédiaires ont été exploitées afin de mieux comprendre les implications structurales de l'étape d'échange de la molécule de cofacteur dans la catalyse.

## ***A.2. Structure de l'intermédiaire réactionnel thioacylenzyme de la GAPDH de type sauvage***

### ***A.2.1. Stratégie***

Une structure cristallographique ou plus exactement la densité électronique dans laquelle est modélisée la structure protéique, correspond en fait à une densité électronique moyenne résultant de la contribution de l'ensemble des molécules composant le cristal. De ce fait, pour qu'un élément soit visible dans les cartes de densité électronique, celui-ci doit adopter une conformation identique dans une majorité des molécules du cristal. Aussi, résoudre la structure d'un intermédiaire réactionnel implique d'avoir préalablement réussi à trouver des conditions telles que l'intermédiaire constitue l'espèce majoritaire du cristal et que son taux d'occupation soit suffisamment élevé pour permettre une interprétation non ambiguë des cartes de densité électronique. L'approche retenue pour isoler cet intermédiaire a consisté à mettre à profit le fait qu'en absence de phosphate, l'étape de désacylation est fortement ralentie, car elle consiste dans ce cas en une réaction d'hydrolyse, (attaque nucléophile d'une molécule d'eau) qui est peu efficace par rapport à la réaction de phosphorylation. Ainsi, à pH 8,2 la vitesse d'acylation est  $6 \cdot 10^4$  fois plus rapide que l'hydrolyse ( $k_{\text{acylation}} = 850 \text{ s}^{-1}$  contre  $k_{\text{hydrolyse}} = 1,4 \cdot 10^{-2} \text{ s}^{-1}$ , Michels *et al.*, 1996) et l'accumulation de l'intermédiaire dans le site actif de l'enzyme est théoriquement possible. Une fois accumulé, l'intermédiaire est piégé par congélation des cristaux à 100 K et les données de diffraction collectées.

La première contrainte pour la recherche de conditions de cristallisation de la GAPDH était donc l'absence de phosphate dans le milieu. Par ailleurs, la vitesse de désacylation par l'eau étant proportionnelle à la concentration de l'espèce  $\text{OH}^-$  et donc directement dépendante du pH, le choix

du pH de cristallisation s'est porté sur des pH non basiques. Cette valeur de pH ne pouvant toutefois être trop abaissée pour préserver une réactivité suffisante de la cystéine catalytique (pKa de 6,2 Talfournier *et al* (1998)), une valeur de pH de 6,9 nous a semblé être un compromis acceptable. Enfin, plutôt que d'utiliser la condition de cristallisation déjà décrite par Skarzynski *et al.* (1987), dont les anions sulfate, en forte concentration, auraient pu entrer en compétition avec le phosphate du substrat pour la fixation dans les sites de reconnaissance anionique de l'enzyme, le choix du PEG comme agent précipitant nous avait semblé préférable.

### ***A.2.2. Cristallogenèse et essais préliminaires de trempage des cristaux d'holoenzyme en présence de substrat (G3P)***

Trois conditions de cristallisation répondant aux critères énoncés ci-dessus ont été isolées de la matrice d'échantillonnage Clear Strategy (Molecular Dimensions) puis optimisées (*cf.* Tableau II-1, conditions A1, A2 et A3). Les cristaux, de forme cubique ou parallélépipédique, présentent tous les mêmes caractéristiques que ceux décrits par Didierjean *et al.* (2003). Ils appartiennent au groupe d'espace  $C2$  avec des paramètres de maille similaires du type  $a = 140 \text{ \AA}$ ,  $b = 88 \text{ \AA}$ ,  $c = 120 \text{ \AA}$ ,  $\alpha = 90^\circ$ ,  $\beta = 120^\circ$ ,  $\gamma = 90^\circ$ . Basés sur les conditions décrites pour l'obtention des complexes ternaires C149A/S - NAD<sup>+</sup> - G3P (Didierjean *et al.*, 2003), les premiers essais de trempage des cristaux en présence de 1 mM de substrat, ont tout d'abord mis en évidence l'extrême fragilité de ces cristaux. En effet, dès l'ajout du substrat dans le milieu, des fissures (visibles lors de l'inspection sous la loupe binoculaire) apparaissent dans les cristaux, rendant impossible la collecte de données de diffraction. A ce stade, les cristaux issus des conditions A2 et A3 ne présentant aucun avantage par rapport aux cristaux issus de la condition A1, toutes les expériences décrites par la suite se sont focalisées sur des cristaux obtenus par la condition A1. Le protocole de trempage initial, présentant un nombre important de manipulations potentiellement néfastes pour l'ordre cristallin, a été modifié de façon à ce que l'incubation en présence du substrat et la cryoprotection des cristaux soient réalisées en une seule et même étape. Dans le nouveau protocole, la concentration en PEG 4000 de la goutte de cristallisation est tout d'abord graduellement augmentée (jusqu'à 30 %) en vue de la cryoprotection. Puis, cette solution cryoprotectrice (30 % PEG 4000, 0,3 M Na acétate, 0,1 M Tris HCl pH 6,9) est peu à peu échangée contre une solution identique mais contenant le substrat. Le cristal est alors prélevé puis congelé dans l'azote liquide. Il n'est donc plus manipulé qu'une seule fois contre trois auparavant. Les expériences de diffraction consécutives ont confirmé que ce protocole de trempage était moins néfaste pour les cristaux. En effet, bien qu'une proportion importante de cristaux trempés en présence du substrat conduise toujours à des clichés de

diffraction de piètre qualité, c'est-à-dire avec des taches diffuses et des limites de résolution faibles (4 à 6 Å), certains d'entre eux diffractent et ont permis de collecter deux jeux de données de qualité acceptable (non montré ici). Le traitement de ces données a révélé que les cristaux, appartenant initialement au groupe d'espace  $C2$ , présentaient, après trempage en présence du substrat, un mode de symétrie inférieure,  $P2_1$ . Parallèlement, l'un des paramètres de dimension de la maille a doublé, illustrant la perte de symétrie ( $a = 120 \text{ \AA}$  ;  $b = 88 \text{ \AA}$  ;  $c = 270 \text{ \AA}$  ;  $\alpha = 90^\circ$  ;  $\beta = 102^\circ$  ;  $\gamma = 90$ ). Bien que ces observations semblaient indiquer qu'un événement, pouvant correspondre à la fixation du substrat, s'était produit au sein du cristal, la résolution des structures et l'analyse de la densité électronique au niveau du site actif n'ont pas permis de mettre en évidence la présence du substrat.

### ***A.2.3. Expériences de microspectrophotométrie - Optimisation des conditions de trempage***

#### ***A.2.3.1. Préambule***

Suite aux résultats négatifs obtenus lors des premiers essais de résolution de la structure de l'intermédiaire thioacylenzyme (absence de signal attribuable à la présence de substrat dans les cartes de densité électronique), il paraissait nécessaire de s'assurer de la faisabilité de l'obtention d'une telle structure. Nous avons alors pris contact avec le Professeur Andrea Mozzarelli du Laboratoire de Biochimie et de Biologie Moléculaire de l'Université de Parme en Italie, spécialiste de la microspectrophotométrie sur monocristaux (Mozzarelli and Rossi, 1996 ; Pearson *et al.*, 2004 ), et qui avait par ailleurs déjà réalisé des études de ce type sur la GAPDH (Mozzarelli *et al.*, 1982). Une collaboration a donc été engagée, dont le but était tout d'abord de suivre, par microspectrophotométrie, la formation de l'intermédiaire réactionnel directement à l'intérieur des cristaux *via* la formation du NADH à 340 nm, concomitante à l'apparition du thioacylenzyme, et de mettre au point des conditions de trempage adéquates pour l'accumulation de l'intermédiaire. Etant donné la fragilité des cristaux obtenus en présence de PEG, nous avons également décidé d'utiliser la condition de cristallisation connue de la GAPDH en sulfate d'ammonium et de l'adapter à notre objectif en éliminant le tampon phosphate, incompatible avec l'isolement de l'intermédiaire thioacylenzyme, et en le remplaçant par un tampon Tris.

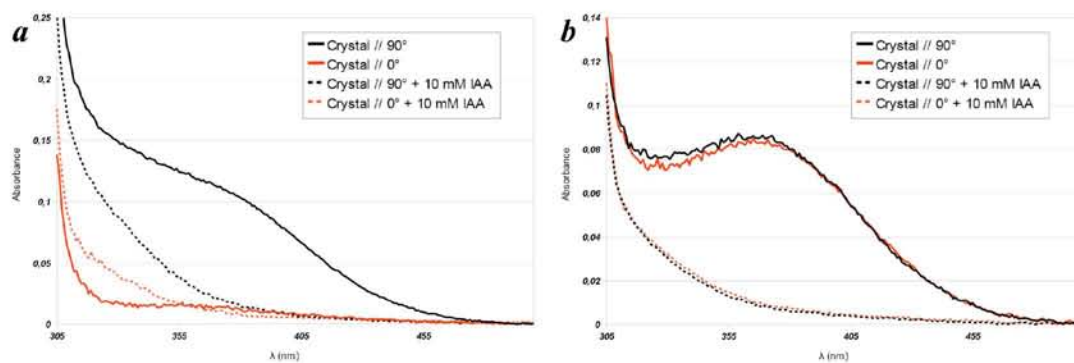
## *A.2.3.2. Résultats des expériences de microspectrophotométrie*

### *A.2.3.2.1. Expériences d'alkylation des cristaux de GAPDH*

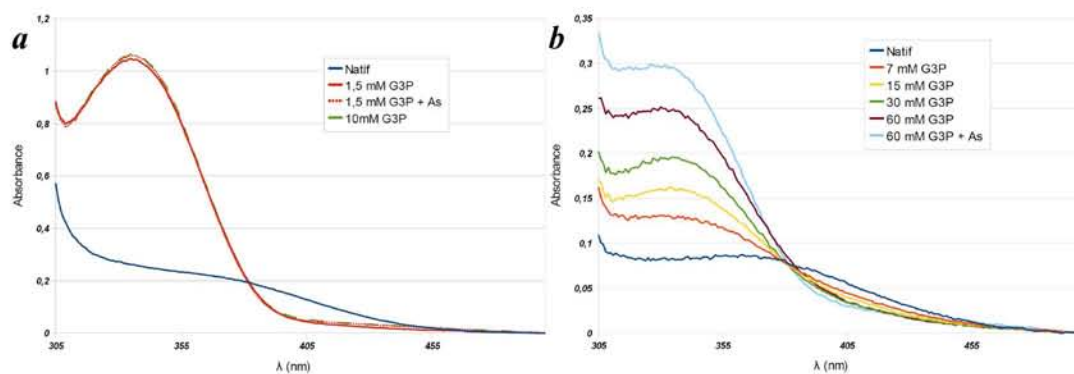
Dans le but de vérifier la réactivité de la cystéine catalytique au sein du cristal, les cristaux de GAPDH holoenzyme ont d'abord été soumis à l'acide iodoacétique (IAA). En effet, la GAPDH présente des caractéristiques spectroscopiques particulières avec l'existence, autour de 365 nm, d'une bande d'absorption caractéristique, appelée bande de Racker (Racker and Krimsky, 1952). Cette propriété est attribuée à la proximité structurale existant entre la charge positive globale du noyau pyridinium et la charge négative portée par la cystéine sous forme thiolate, activée. Lors de l'alkylation de la cystéine ou de son acylation par le substrat, cette bande d'absorption est abolie. L'incubation des cristaux en présence d'IAA (Figure III-2) abolit très nettement la bande de Racker et confirme ainsi que la cystéine (Cys149) du site actif de l'enzyme est bien sous forme réactive.

### *A.2.3.2.2. Incubation des cristaux de GAPDH holoenzyme en présence du substrat*

Les spectres d'absorption ont été ensuite enregistrés sur les cristaux de GAPDH incubés en présence de G3P (Figure III-3). Préalablement à l'incubation en présence du substrat, l'enregistrement d'un spectre en condition "native" permet de s'assurer de la présence de la bande de Racker. Ensuite, les cristaux sont incubés en présence du substrat et l'apparition du NADH est suivie à une longueur de 340 nm. L'incubation des cristaux PEG avec 1,5 mM de *D,L*-G3P conduit à l'apparition d'un signal à 340 nm. Ce signal augmente au cours du temps et se stabilise ensuite pour rester constant pendant au moins 3h à un pH de 8,2. Des expériences similaires réalisées à un pH de 6,9 montrent, au contraire, qu'après avoir atteint un maximum, le signal à 340 nm décroît. Ce phénomène a été attribué à l'instabilité du NADH à ce pH (résultats non montrés). L'ajout final d'arséniate dans le milieu, qui déplace entièrement la réaction et consomme la totalité du  $\text{NAD}^+$  présent dans les cristaux pour former du NADH, permet d'obtenir le signal maximum possible à 340 nm (réaction totale). Le spectre obtenu après 3h d'incubation en présence du substrat est équivalent à celui obtenu après incubation avec de l'arséniate, ceci confirmant que la réaction avec le substrat est totale. Des résultats similaires sont obtenus pour l'incubation des cristaux avec 10 mM de substrat (Figure III-3), mais le signal à 340 nm atteint plus rapidement son maximum (<10 minutes) et l'ajout d'arséniate confirme également que la totalité du  $\text{NAD}^+$  a été réduit en NADH lors de la



**Figure III-2 : Spectres d'absorption des cristaux de GAPDH obtenus en présence de PEG (a) ou de sulfate d'ammonium (b) en condition native (tracé continu) et après alkylation (tracé pointillé) par 10 mM d'acide iodoacétique. Deux spectres d'absorption sont enregistrés pour chaque condition, le premier (noir) avec l'axe b du cristal orienté parallèlement au vecteur électrique de la lumière polarisée, le second (rouge) après une rotation de 90° du cristal.**

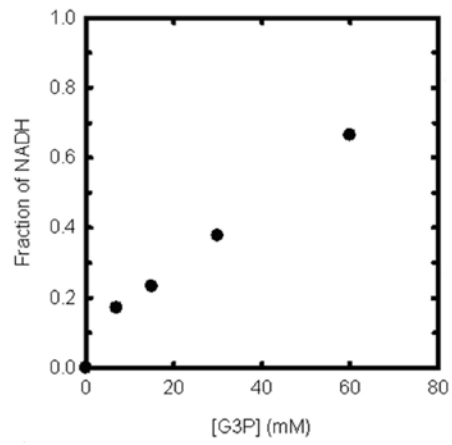


**Figure III-3 : Spectres d'absorption mesurés sur les cristaux de GAPDH holoenzyme obtenus en présence de PEG (a) ou de sulfate d'ammonium (b) et incubés en présence de G3P.**

réaction avec le substrat.

La même expérience a été réalisée avec les cristaux obtenus en présence de sulfate d'ammonium. Après confirmation de la présence de la bande de Racker, les cristaux ont été incubés en présence du substrat. Les premiers essais, réalisés avec 1,5 mM de *D,L*-G3P, ont montré que le signal à 340 nm augmentait au cours du temps mais qu'il stagnait à des valeurs bien inférieures au signal maximal (obtenu en présence d'arséniate). Aussi, un cristal a-t-il été incubé en présence de concentrations croissantes de substrat (7, 15, 30 et 60 mM), et le spectre d'absorption correspondant a été enregistré pour chaque concentration après équilibration du signal à 340 nm (Figure III-3). Ces expériences ont montré que la quantité de NADH formé était proportionnelle à la quantité de substrat présente dans le milieu (Figure III-4). Pour une concentration de 60 mM en *D,L*-G3P (qui pour des raisons techniques peut difficilement être augmentée au-delà), le signal à 340 nm demeure cependant toujours inférieur au signal obtenu après ajout d'arséniate. La représentation graphique de la fraction de NADH formé en fonction de la concentration en substrat (Figure III-4) permet de quantifier le taux d'avancement de la réaction. Ainsi, en présence de 60 mM de *D,L*-G3P, environ 70 % du NAD<sup>+</sup> présent dans le cristal a été réduit en NADH et la réaction n'est donc pas complète. Ce phénomène a été attribué à une compétition existant entre les ions sulfates présents dans le milieu à une concentration élevée (2,7 M) et le phosphate du substrat pour la fixation dans le site actif de l'enzyme.

Différents protocoles d'incubation des cristaux en présence du substrat ont alors été établis. Deux protocoles dits de trempage "court" ou "long" ont été définis pour les cristaux PEG. Le trempage long consiste à incuber pendant 1 h les cristaux en présence de 1,5 mM de substrat et le trempage court à incuber pendant 10 minutes les cristaux avec 10 mM de *D,L*-G3P. Chacun des deux protocoles a été suivi par la congélation des cristaux dans l'azote liquide. Bien qu'il soit impossible techniquement d'obtenir une réaction totale avec les cristaux SA, un protocole de trempage a néanmoins été mis au point. Celui-ci a consisté à incuber les cristaux pendant 1 à 10 minutes en présence de la concentration maximale de substrat (environ 60 mM) dans une solution contenant 2,7 M de sulfate d'ammonium et 100 mM de tampon Tris pH 8,2. Pour limiter la dilution du substrat par la solution de trempage, le sulfate d'ammonium et le tampon Tris, une fois pesés, ont été dilués directement dans la solution de *D,L*-G3P et le pH de la solution a été ajusté. Après trempage, les cristaux ont été rapidement passés dans la solution cryoprotectrice avant leur congélation.



**Figure III-4 : Représentation graphique de la fraction de NADH formé en fonction de la concentration en D,L-G3P utilisé pour le trempage des cristaux.**

## A.2.4. Enregistrement des données de diffraction

### A.2.4.1. Cristaux PEG

Malgré l'optimisation des conditions, le trempage des cristaux obtenus en présence de PEG reste très délicat. En effet, une forte proportion de cristaux se fissure lors du trempage en présence du substrat et devient inutilisable pour la collecte de données de diffraction. La taille des cristaux semble déterminante, les plus petits résistant plus facilement aux conditions de trempage. Sept jeux de données de diffraction ont été néanmoins collectés sur des cristaux PEG incubés selon les différents protocoles établis. La première collecte, effectuée sur FIP, présente une faible complétude et une résolution très moyenne de 3,3 Å (Tableau III-1), cependant les données ont été utilisées avec succès pour établir la première solution de remplacement moléculaire. Quatre tétramères du modèle holoenzyme (1gd1, Skarzynski *et al.*, 1987) ont ainsi été positionnés dans l'unité asymétrique grâce au logiciel Molrep (Vagin and Teplyakov, 1997). Les jeux de données collectés ensuite sur les lignes de lumières X11 et X13 de l'EMBL de Hambourg présentent des statistiques satisfaisantes (Tableau III-1). Chacun de ces jeux de données a été utilisé dans un processus d'affinement d'une structure (jusqu'à des valeurs de  $R_{\text{free}}$  de l'ordre de 25 %) permettant une interprétation explicite de la densité électronique notamment au niveau du site actif de l'enzyme. Dans tous les cas, aucun signal pouvant être clairement attribué à la présence du substrat n'a été observé dans la continuité de la cystéine catalytique ou au niveau des sites de reconnaissance anionique.

Dans tous les cas, les cristaux appartenant initialement au groupe d'espace  $C2$  (Didierjean *et al.* (2003), et essais sur des cristaux non trempés) passent à un groupe d'espace  $P2_1$ . La nouvelle maille dérive d'une transformation du type  $\vec{a}_{P21} = \vec{c}_{C2}$  ;  $\vec{c}_{P21} = 2\vec{a}_{C2} + \vec{c}_{C2}$  mais a également subi une déformation correspondant à une expansion. Cette réduction de symétrie pourrait être attribuée aux conditions de trempage. Néanmoins, le fait que des cristaux soumis uniquement à la cryoprotection conservent les paramètres cristallins initiaux, suggère fortement que le substrat ou même sa réaction avec l'enzyme sont à l'origine de ce changement. En effet, la structure du complexe ternaire GAPDH-G3P-NAD<sup>+</sup> a été obtenue avec le même type de cristaux, et ceci sans aucun changement de groupe d'espace. Ainsi, la seule présence du substrat, ou sa fixation dans le site actif de l'enzyme ne peuvent être à l'origine de cette réduction de symétrie des cristaux. C'est donc la réaction entre l'enzyme et son substrat, ou encore un événement structural qui lui est associé, qui est



**Tableau III-1 : Statistiques des jeux de données de diffraction collectés sur des cristaux de BstGAPDH obtenus en présence de PEG et incubés en présence de G3P**

<b>Synchrotron</b>	ESRF, Grenoble	DESY, Hambourg					
<b>Ligne de lumière</b>	FIP BM30A	X11					X13
<b>Longueur d'onde (Å)</b>	0,9796	0,814	0,816				0,806
<b>Trempage</b>	"long"		"court"				
	1h30 1,5 mM <i>D,L</i> -G3P		10'	10'	12'	20'	4'
	10 mM <i>D,L</i> -G3P						
<b>Groupe d'espace</b>	P 2 <sub>1</sub>						
<b>Maille (a, b, c et β)</b>	~ 120 Å ~ 88 Å ~ 266 Å ; 102,4°						116 Å 173 Å 130 Å ; 112,3°
<b>Limite de résolution (Å)</b>	47 - 3,3 (3,5-3,3)	27 - 2,8 (3,0-2,8)	30 - 2,15 (2,20-2,15)	55 - 2,18 (2,23-2,18)	36 - 2,55 (2,61-2,55)	34 - 2,4 (2,46-2,40)	30 - 2,36 (2,40-2,36)
<b>Réflexions uniques</b>	68359	131765	293851	245257	173452	293851	195410
<b>Redondance</b>	1,7	3,8	3,0	2,5	2,6	2,3	4,2
<b>Complétude (%)</b>	79,8 (83,2)	98,2 (91,7)	99,2 (92,2)	86,9 (56,8)	99,1 (98,6)	96,8 (86,2)	99,8 (99,8)
<b>I/σ(I)</b>	7,0 (2,2)	17,2 (7,5)	16,9 (3,7)	15,9 (3,7)	11,7 (2,5)	18,4 (4,4)	20,6 (6,5)
<b>Rmerge (%)</b>	10 (37,2)	6,8 (20,3)	3,9 (15,4)	3,3 (14,1)	4,3 (18,0)	2,7 (10,0)	5,1 (21,8)

Les valeurs entre parenthèses font référence à la dernière couronne de résolution

vraisemblablement à l'origine de ce phénomène. Lors de la transition du cristal de  $C 2$  vers  $P 2_1$ , le coefficient de Matthews passe de 2,19 à 2,35 Å<sup>3</sup>/Da. Il semblerait donc que la réaction entraîne une "dilatation" du cristal à l'origine, sur le plan microscopique, de la réduction de symétrie et, sur le plan macroscopique, de l'apparition des fissures.

#### A.2.4.2. Cristaux SA

Les cristaux SA ont également été soumis aux rayons X après incubation avec le substrat. Bien qu'il soit impossible d'obtenir une réaction totale avec ces cristaux SA, ceux-ci présentent cependant une bonne résistance aux conditions de trempage. Aucun signe de dégradation visuelle des cristaux n'est, en effet, observé durant l'incubation en présence du substrat, et les paramètres cristallins ne sont pas modifiés. Cinq jeux de données ont été collectés sur ces cristaux incubés de 1,5 à 10 minutes en présence de forte concentration de substrat (60 mM *D,L*-G3P). Leurs statistiques sont présentées dans le Tableau III-2.

Comme indiqué précédemment, les cristaux SA, caractérisés initialement par Skarzynski *et al.* (1987), conservent les mêmes paramètres cristallins. Monocliniques, ils appartiennent au groupe d'espace  $P2_1$  et présentent des paramètres de maille du type  $a = 81,7 \text{ \AA}$ ,  $b = 122,5 \text{ \AA}$ ,  $c = 81,8 \text{ \AA}$ ,  $\alpha = 90^\circ$ ,  $\beta = 111,2^\circ$  et  $\gamma = 90^\circ$ . La méthode du remplacement moléculaire en utilisant le tétramère de référence (1gd1, Skarzynski *et al.*, 1987) et le logiciel Molrep ou un simple affinement en corps rigide (Refmac5) a permis de positionner le modèle dans la maille et conduit à des valeurs de R de l'ordre de 35%. Un processus d'affinement a été conduit sur l'ensemble de ces structures jusqu'à un stade relativement avancé. Néanmoins, la structure dérivant du 4<sup>e</sup> jeu de données, qui présentait la densité la plus claire pour l'intermédiaire réactionnel, est la seule pour laquelle l'affinement a été mené à terme et dont le modèle, déposé dans la Protein Data Bank, a été publié (code pdb 3cmc, *cf.* Annexe 1). Les statistiques d'affinement de la structure sont présentées dans le Tableau III-3. L'affinement initial a été effectué avec le logiciel Refmac5 jusqu'à la convergence des valeurs de R et  $R_{\text{free}}$  à des valeurs respectives de 18,2 et 23,7 %. La structure présentant deux états différents de l'enzyme (*cf.* ci-dessous), les taux d'occupation respectifs de chacun de ces états ont été estimés en utilisant à la fois les facteurs d'agitation thermique et les cartes différences. Dans un souci de précision et pour éliminer toute part de subjectivité, nous nous sommes ensuite adressés au Docteur Gérard Bricogne (Global Phasing Ltd., Cambridge) qui développe le logiciel Buster (suite de programme TNT ; Bricogne and Irwin, 1996). En effet, par rapport à Refmac5 qui ne permet d'affiner que les coordonnées atomiques et les facteurs d'agitation thermique, le programme Buster-TNT permet, en parallèle, d'affiner les facteurs d'occupation avec exactitude. Les cartes de densité

**Tableau III-2 : Statistiques des jeux de données de diffraction collectés sur des cristaux de BstGAPDH obtenus en présence de sulfate d'ammonium et incubés en présence de G3P.**

<b>Synchrotron</b>	DESY, Hambourg				
<b>Ligne de lumière</b>	X11	X13			
<b>Longueur d'onde (Å)</b>	0,814	0,806			
<b>Trempage</b>	10' - 67 mM <i>D,L</i> -G3P	5'	5'	3'	1'30"
		70 mM <i>D,L</i> -G3P			
<b>Groupe d'espace</b>	P 2 <sub>1</sub>				
<b>Maille (a, b, c et β)</b>	~81,7 Å ~122,7 Å ~81,8 Å ~111,1°				
<b>Limite de résolution (Å)</b>	48,6 - 2,11 (2,24 - 2,11)	29,3 - 2,0 (2,1 - 2,0)	30,2 - 2,03 (2,13 - 2,03)	29,5 - 1,77 (1,88 - 1,77)	29,8 - 2,02 (2,10 - 2,02)
<b>Réflexions totales</b>	351730	370230	530453	528112	449508
<b>Réflexions uniques</b>	84693	98306	96372	142733	97991
<b>Redondance</b>	4,2	3,8	5,5	3,7	4,6
<b>Complétude (%)</b>	97,6 (86,2)	99,1 (94,0)	99,7 (98,8)	97,7 (92,9)	98,6 (90,1)
<b>I/σ(I)</b>	19,6 (6,8)	42,7 (11,0)	36,1 (11,0)	32,1 (6,5)	28,2 (6,8)
<b>Rmerge (%)</b>	5,7 (24,9)	2,5 (11,1)	3,5 (15,2)	2,8 (18,9)	3,8 (19,3)

Les valeurs entre parenthèses font référence à la dernière couronne de résolution

**Tableau III-3 : Statistiques du jeu de données de diffraction et d'affinement de la structure de l'intermédiaire thioacylenzyme de la BstGAPDH.**

<b>Données de diffraction</b>	
Groupe d'espace	<i>P</i> 2 <sub>1</sub>
Maille (Å)	<i>a</i> = 81,7 ; <i>b</i> = 122,6 ; <i>c</i> = 81,8 ; $\beta$ = 111,1°
Z	2
Limite de résolution (Å)	29,50 - 1,77
Couronne de plus haute résolution (Å)	1,88 - 1,77
Température (K)	100
Réflexions uniques	142733 (20259)
Complétude (%)	97,7 (92,9)
Redondance	3,7 (3,2)
R-merge <sup>a</sup> (%)	2,8 (18,9)
I/ $\sigma$ (I)	32,1 (6,5)
<b>Affinement</b>	
R <sup>b</sup> (%)	16,1
R <sub>free</sub> <sup>c</sup> (%)	19,8
<b>r.m.s.d. <sup>d</sup> par rapport à la géométrie standard</b>	
Longueur de liaisons (Å)	0,011
Angles de valence (°)	1,27
Angles dièdres (°)	17,0
<b>Facteur B moyens (Å<sup>2</sup>)</b>	
Modèle	25,7
Atomes de la protéine	23,6
NAD	17,1
Intermédiaire réactionnel	27,4
Anions sulfates (sites Pi et Ps)	40,4 (45,8 – 34,7)
Molécules d'eau	37,7

<sup>a</sup> Facteur R pour les intensités reliées par symétrie. <sup>b</sup> Facteur R cristallographique. <sup>c</sup> Facteur R cristallographique calculé sur 5 % des réflexions choisies aléatoirement et non incluses dans l'affinement.

<sup>d</sup> Écart quadratique moyen.

électronique issues de ce programme sont par ailleurs globalement d'une meilleure qualité que celles produites par les programmes conventionnels, autorisant à la fois un meilleur niveau de détail visible, et l'obtention d'un modèle final de meilleure qualité.

### ***A.2.5. Présentation de la structure et qualité du modèle***

Le modèle final est composé d'un homotétramère dont chacune des sous-unités de 334 résidus contient une molécule de cofacteur. Le modèle comporte également quatre molécules de G3P (en occupation partielle) liées à l'enzyme, 18 anions sulfates (dont deux dans chacun des sites actifs), 8 molécules de glycérol et 1740 molécules d'eau. A l'exception de quelques chaînes latérales exposées au solvant, les résidus d'acides aminés de la protéine sont bien définis dans la densité électronique et seuls quelques pics de densité demeurant dans les cartes différences ( $F_{\text{obs}} - F_{\text{calc}}$ ) n'ont pas pu être expliqués. Le modèle présente une stéréochimie de bonne qualité avec notamment des valeurs de rmsd sur les longueurs de liaison ou sur les angles de valence de 0,011 Å et 1,576°, respectivement (Tableau III-3). Le diagramme de Ramachandran (Ramakrishnan and Ramachandran, 1965) indique que l'ensemble des résidus non glycine à l'exception de l'Asp186 et de la Val237 est situé dans des régions favorables du diagramme. Les deux résidus particuliers sont cependant systématiquement localisés en dehors des régions admises du diagramme de Ramachandran, dans les structures cristallines de GAPDH issues de *Bacillus stearothermophilus* ou d'autres sources (Duée *et al.*, 1996).

L'un des quatre monomères a été arbitrairement noté O puis, en suivant la nomenclature établie pour la GAPDH de *Homaricus americanus* par Buehner *et al.* (1974), les 3 autres ont été nommés P, Q et R. La superposition des carbones  $\alpha$  des différentes sous-unités conduit à une valeur de rmsd moyenne de 0,125 Å qui indique que la conformation des quatre monomères est très similaire (Tableau III-4). En utilisant le monomère O comme référence et en le superposant aux différentes sous-unités de la structure holoenzyme de la *Bst*GAPDH (1gd1, Skarzynski *et al.*, 1987) on obtient, pour les carbones  $\alpha$ , une valeur de rmsd moyenne de 0,25 Å, qui montre que la réaction d'acylation n'induit pas de changement conformationnel majeur de la structure du monomère. La seule différence notable existant entre les deux structures se situe au niveau de la boucle 206-212 qui montre des écarts représentant jusqu'à 15 fois la valeur de rmsd. Cette particularité de la structure de l'intermédiaire thioacylenzyme par rapport à la forme holoenzyme sera développée plus tard. Lorsque l'on s'intéresse au tétramère dans son ensemble, et qu'on le compare à la structure du tétramère holoenzyme (1gd1), la valeur de rmsd sur les 1336 carbones  $\alpha$  est de 0,48 Å. Cette faible

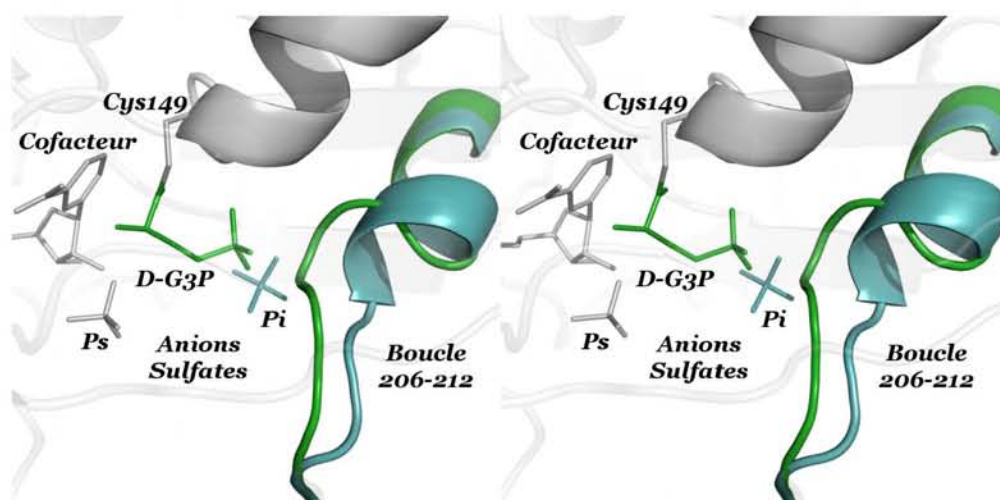
**Tableau III-4 : Valeurs de rmsd résultant de la superposition des monomères de la structure de l'intermédiaire thioacylenzyme et comparaison avec la structure holoenzyme.** Sur fond grisé sont indiquées les valeurs de rmsd (en Å) résultant de la superposition des carbones  $\alpha$  des 4 sous-unités de la structure du thioacylenzyme 2 à 2 et sur fond bleu sont indiquées les valeurs de rmsd résultant de la superposition des carbones  $\alpha$  du monomère O (pris comme référence) sur les différentes sous-unités de la structure la GAPDH de *B. stearothermophilus* sous forme holoenzyme (Skarzynski et al., 1987; code pdb 1gd1). Les résidus 206-212, modélisés selon deux conformations alternatives dans la structure du thioacylenzyme, ne sont pas considérés dans le calcul.

	P	Q	R	O
O	0,156	0,148	0,117	0,254
	P	0,128	0,126	0,244
		Q	0,142	0,233
			R	0,274

valeur suggère que la structure quaternaire de l'enzyme n'est pas différente dans les formes holoenzyme et thioacylenzyme.

Comme attendu d'après la stœchiométrie de la réaction entre les cristaux SA et les 60 mM de *D,L*-G3P déterminée au travers des expériences de microspectrophotométrie (i.e. environ 70 % du  $\text{NAD}^+$  transformé en NADH ; Figure III-4), le site actif de l'enzyme est observé dans deux états chimiques distincts et qui correspondent respectivement à l'holoenzyme sous forme native (cystéine catalytique libre et un ion sulfate dans le site  $\text{P}_i$ ) et à l'intermédiaire thioacylenzyme (cystéine catalytique acylée par le *D*-G3P). Ces deux états ont été respectivement notés conformation A et conformation B. La différence entre les deux états inclut également la boucle 206-212 qui borde le site de reconnaissance anionique  $\text{P}_i$  et participe à sa formation, et qui a été modélisée dans deux conformations alternatives, A et B, chacune associée à l'un des états présentés ci-dessus. Ainsi la conformation A de la boucle est-elle associée à la présence d'un anion sulfate dans le site  $\text{P}_i$  (holoenzyme) et la conformation B, à la présence du groupement phosphate du substrat (Figure III-5). La proximité structurale existant entre la boucle en conformation B et l'anion sulfate en  $\text{P}_i$  (moins de 2 Å) exclut en effet toute coexistence de ces deux éléments et confirme que la présence de l'anion sulfate est associée à la conformation A de la boucle alors que la conformation B est, elle, associée à la présence du groupement phosphate du substrat (Figure III-5).

Les facteurs d'occupation et les facteurs d'agitation thermique des ligands (valeurs affinées par Clemens Vornrhein avec le logiciel Buster-TNT 2.0) sont similaires dans les quatre sous-unités de l'enzyme et traduisent une distribution équivalente des ligands dans tous les monomères (Tableau III-5). L'affinement des facteurs d'occupation de la boucle 206-212, en revanche, suggère l'existence de deux types de sous-unités différentes. Ainsi, alors que dans les monomères O et R, les deux conformations de la boucle ont des facteurs d'occupation sensiblement identiques, la conformation A (conformation holoenzyme) semble prédominer dans les monomères P et Q. Bien que ces boucles ne soient pas directement impliquées dans les interactions cristallines, de telles différences pourraient provenir de l'environnement cristallin particulier de chacune des sous-unités (pour une description plus complète, se reporter à Skarzynski *et al.*, 1987). Malgré ces quelques variations des facteurs d'occupation, les conformations de la boucle 206-212 et des ligands sont très similaires dans les quatre monomères. C'est pourquoi le monomère O, sauf mention contraire, sera utilisé comme référence dans la suite de ce manuscrit.



**Figure III-5 : Vue stéréoscopique des deux états, holoenzyme et thioacylzyme, du site actif.** Sont représentées en blanc les parties invariantes du modèle comme le cofacteur, la cystéine catalytique ou l'anion sulfate présent dans le site Ps. La boucle 206-212 en conformation "holoenzyme" (associée à la présence d'un anion sulfate en Pi) est représentée en bleu alors qu'elle est représentée en vert dans sa conformation "thioacylzyme" (associée à la présence du groupement phosphate du substrat).

**Tableau III-5 : Taux d'occupation et facteur d'agitation thermique du modèle affiné par Buster-TNT (Bricogne and Irwin, 1996).**

	Sous-unités	O	P	Q	R
<b>D-G3P</b>	Occupation	0,64	0,59	0,62	0,67
	Facteur B moyen (Å <sup>2</sup> )	27,3	27,9	26,8	27,6
	Coefficient de corrélation dans l'espace réel*	0,95	0,91	0,94	0,94
<b>Sulfate (site Pi)</b>	Occupation	0,33	0,47	0,41	0,27
	Facteur B(Å <sup>2</sup> )	46,1	45,6	45,1	46,3
<b>Boucle 206-212 (confA/confB)</b>	Occupation	0,47/0,53	0,66/0,34	0,65/0,34	0,44/0,56
	Facteur B moyen (Å <sup>2</sup> )	24,8/25,6	24,5/25,2	23,8/24,5	24,9/25,5
	Coefficient de corrélation dans l'espace réel*	0,93/0,87	0,95/0,90	0,95/0,89	0,94/0,88

\* Corrélation existant entre la densité électronique du modèle moléculaire et la densité électronique calculée (2mFo-DFc) à partir des données de diffraction.

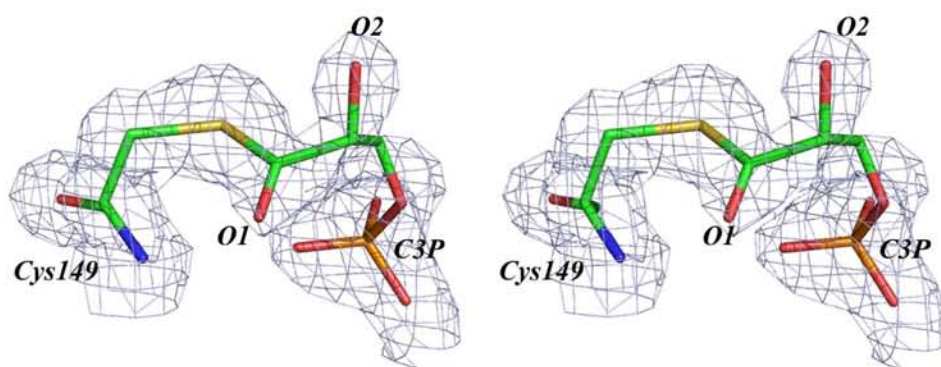


## A.2.6. Étude du site actif de l'enzyme

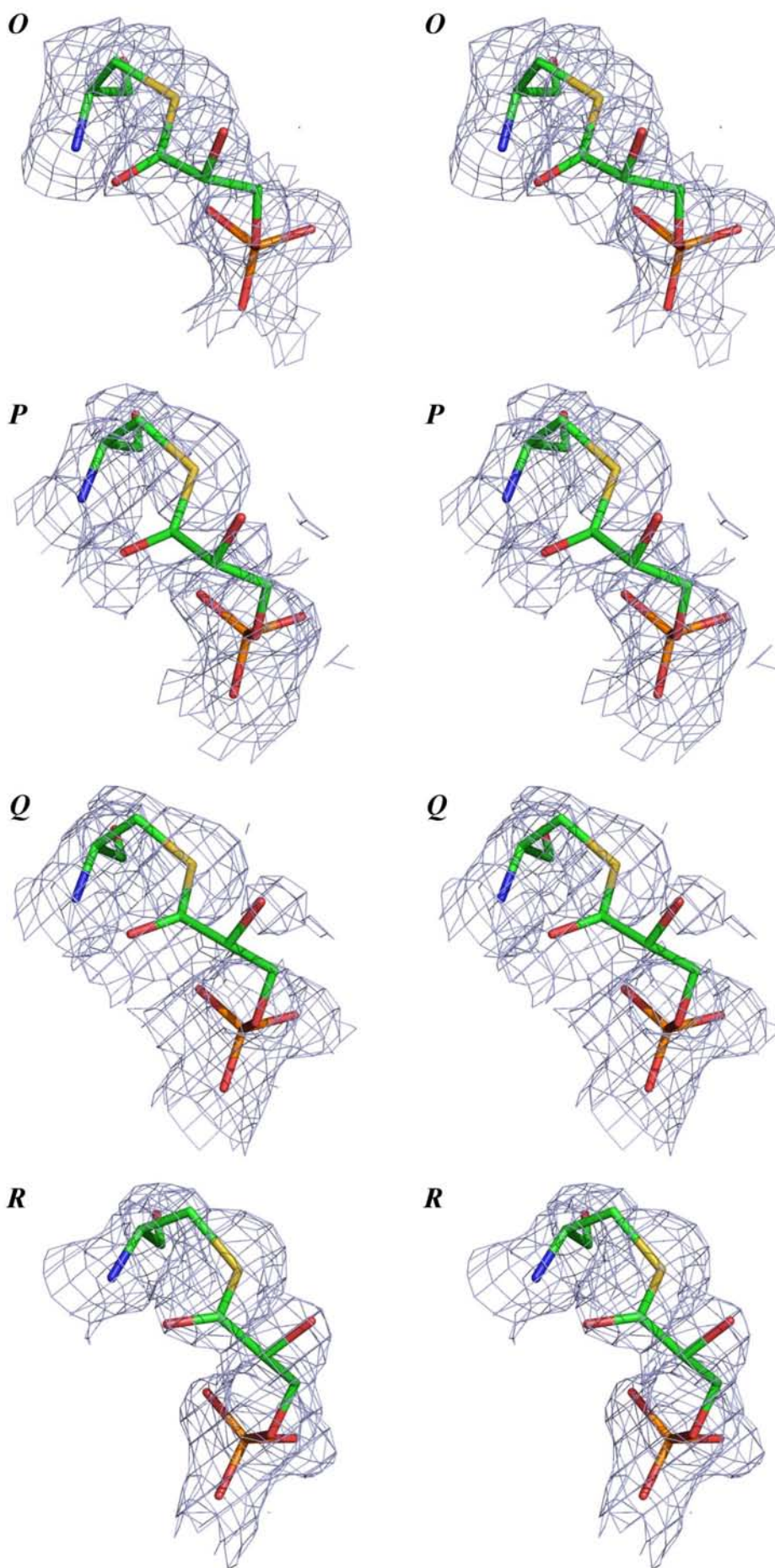
### A.2.6.1. Conformation de l'intermédiaire réactionnel GAPDH-G3P-NADH

La molécule de substrat est présente dans chacune des sous-unités avec un taux d'occupation variant de 0,59 à 0,67 (Tableau III-5). Dans les deux monomères O et R, les atomes composant la molécule de G3P sont assez bien définis dans les cartes de densité électronique finales ( $2mF_o-DF_c$ ) avec une connectivité à 0,8 et 0,9  $\sigma$  et des coefficients de corrélation dans l'espace réel de 0,95 et 0,94, respectivement (Figure III-6). Dans les monomères P et Q, la densité autour du substrat modélisé est de moins bonne qualité, surtout au niveau de l'atome C2. L'analyse des cartes "omits", calculées à l'aide du logiciel Buster-TNT, conduit au même type d'observation (Figure III-7). Bien qu'aucune contrainte de symétrie non cristallographique n'ait été appliquée aux atomes du substrat, la conformation de l'intermédiaire tout comme le réseau d'interactions formé avec la protéine est sensiblement identique dans les quatre sous-unités.

La molécule de 3-phosphoglyceroyl est liée à l'enzyme au travers d'une liaison thioester, formée entre l'atome C1 (*sp*<sup>2</sup>) du substrat et l'atome S $\gamma$  de la cystéine catalytique (Cys149) (Figures III-8). La liaison thioester adopte une conformation *trans* avec l'angle de torsion C $\beta$ -S $\gamma$ -C1-C2 oscillant autour de -155°. Le plan imposé par la liaison thioester est approximativement parallèle au plan défini par le noyau pyridinium du nicotinamide (Figure III-8). La fonction carbonyle portée par l'atome C1 pointe dans une direction opposée à l'histidine catalytique pour former des liaisons hydrogène avec l'atome d'azote de la chaîne principale du résidu Cys149 et avec une molécule d'eau. Bien qu'un mélange racémique du glycéraldéhyde-3-phosphate (*D,L*-G3P) ait été utilisé pour le trempage des cristaux, l'atome de carbone 2 (C2) du G3P est retrouvé dans une configuration R, confirmant la spécificité de l'enzyme pour le *D*-G3P (Byers, 1978). Oscillant autour de -130° dans les différentes sous-unités, l'angle de torsion O1-C1-C2-O2 adopte une conformation *trans*, énergétiquement favorable. L'atome O2 forme des liaisons hydrogène avec le cofacteur (atome NO7) et avec l'histidine catalytique (His176 atome N $\epsilon_2$ ). Le groupement phosphate (C3P) porté par l'atome C3 du substrat est localisé au niveau du nouveau site Pi, délimité notamment par la boucle 206-212 dans sa conformation B (*cf.* ci-dessous). Les quatre atomes d'oxygène du C3P forment des liaisons hydrogène avec les chaînes latérales des résidus Ser148, Thr150, Thr208 ainsi qu'avec les

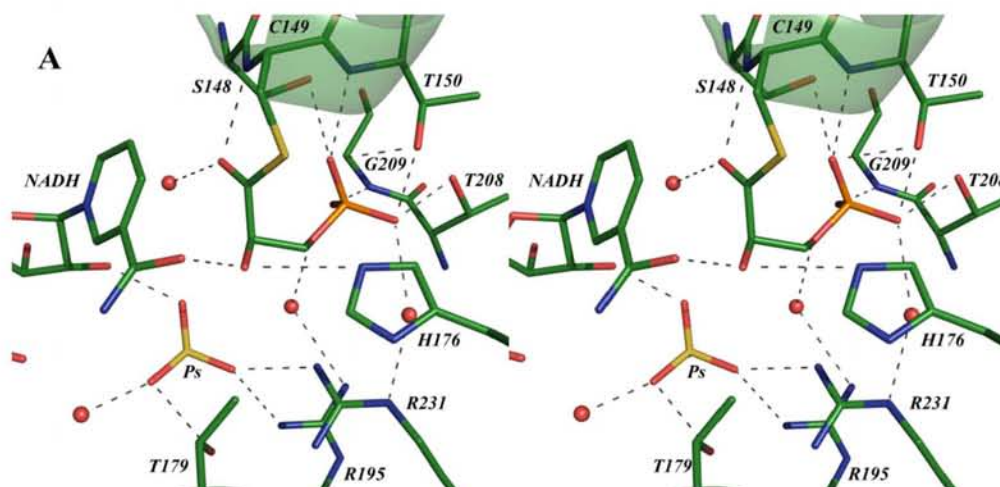


*Figure III-6 : Vue stéréoscopique de l'intermédiaire réactionnel de la sous-unité R dans sa carte de densité électronique. La carte est de type 2mFo-DFc et son niveau de contour de 1 sigma. De gauche à droite : la cystéine catalytique (C149) en liaison thioester avec le glyceroyl-3-phosphate.*



**Figure III-7 : Vues stéréoscopiques de l'intermédiaire thioacylzyme dans des cartes de densité électronique "omit".**

Les cartes ont été calculées avec le logiciel Buster-TNT (Bricogne and Irwin, 1996) après avoir éliminé du fichier de coordonnées les atomes de l'intermédiaire et les atomes de son environnement (à une distance maximale de 3,6 Å de celui-ci). Les cartes sont de type 2mFo-DFc et sont contourées à 1  $\sigma$ .



**Figure III-8 : Conformation de l'intermédiaire thioacylenzyme et interactions avec le site actif de la GAPDH.** Vue stéréoscopique du site actif de l'enzyme dans lequel le substrat, l'anion sulfate, le cofacteur ainsi que les résidus impliqués dans leur fixation sont présentés en mode bâton. Les liaisons hydrogène sont représentées en pointillés. Pour la longueur des liaisons, se reporter au Tableau III-6.

**Tableau III-6 : Distances interatomiques des liaisons hydrogène formées entre l'intermédiaire thioacylenzyme et l'enzyme.** Les chiffres entre parenthèses correspondent aux numéros de la molécule d'eau tels que définis dans le fichier de coordonnées atomiques

	O	P	Q	R
<b>O1-G3P --- N-Cys149</b>	2,7	2,6	2,7	2,7
<b>O1-G3P --- H<sub>2</sub>O</b>	3,0 (W558)	2,7 (W1547)	3,0 (W1056)	2,7 (W1213)
<b>O2-G3P --- NO7-NAD</b>	2,7	2,7	2,7	2,8
<b>O2-G3P --- N<math>\epsilon</math><sub>2</sub>-His176</b>	3,1	3,1	3,2	3,1
<b>O1P-G3P --- H<sub>2</sub>O</b>	3,0 (W395)	3,0 (W631)	2,9 (W1743) 3,0 (W367)	3,0 (W111)
<b>O2P-G3P --- O<math>\gamma</math><sub>1</sub>-Thr150</b>	2,6	2,5	2,5	2,5
<b>O2P-G3P --- O<math>\gamma</math><sub>1</sub>-Thr208</b>	2,7	2,8	2,9	2,7
<b>O2P-G3P --- H<sub>2</sub>O</b>	2,8	2,7	2,6	2,8
<b>O3P-G3P --- N-Gly209</b>	2,3	2,4	2,4	2,5
<b>O3P-G3P --- H<sub>2</sub>O</b>	/	/	3,2 (W367)	3,1 (W111) 3,1 (W1701)
<b>O4P-G3P --- O<math>\gamma</math>-Ser148</b>	3,0	3,0	3,2	2,9
<b>O4P-G3P --- N-Thr150</b>	2,9	3,1	3,0	3,0
<b>O4P-G3P --- O<math>\gamma</math><sub>1</sub>-Thr150</b>	2,7	2,8	2,6	2,7

chaînes principales des résidus Thr150 et Gly209, tous ces résidus étant invariants dans les séquences de GAPDH. Deux interactions supplémentaires sont créées au travers des molécules d'eau avec les chaînes latérales des résidus Arg195 (non conservé) et Arg231 (invariant). Les interactions entre le substrat et la protéine sont reportées dans le Tableau III-6.

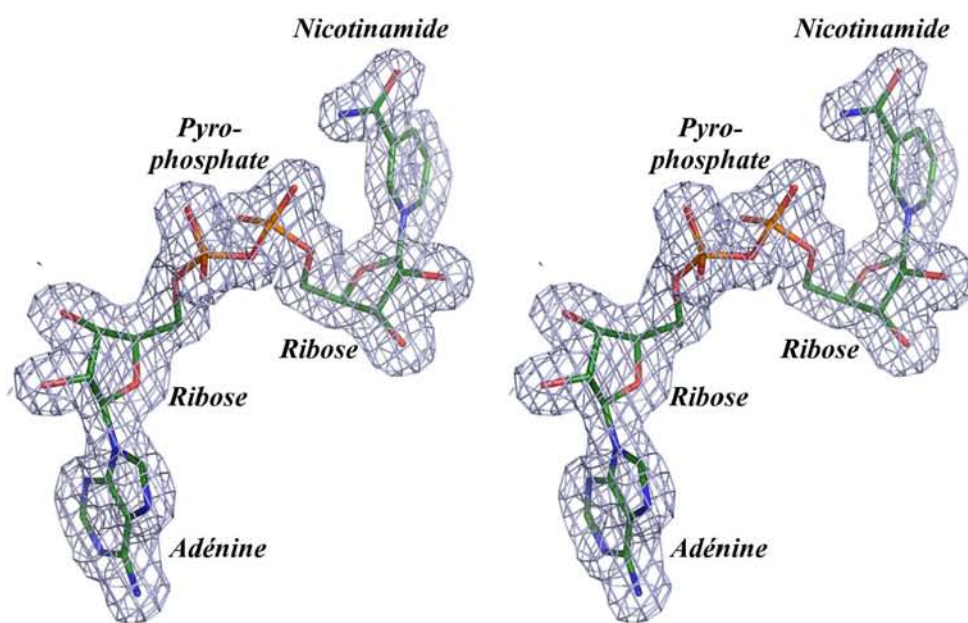
Étant donnée l'occupation partielle de l'intermédiaire, affinée à une valeur moyenne de 0,63, il faut considérer que la cystéine catalytique (Cys149) existe aussi sous forme libre. En se basant sur la forme de la densité et sur les différents essais d'affinement, il n'a pas été possible de distinguer deux conformations distinctes de la chaîne latérale de la cystéine. Celle-ci a donc été finalement modélisée dans une conformation unique, représentative des deux états, acylé et libre. Dans cette conformation, la chaîne latérale de la cystéine catalytique se superpose parfaitement à celle de la structure holoenzyme connue (Skarzynski *et al.*, 1987) et est donc parfaitement positionnée pour la formation d'une paire d'ions entre l'atome de soufre et l'atome N $\epsilon_2$  de l'His176 distant de 3,32 Å.

#### *A.2.6.2. La molécule de cofacteur*

Dans l'ensemble des monomères, la molécule de cofacteur, modélisée avec un taux d'occupation de 1, est parfaitement définie dans les cartes de densité électronique (Figure III-9). La réaction avec le G3P étant incomplète dans le cristal, il existe donc un mélange entre les formes réduite (après réaction) et oxydée (n'ayant pas réagi) du cofacteur. Toutefois, les deux états redox du NAD ne peuvent être distingués à partir de la carte de densité électronique calculée à 1,78 Å de résolution et il n'existe aucune indication pour soutenir un quelconque changement de conformation de la molécule de cofacteur après l'étape du transfert d'hydrure.

#### *A.2.6.3. Les sites de reconnaissance anionique*

Un ion sulfate a été modélisé dans le site Ps avec un taux d'occupation de 1. Ses atomes d'oxygène forment 5 liaisons hydrogène avec les chaînes latérales des résidus Arg231, Arg195 et Thr179, avec la fonction hydroxyle en 2' du ribose adjacent au nicotinamide et avec une molécule d'eau (Figure III-8). Une seconde molécule de sulfate, en occupation partielle, localisée au niveau du site Pi (*cf.* ci-dessous), développe des interactions avec les chaînes latérales de la Thr208 et de la Ser148, avec l'atome d'azote de la chaîne principale du résidu Gly209 (avec la boucle 206-212 en conformation A). Cet anion présente un facteur B élevé (46,6 Å<sup>2</sup>) en comparaison à la valeur moyenne du modèle (24,5 Å<sup>2</sup>) et n'est pas très bien défini dans les cartes de densité électronique.



*Figure III-9 : Vue stéréoscopique de la molécule de cofacteur de la structure de l'intermédiaire thioacylenzyme de la BstGAPDH. La molécule de cofacteur de la sous-unité O est représentée dans une carte de densité électronique 2mFo-DFc contournée à 1,2 sigma. La réaction avec le G3P étant incomplète, le modèle contient un mélange entre les formes réduite et oxydée du NAD.*

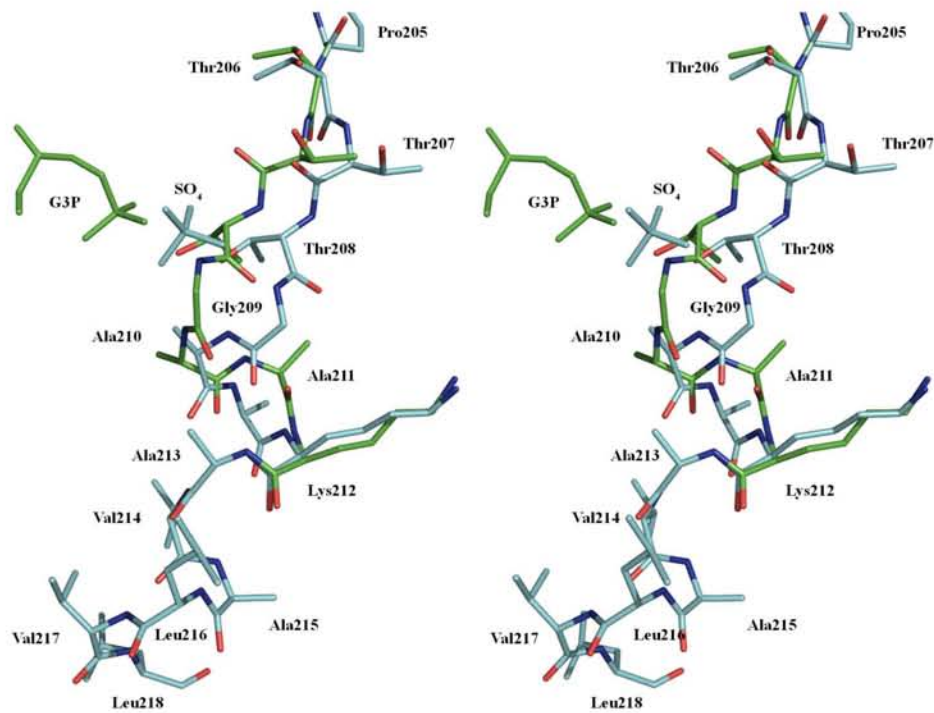
Cependant, une simple molécule d'eau ne suffit pas à expliquer la densité électronique observée et la position de cet anion correspond parfaitement à la position du site Pi observée dans la structure holoenzyme décrite par Skarzynski *et al.* (1987).

#### *A.2.6.4. Conformations alternatives de la boucle 206-212*

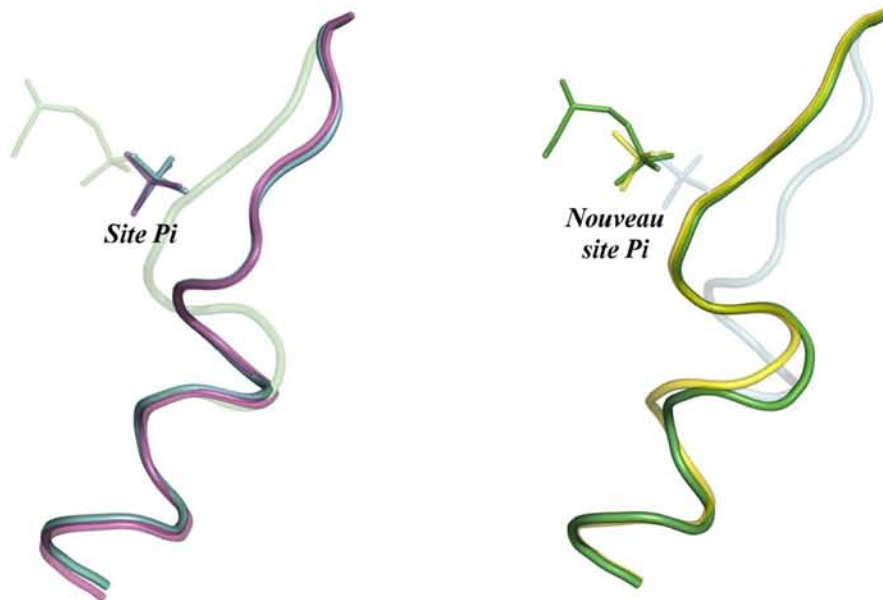
Délimitant l'un des côtés du site Pi, la boucle 206-212 contribue à la formation du site de reconnaissance anionique Pi. Comme évoqué précédemment, cette boucle a été construite sous deux conformations alternatives (Figure III-10), construites de façon non équivoque dans la densité électronique. La première conformation, notée A, est strictement équivalente à la conformation décrite dans les structures holo de GAPDH issue de *B. stearothermophilus* (Skarzynski *et al.*, 1987, Didierjean *et al.*, 2003) (Figure III-11) et participe à la fixation de l'anion sulfate dans le site Pi "classique". Avec un taux d'occupation complémentaire, la seconde conformation de la boucle (conformation B), présente un déplacement maximum de 3,4 Å au niveau du résidu Gly209, dirigé vers la cystéine catalytique. Dans cette conformation alternative, la boucle 206-212, est trop proche de l'anion sulfate (Pi) pour que les deux puissent coexister dans le modèle moléculaire (distance <2 Å),=ce qui confirme que la présence d'un anion sulfate au niveau du site Pi est associée à la conformation A de la boucle, précédemment décrite. La conformation B est, elle, très similaire à la conformation initialement décrite dans la structure de GAPDH de *Thermotoga maritima* (Korndörfer *et al.*, 1995) (Figure III-11) et participe à la formation du site de reconnaissance anionique appelé nouveau site Pi ("new Pi site") (Kim *et al.*, 1995). Les deux conformations présentent des facteurs B similaires, 20,6 et 22,2 Å<sup>2</sup> respectivement pour les conformations A et B. Ces valeurs sont comparables au facteur B moyen de la protéine (21,4 Å<sup>2</sup>) (Tableau III-5) suggérant, tout comme la relativement bonne définition des cartes de densité électronique dans cette région, que les deux conformations coexistent sous forme stable dans le cristal.

#### *A.2.7. Comparaison avec quelques structures représentatives des GAPDH*

La position du nouveau site Pi, observé ici pour la première fois dans la GAPDH de *B. stearothermophilus*, est équivalente à celle décrite pour les structures d'autres sources (Korndörfer *et al.*, 1995 ; Kim *et al.*, 1995 ; Yun *et al.*, 2000) (Figure III-11). La superposition des atomes de la chaîne principale de la sous-unité O de notre modèle avec ceux de la structure de l'intermédiaire apo-hémithioacétal (1dc4, chaîne A) (Figure III-12) positionne les deux atomes de phosphore (C3P)

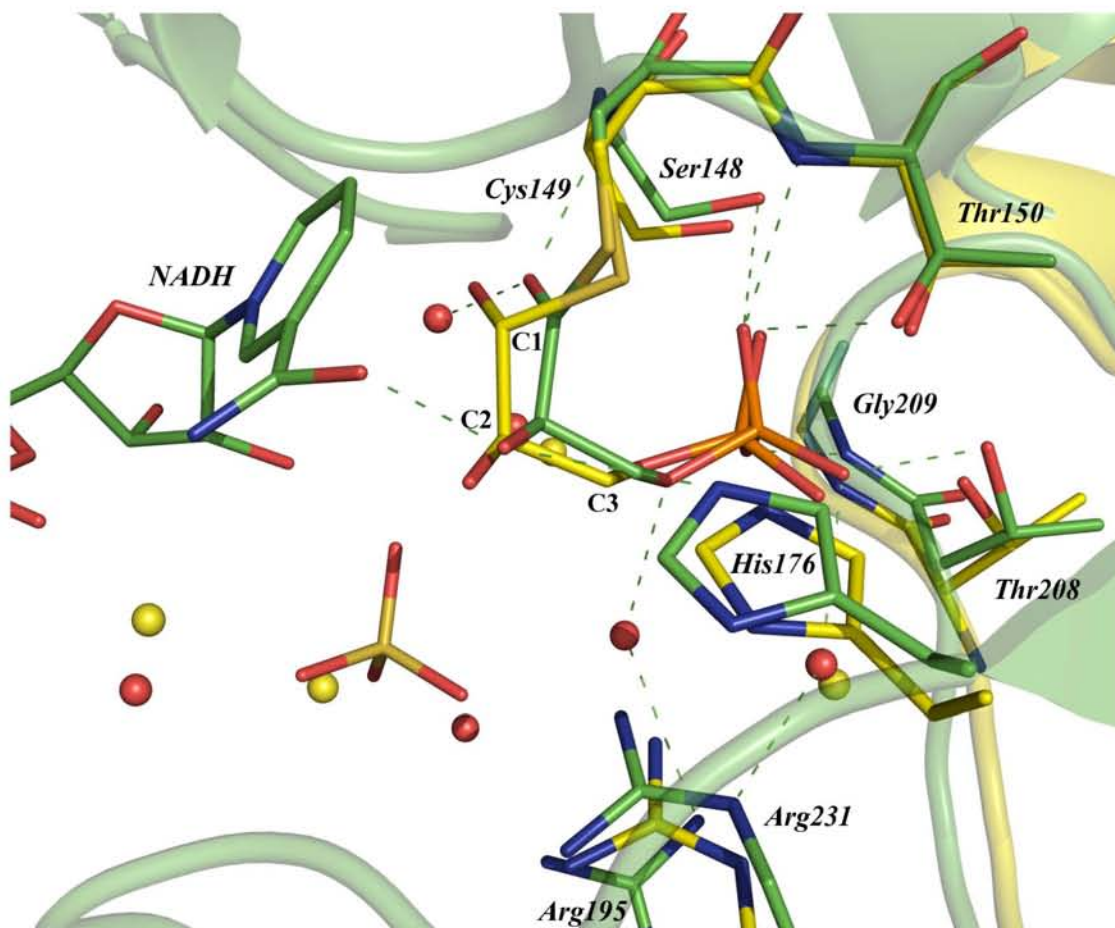


**Figure III-10 : Vue stéréoscopique des deux conformations alternatives de la boucle 206-212.** La région 205-218 de la sous-unité O est présentée avec les deux conformations alternatives de la boucle 206-212, la conformation A, en bleu, et la conformation B en vert. Sont représentés également les ligands (G3P et sulfate) associés à chacune des deux formes, et colorés en conséquence.



**Figure III-11 : Superposition de la région 205-218 avec des structures représentatives de GAPDH et Re-localisation du site Pi.** Les superpositions de notre modèle (en bleu clair ou vert respectivement pour les conformations A et B de la boucle 206-212) avec la région correspondante de la structure holoenzyme issue de *B. stearothermophilus* (en mauve, code pdb 1gd1) et de la structure de *Thermotoga maritima* (en jaune, code pdb 1hdg), sont représentées à gauche et à droite, respectivement. Pour faciliter la comparaison, les deux conformations de notre modèle sont toujours représentées avec, en transparence, la conformation non considérée.

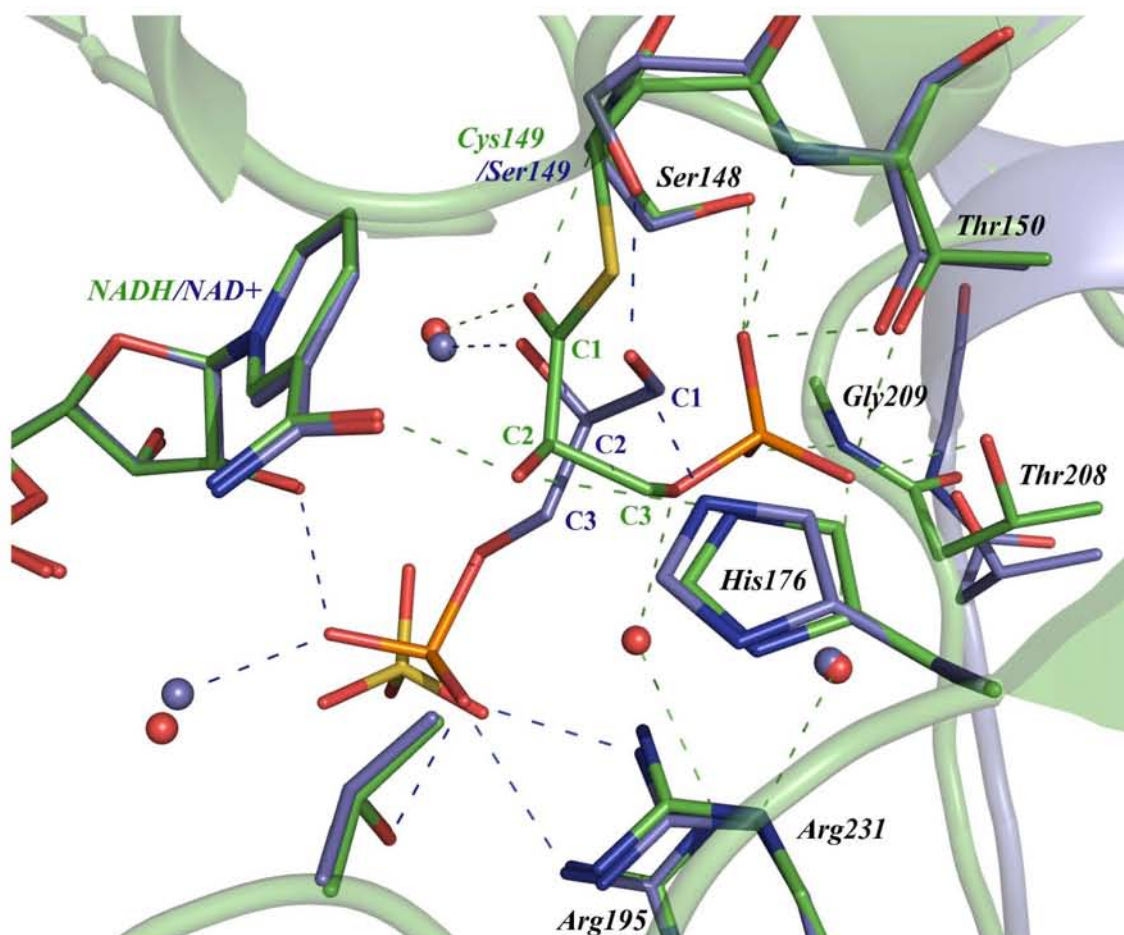




**Figure III-12 : Superposition des intermédiaires NADH - thioacylzyme et apo-hémithioacétal de la GAPDH.** Les modèles du thioacylzyme et de l'hémithioacétal (code pdb 1dc4 ; source E. coli) sont colorés en vert et jaune, respectivement. Le réseau de liaisons hydrogène de l'intermédiaire thioacylzyme est représenté par des pointillés. Les résidus impliqués dans la fixation du substrat et la chaîne carbonée du G3P sont annotés.

à seulement 0,33 Å l'un de l'autre. Les atomes centraux des anions sulfate et phosphate respectivement fixés dans le nouveau site Pi des structures de GAPDH issue de *T. maritima* et *Leishmania mexicana*, après une procédure de superposition similaire, sont situés à 0,17 et 0,42 Å de l'atome de phosphore de notre modèle.

La position et la géométrie du substrat observées dans l'intermédiaire thioacylenzyme sont très similaires à celles décrites par Yun *et al.* (2000) pour l'intermédiaire hémithioacétal (intermédiaire obtenu avec une enzyme sous forme apo, code pdb 1dc4) (Figure III-12). Seul un léger décalage est observé entre les deux molécules de G3P qui peut être expliqué, tout d'abord, par la différence d'hybridation des atomes de carbone C1 (*sp*<sup>3</sup> et *sp*<sup>2</sup> respectivement dans les états hémithioacétal et thioacylenzyme) mais aussi probablement par l'influence stérique de la molécule de cofacteur, absente de la structure de l'intermédiaire hémithioacétal. Comme précédemment décrit par Yun *et al.* (2000) pour l'état l'hémithioacétal, l'atome d'oxygène O1 pointe dans une direction opposée à la fonction hydroxyle O2 pour adopter une conformation *trans* (angle de torsion O1-C1-C2-O2) (Figure III-12). Il est important de préciser que dans cette conformation, l'atome O1 pointe dans une direction opposée à l'His176 dont le rôle a pourtant été clairement démontré dans le mécanisme catalytique (Soukri *et al.*, 1989 ; Talfournier *et al.*, 1998 ; *cf.* Partie B : Discussion). Lorsque l'on compare la structure du thioacylenzyme de la GAPDH au complexe de type michaelien de la GAPDH de *B. stearothermophilus* déterminé par Didierjean *et al.* (2003) (1nqo) (Figure III-13), la principale différence concerne la position du groupement C3P du substrat qui, dans le complexe non covalent, est localisé dans le site Ps, à plus de 6 Å du nouveau site Pi. Ce changement de position du groupement C3P est également accompagné d'une réorganisation complète du réseau d'interactions formé entre le substrat et la protéine. Ainsi les orientations des liaisons C1-O1 et C2-O2 diffèrent substantiellement entre les deux structures. Alors que l'atome O1 pointe en direction de l'histidine catalytique et que la liaison C2-O2 est quasiment perpendiculaire au plan du nicotinamide dans le complexe michaelien, l'atome O1 pointe dans une direction opposée à l'His176 et la liaison C2-O2 est pratiquement parallèle au plan du nicotinamide dans le thioacylenzyme (Figure III-13).



**Figure III-13 : Superposition de l'intermédiaire thioacylzyme - NADH et du complexe michaelien de la GAPDH de *B. stearothermophilus*.** Les modèles du thioacylzyme et du complexe michaelien (code pdb Inqo) sont colorés en vert et bleu, respectivement. Les deux réseaux de liaisons hydrogène impliquant le G3P ont été représentés par des pointillés de la couleur du modèle respectif. Les résidus impliqués dans la fixation du substrat et la chaîne carbonée du G3P sont annotés.

## ***A.3. Structure de complexes entre la GAPDH de *B. stearothermophilus*, le NAD<sup>+</sup> et le 1,3-DPG***

### ***A.3.1. Stratégie***

La démarche utilisée pour obtenir les structures d'un complexe ternaire GAPDH-NAD<sup>+</sup>-1,3DPG est la même que celle développée par Didierjean *et al.* (2003), à savoir, la substitution de la cystéine catalytique par un résidu alanine ou sérine, ces deux mutations entraînant, respectivement, une perte totale ou une réduction très forte de l'activité de l'enzyme ( $k_{\text{cat}}$  de  $8 \cdot 10^{-3} \text{ s}^{-1}$  contre  $88 \text{ s}^{-1}$  pour l'enzyme sauvage, Boschi-Muller and Branlant, 1999). Les mutants C149A et C149S de la *Bst*GAPDH ont été cocrystallisés avec du NAD<sup>+</sup> (2mM) dans la condition A1 (PEG 4000) établie pour les essais d'obtention de l'intermédiaire thioacylenzyme (*cf.* Tableau II-1). Les cristaux ont ensuite été incubés en présence de 1 mM de 1,3-DPG puis congelés dans l'azote liquide avant la collecte des données de diffraction. Cette concentration en substrat, limitée principalement pour des raisons techniques, est néanmoins suffisante en théorie pour obtenir la saturation de l'enzyme au regard de l'affinité particulièrement élevé de la GAPDH pour ce substrat ( $K_M$  de 16  $\mu\text{M}$  pour l'enzyme sauvage, Michels *et al.*, 1996). Deux jeux de données sont présentés dans ce manuscrit avec les structures correspondantes. Il s'agit des structures obtenues à partir de deux cristaux des mutants C149S et C149A incubés en présence de 1,3-DPG pendant respectivement 10 et 8 min. Les statistiques de collecte de données et d'affinement des structures sont regroupées dans le Tableau III-7.

### ***A.3.2. Structure des complexes ternaires C149S/A-1,3DPG-NAD<sup>+</sup>.***

#### ***A.3.2.1. Présentation des structures et qualité des modèles***

##### ***A.3.2.1.1. Structure du mutant C149S***

La structure, obtenue à 1,66 Å de résolution, a été résolue par un affinement en corps rigide à partir de la structure du mutant C149S de la *Bst*GAPDH publiée par Didierjean *et al.* (2003).

**Tableau III-7 : Statistiques des jeux de données de diffraction et d'affinement des structures de complexes ternaires entre la BstGAPDH, le NAD<sup>+</sup> et le 1,3-DPG.**

Données de diffraction	C149S	C149A
Groupe d'espace	<i>C</i> 2	<i>P</i> 2 <sub>1</sub>
Maille (Å)	<i>a</i> = 140,0 ; <i>b</i> = 87,8 ; <i>c</i> = 119,1 ; $\beta$ = 119,0°	<i>a</i> = 121,2 ; <i>b</i> = 89,5 ; <i>c</i> = 266,9 ; $\beta$ = 102,2°
Z	4	8
Limite de résolution (Å)	52,3 - 1,66	45,0 - 2,14
Couronne de plus haute résolution (Å)	1,75 - 1,66	2,26 - 2,14
Température (K)	100	100
Réflexions uniques	136492 (14492)	278076 (27380)
Complétude (%)	91,7 (66,8)	90,7 (61,8)
Redondance	3,7 (3,4)	3,9 (4,0)
R-merge <sup>a</sup> (%)	4,7 (21,4)	7,0 (27,3)
I/ $\sigma$ (I)	17,9 (6,1)	14,5 (4,8)
Affinement		
R <sup>b</sup> (%)	18,3	18,6
R <sub>free</sub> <sup>c</sup> (%)	23,0	25,2
<b>r.m.s.d.<sup>d</sup> par rapport à la géométrie standard</b>		
Longueur de liaisons (Å)	0,012	0,011
Angles de valence (°)	1,50	1,40
chiralité (°)	0,09	0,09
<b>Facteur B moyens (Å<sup>2</sup>)</b>		
Modèle	20,5	34,7
Atomes de la protéine	18,75	34,9
NAD	13,0	27,1
1,3-DPG	20,5	32,6
Molécules d'eau	32,6	34,4

<sup>a</sup> Facteur R pour les intensités reliées par symétrie.

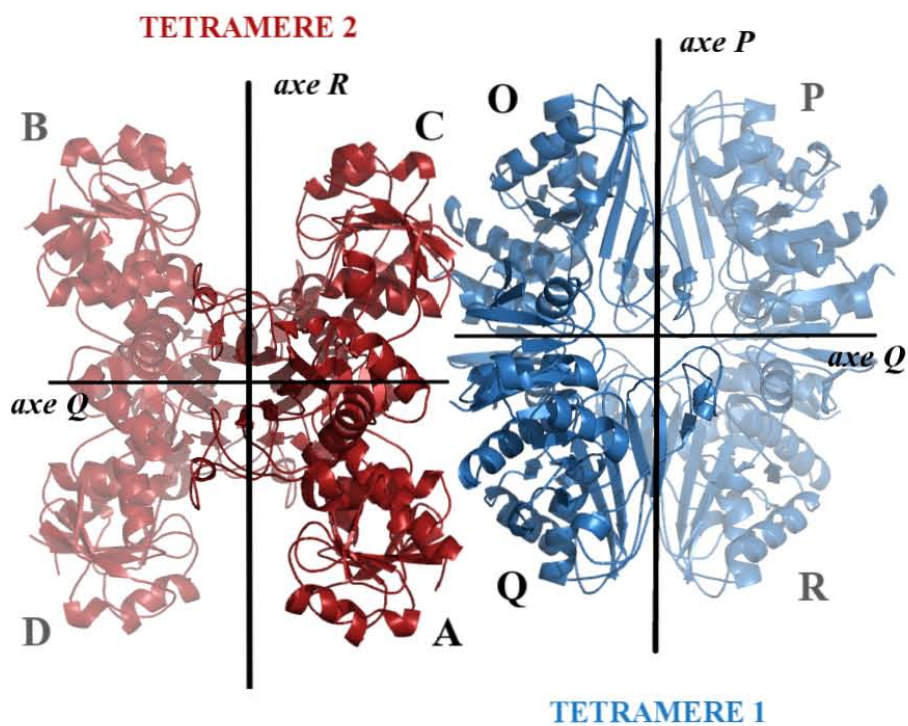
<sup>b</sup> Facteur R cristallographique. Facteur R cristallographique calculé sur 5 % des réflexions choisies aléatoirement et non incluses dans l'affinement.

<sup>d</sup> Écart quadratique moyen.

L'unité asymétrique ( $a = 140,0 \text{ \AA}$  ;  $b = 87,8 \text{ \AA}$  ;  $c = 119,1 \text{ \AA}$  ;  $\alpha = 90,0^\circ$  ;  $\beta = 119,0^\circ$  ;  $\gamma = 90,0^\circ$ ) contient 2 dimères indépendants, appartenant chacun à un tétramère physiologique distinct. Le premier dimère contient les sous-unités O et Q et le second les sous-unités A et C. Deux tétramères physiologiques OPQR et ABCD sont générés grâce à l'axe binaire cristallographique confondu à l'un des axes de symétrie interne de chacun des tétramères de GAPDH (Figure III-14). Le modèle comporte également 4 molécules de cofacteur  $\text{NAD}^+$ , 4 molécules de 1,3-diphosphoglycérate (1,3-DPG), 5 ions sodium et près de 1500 molécules d'eau. Tout comme dans le cas de l'intermédiaire thioacylenzyme, la boucle 206-212 a été modélisée sous deux conformations alternatives (*cf.* ci-dessous). L'affinement par le logiciel Refmac5 a convergé vers des valeurs de R et  $R_{\text{free}}$ , respectivement de 18,3 et 23,0%. A part quelques chaînes latérales exposées au solvant, le modèle protéique est parfaitement défini dans les cartes de densité électronique à 1,66 Å de résolution et présente des propriétés géométriques satisfaisantes (Tableau III-7). L'étude du diagramme de Ramachandran montre qu'en plus des 2 résidus connus pour être en dehors des zones acceptables dans l'ensemble des structures de GAPDH connues à ce jour (Asp186 et Val237), la Lys212 du monomère C est également à l'extérieur des zones admises. Ce résidu appartient à la boucle qui a été modélisée sous 2 conformations alternatives et qui est moins bien définie dans les cartes de densité électronique que le reste du modèle.

#### ***A.3.2.1.2. Structure du mutant C149A***

La structure du complexe ternaire C149A-1,3DPG- $\text{NAD}^+$  présentée est issue d'un jeu de données de diffraction à 2,14 Å de résolution (Tableau III-7). La structure a été obtenue par remplacement moléculaire en utilisant comme modèle l'une des structures monocliniques à 4 tétramères de l'enzyme de type sauvage déterminée précédemment. En effet, comme dans le cas de l'incubation des cristaux PEG en présence de G3P, l'incubation en présence de 1,3-DPG a entraîné une réduction de symétrie de  $C2$  vers  $P2_1$  (*cf.* en A.2.4.1). L'unité asymétrique du modèle contient 4 tétramères de GAPDH (16 monomères) dénommés OPQR, ABCD, EFGH et IJKL, 16 molécules de cofacteur, 16 molécules de 1,3-diphosphoglycérate (1,3-DPG) et plus de 2800 molécules d'eau. L'affinement a été conduit avec le logiciel Refmac5 jusqu'à la convergence du modèle à des valeurs de R et de  $R_{\text{free}}$  de 18,6 et 25,2%, respectivement (Tableau III-7). Si le modèle présente une géométrie acceptable (Tableau III-7), l'étude du diagramme de Ramachandran met en évidence, 7 résidus non-glycine en dehors des régions non autorisées (en plus des résidus Asp186 et Val237). Il s'agit notamment des résidus de la boucle 206-212, à nouveau en double conformation, de différentes sous-unités. Toutefois ces déviations ne sont que très légères et n'ont à priori aucune



**Figure III-14 :** Contenu de l'unité asymétrique de la forme monoclinique C 2. Les deux dimères indépendants, OQ et AC, composant l'unité asymétrique, sont représentés en bleu et en rouge, respectivement. Les dimères PR et BD, symétriquement reliés aux précédents par les opérateurs  $(-x, y, 1-z)$  et  $(2-x, y, 1-z)$ , respectivement, sont représentés en transparence et permettent d'obtenir les tétramères physiologiques 1 et 2. Dans cette figure, l'axe cristallographique d'ordre 2 est vertical, parallèle à l'axe P du tétramère 1 et à l'axe R du tétramère 2.

influence sur les interprétations qui seront faites à partir du modèle.

La qualité du modèle est inégale en fonction de la région considérée. En effet, les cartes de densité électronique des monomères J, K et L notamment, sont d'assez mauvaise qualité comparées au reste du modèle. La comparaison des valeurs moyennes de facteurs B par monomère et par tétramère (Tableau III-8), révèle qu'en fait, deux des tétramères (ABCD et surtout IJKL) sont globalement plus "agités" que le reste de la structure. Il est néanmoins plus vraisemblable que cette observation soit le résultat d'un désordre statique interne au cristal et directement lié à la réduction de symétrie des cristaux PEG lors de leur incubation en présence de substrats plutôt que d'une plus grande mobilité de ces régions. L'analyse du modèle a donc été basée principalement sur les tétramères OPQR et EFGH qui sont globalement de meilleure qualité que les deux autres.

### *A.3.2.2. Comparaison des structures ternaires et quaternaires des complexes GAPDH-1,3DPG-NAD<sup>+</sup>.*

#### *A.3.2.2.1. Structure du mutant C149S*

Comme l'indiquent les faibles valeurs de rmsd résultant de la superposition des atomes de Ca des 4 sous-unités considérées deux à deux (de 0,20 à 0,38 Å) (Tableau III-9a), le repliement global des quatre monomères de l'unité asymétrique est équivalent. Lorsque l'on superpose les quatre sous-unités O, Q, A et C au modèle précédemment publié par Didierjean *et al.* (2003) (code pdb 1nqo) qui présente une unité asymétrique équivalente, la valeur de rmsd sur les 1336 carbones  $\alpha$  est de 0,39 Å. Les superpositions des tétramères physiologiques 1 et 2, reconstitués grâce aux symétries cristallographiques, sur le tétramère holoenzyme (1gd1, Skarzynski *et al.*, 1987) donnent un rmsd moyen de 0,72 Å qui suggère que la fixation de la molécule de 1,3-DPG sur l'holoenzyme n'entraîne pas de changements majeurs de la structure.

#### *A.3.2.2.2. Structure du mutant C149A*

La superposition des atomes de Ca des 16 sous-unités composant le modèle conduit à une valeur de rmsd moyenne de 0,31 Å et ce, en l'absence de toute contrainte de symétrie lors de l'affinement. Les faibles valeurs de rmsd obtenues lorsque les tétramères et monomères du modèle sont superposés sur la structure holoenzyme 1gd1 (Tableau III-9b) amènent aux mêmes types de commentaires que pour la structure du mutant C149S.



**Tableau III-8 : Analyse des valeurs moyennes des facteurs B par monomère et par tétramère du complexe ternaire C149A-1,3DPG-NAD<sup>+</sup>.**

	O	P	Q	R	A	B	C	D
B moyen (Å <sup>2</sup> )	23,5	33,2	24,4	25,9	40,4	37,6	36,2	35,1
	26,7				37,3			
	E	F	G	H	I	J	K	L
B moyen (Å <sup>2</sup> )	21,4	27,8	31,6	29,4	38,4	59,3	53,1	41,1
	27,5				48,0			

**Tableau III-9a : Valeurs de rmsd (Å) résultant de la superposition des carbonés α des différentes sous-unités du modèle C149S-1,3DPG-NAD<sup>+</sup> deux à deux.**

	Q	A	C
O	0,25	0,29	0,31
Q		0,33	0,20
A			0,38

**Tableau III-9b : Valeurs de rmsd (Å) résultant de la superposition des carbonés α des différentes sous-unités du modèle C149A-1,3DPG-NAD<sup>+</sup> deux à deux.**

	P	Q	R	A	B	C	D	E	F	G	H	I	J	K	L
O	0,19	0,18	0,18	0,18	0,21	0,20	0,18	0,18	0,17	0,25	0,21	0,22	0,32	0,29	0,25
P		0,17	0,17	0,20	0,20	0,17	0,21	0,20	0,19	0,26	0,19	0,25	0,32	0,28	0,29
Q			0,17	0,20	0,21	0,16	0,19	0,19	0,19	0,24	0,18	0,24	0,31	0,30	0,28
R				0,19	0,21	0,19	0,20	0,20	0,18	0,25	0,20	0,24	0,31	0,30	0,30
A					0,24	0,21	0,21	0,21	0,20	0,27	0,22	0,23	0,31	0,30	0,27
B						0,22	0,25	0,24	0,24	0,33	0,23	0,28	0,35	0,35	0,33
C							0,22	0,19	0,19	0,24	0,18	0,26	0,31	0,28	0,30
D								0,23	0,21	0,27	0,23	0,25	0,32	0,31	0,26
E									0,19	0,24	0,18	0,22	0,31	0,31	0,29
F										0,23	0,20	0,24	0,30	0,29	0,28
G											0,25	0,30	0,32	0,29	0,29
H												0,26	0,33	0,29	0,30
I													0,32	0,35	0,26
J														0,38	0,36
K															0,36

La superposition des 2 tétramères physiologiques reconstitués par symétrie à partir de la structure du mutant C149S avec les 2 tétramères bien définis de l'unité asymétrique de la structure du mutant C149A conduit à un rmsd moyen de 0,52 Å et suggère qu'aucune différence majeure n'existe entre les structures de ces deux mutants. Ainsi, c'est la structure du mutant C149S qui sera principalement détaillée dans la suite de cette partie résultats.

### *A.3.2.3. Conformation du complexe ternaire C149S/A-1,3DPG-NAD<sup>+</sup>.*

#### *A.3.2.3.1. La molécule de cofacteur*

Comme précédemment évoqué, le site actif de chacune des sous-unités des modèles C149S et C149A, contient une molécule de NAD<sup>+</sup>. La molécule de cofacteur qui a été modélisée avec un facteur d'occupation de 1 est parfaitement définie dans la densité électronique (Figure III-15) et présente un facteur B moyen de 12,9 Å<sup>2</sup> dans la structure du mutant C149S. Bien que les atomes de la molécule de cofacteur de la sous-unité A présentent des facteurs B (18,5 Å<sup>2</sup>) significativement plus élevés que dans les autres monomères, cette valeur reste cependant comparable au facteur B moyen du modèle (20,4 Å<sup>2</sup>). La conformation et le réseau de liaisons hydrogène de la molécule de cofacteur sont strictement équivalents à ceux précédemment décrits pour les holo-structures de GAPDH issue de *B. stearothermophilus* (Skarzynski *et al.*, 1987).

#### *A.3.2.3.2. La molécule de 1,3-DPG*

La présence d'un fort signal de densité électronique au niveau du site actif de chacune des sous- unités a été attribuée à la présence de 1,3-DPG. Dans tous les cas, l'occupation de cette molécule a été estimée comme étant inférieure à 1 (Tableau III-10). Deux conformations du ligand (notées 1 et 2 et décrites ci-dessous) ont été modélisées dans les sites actifs de la structure du mutant C149S (Figure III-16). Les facteurs d'occupation du 1,3-DPG, estimés en se basant sur les cartes Fourier différence et sur les facteurs d'agitation thermique, varient entre 0,25 à 0,65 en fonction de la conformation et de la sous-unité considérée (Tableau III-10). Les taux d'occupation moyens des conformations 1 et 2 sont respectivement de 0,42 et 0,32 ce qui donne un taux de “remplissage” moyen de 0,75 pour le 1,3-DPG. Dans le cas du mutant C149A dont le taux de “remplissage” moyen en 1,3-DPG a été estimé à seulement 0,4, seule la conformation 1 a été modélisée car celle-ci est également majoritaire. Des affinements approfondis avec Buster-TNT (Clemens Vornhein, en

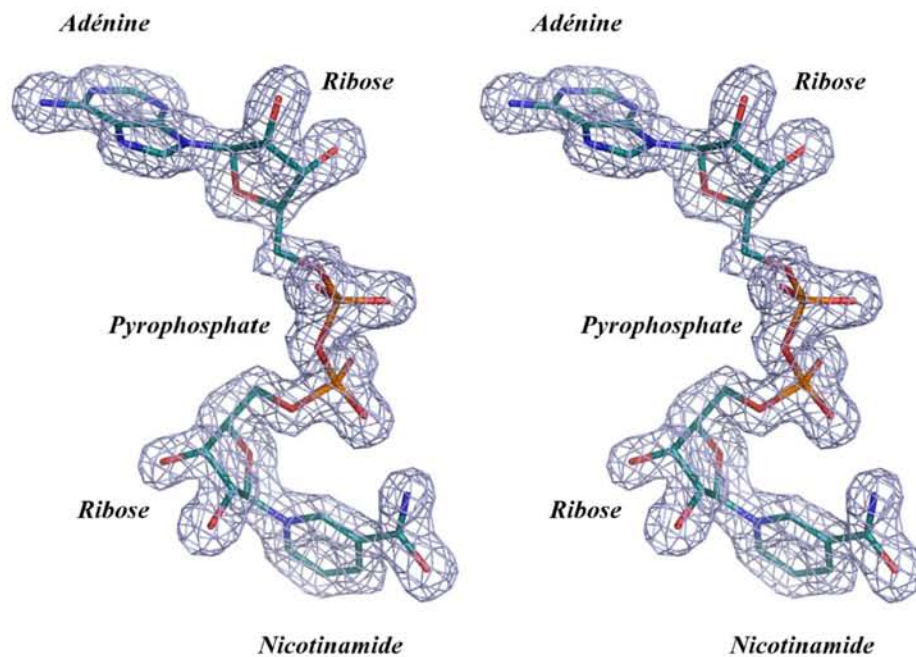


Figure III-15 : Vue stéréoscopique de la molécule de cofacteur  $NAD^+$  de la sous-unité O dans sa carte de densité électronique de type  $2mFo-DFc$  contournée à  $1,2\sigma$ .

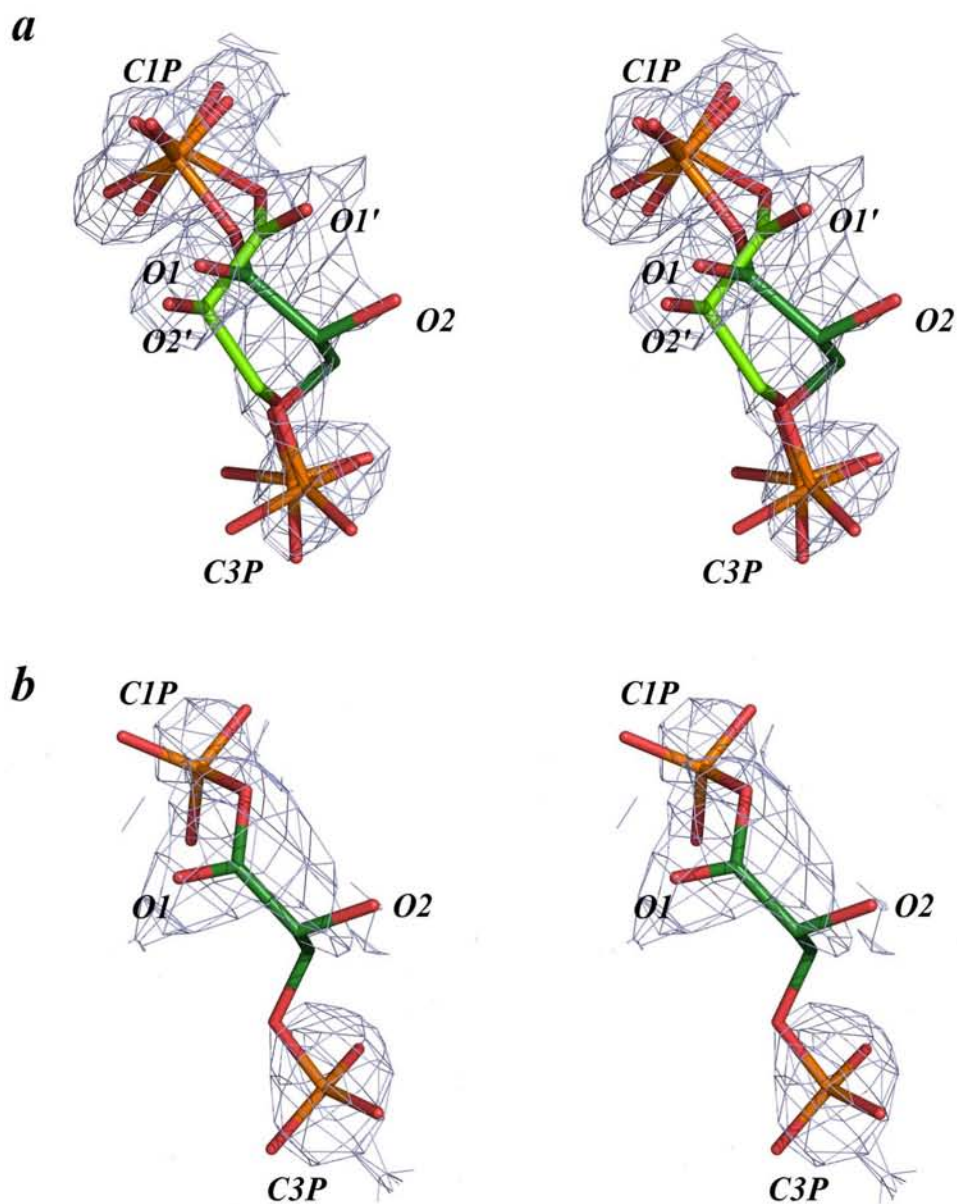
Tableau III-10 : Paramètres cristallographiques de la molécule de 1,3-DPG dans les deux complexes ternaires ( $C149S/A-1,3DPG-NAD^+$ )

**Mutant C149S**

Sous- unités	O		Q		A		C	
	Conf1	Conf2	Conf1	Conf2	Conf1	Conf2	Conf1	Conf2
Conformation	0,35	0,30	0,35	0,40	0,35	0,25	0,65	0,35
Taux d'occupation	20,3	18,2	17,8	18,2	21,8	20,0	24,2	23,5
Facteurs B	0,65		0,75		0,60		1	

**Mutant C149A (taux d'occupation de la conformation 1 = 0,4)**

Sous-unités	O	P	Q	R	A	B	C	D
Facteurs B	21,4	31,8	28,7	25,5	39,3	48,0	32,6	31,4
Sous-unités	E	F	G	H	I	J	K	L
Facteurs B	20,2	24,1	27,3	27,2	33,8	45,7	47,5	36,5



**Figure III-16 : Vue stéréoscopique de la molécule de 1,3-DPG dans des cartes de densité électronique "omit".** La molécule de 1,3-DPG a été supprimée du modèle final qui a ensuite été soumis à un affinement avec Refmac5 avant le calcul de la nouvelle carte de densité de type  $2mFo-DFc$ . Les groupements phosphate C1P et C3P et les atomes O1 et O2 sont annotés. **(a)** Les conformations 1 (vert foncé) et 2 (vert clair) du monomère Q de la structure du mutant C149S sont représentées dans une carte omit contourée à 0,8 sigma. **(b)** La conformation 1, seule modélisée, du monomère Q de la structure du mutant C149A est représentée dans une carte omit contourée à 0,8 sigma.

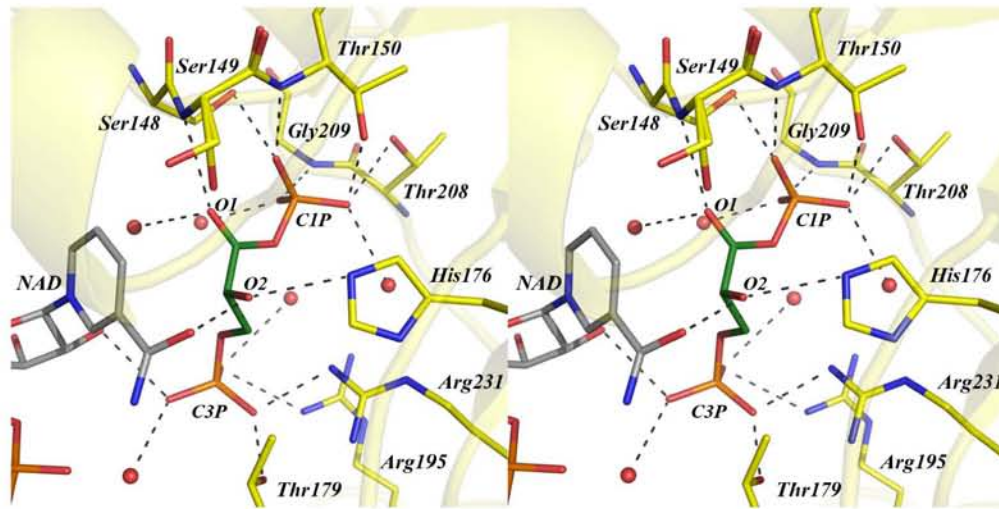
cours de réalisation) devraient permettre de donner de meilleures estimations des taux d'occupation.

Dans la structure des deux mutants, la qualité des cartes de densité électronique ne laisse aucune ambiguïté quant à la position des groupements phosphate du ligand, C1P et C3P (Figure III-16). Quelle que soit la conformation du ligand (1 ou 2), les groupements C1P et C3P du 1,3-DPG sont localisés respectivement dans le nouveau site Pi et dans le site Ps. Le groupement phosphate en C1 (C1P) forme des interactions avec les résidus Ser148, Thr150, Thr208 et Gly209 (Figures III-17 et III-18) et le groupement C3P développe un réseau de liaisons hydrogène avec les résidus Thr179, Arg195 et Arg231 et avec la molécule de cofacteur (partie ribose du nicotinamide) (Figures III-17 et III-18).

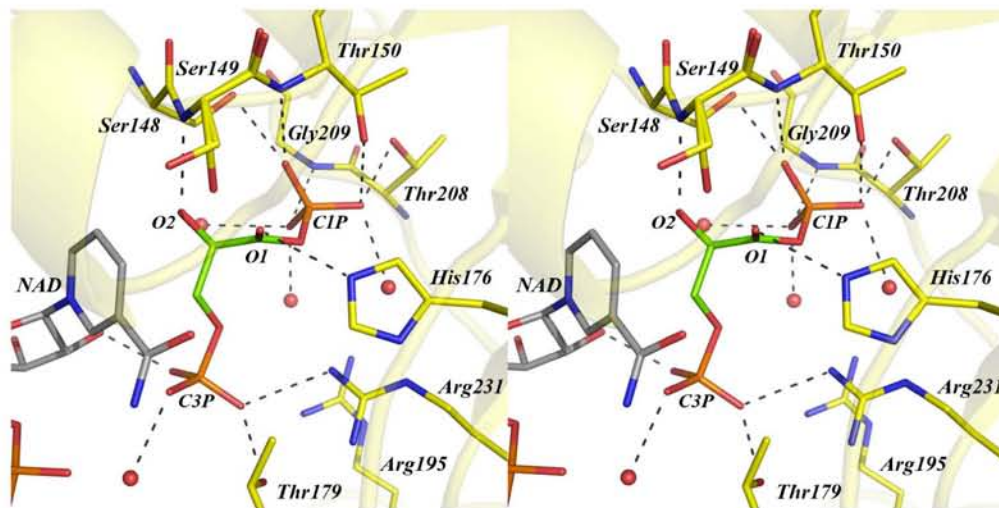
Les conformations 1 et 2 se différencient principalement au travers des positions et orientations des atomes O1 et O2 portés par la chaîne carbonée du 1,3-DPG (Figure III-16). Le passage d'une conformation à l'autre implique principalement une rotation d'environ 100° autour de la liaison C1-O11 (O11 étant l'atome d'oxygène du groupement C1P lié à l'atome C1), d'environ 120° autour de la liaison C1-C2 et d'environ 130° autour de la liaison O3-P. Les atomes O1 et O2 sont, eux aussi, stabilisés par des liaisons hydrogène quelle que soit la conformation considérée. Dans la conformation 1, l'atome O1 interagit avec la chaîne principale du résidu Ser149(N) et l'atome O2 avec le résidu His176(N $\epsilon_2$ ) et la fonction carboxamide du nicotinamide (Figure III-17). Dans la conformation 2, l'atome O1 interagit avec l'His176(N $\epsilon_2$ ) et l'atome O2 avec la chaîne principale de la Ser149(N) (Figure III-18). Au total, la molécule de 1,3-DPG développe près de vingt liaisons hydrogène avec les résidus protéiques, la molécule de cofacteur ainsi qu'avec 4 à 5 molécules d'eau selon la sous-unité considérée. Les distances interatomiques des liaisons hydrogène formées par les deux conformations du ligand dans la sous-unité Q sont indiquées dans le Tableau III-11.

La conformation 1, qui présente une conformation décalée des atomes O1 et O2, semble plus favorable énergétiquement que la conformation 2. Des calculs d'énergie libre réalisés avec Gaussian 03 (DFT, B3LYP, 6-311G++(2d,p)) semblent confirmer ce point et donnent une différence d'énergie de 11,8 kcal.mol<sup>-1</sup> en faveur de la conformation 1.

Le résidu Ser149 adopte deux conformations distinctes (Figures III-17 et III-18), clairement mises en évidence par les cartes de densité électronique. La conformation A, minoritaire (Tableau III-12), est équivalente à la conformation classiquement observée pour la cystéine catalytique. C'est également la conformation de la Ser149 rencontrée dans les structures précédentes du mutant C149S de la *Bst*GAPDH. La conformation B, majoritaire, (Tableau III-12) n'a jamais été décrite et tend à faire pointer la chaîne latérale vers l'arrière du site actif par une rotation autour de la rotule  $\chi_1$ .



**Figure III-17 :** *Vue stéréoscopique du réseau de liaisons hydrogène formé entre la protéine et la molécule de 1,3-DPG en conformation 1 (mutant C149S, monomère Q). Le 1,3-DPG, le NAD ainsi que les résidus directement impliqués dans la fixation du 1,3-DPG sont représentés sous forme de bâtons et les molécules d'eau en interaction directe avec le 1,3-DPG sous forme de sphères. Les liaisons hydrogènes impliquant les atomes du ligand sont représentées en pointillés. Les groupements C1P, C3P et les atomes O1 et O2 du ligand sont annotés.*



**Figure III-18 :** *Vue stéréoscopique du réseau de liaisons hydrogène formé entre la protéine et la molécule de 1,3-DPG en conformation 2 (mutant C149S, monomère Q). Le 1,3-DPG, le NAD ainsi que les résidus directement impliqués dans la fixation du 1,3-DPG sont représentés sous forme de bâtons et les molécules d'eau en interaction directe avec le 1,3-DPG sous forme de sphères. Les liaisons hydrogènes impliquant les atomes du ligand sont représentées en pointillés. Les groupements C1P, C3P et les atomes O1 et O2 du ligand sont annotés.*

**Tableau III-11 : Interactions de type liaisons hydrogène entre la molécule de 1,3-DPG et les résidus du site actif de la sous-unité Q de la structure du mutant C149S (les distances entre atomes sont indiquées en Å). La colonne de gauche indique les atomes du 1,3-DPG et les atomes de la protéine engagés dans des liaisons.**

Conformation 1		Conformation 2	
O1 --- N-Ser149	2,8	O1 --- N $\epsilon_2$ -His176	2,8
O1 --- H <sub>2</sub> O	3,2		
O12 --- N-Gly209	2,7	O12 --- N-Gly209	2,9
O12 --- H <sub>2</sub> O	2,7	O12 --- H <sub>2</sub> O	2,5 / 2,9
O13 --- O $\gamma_1$ -Thr150	2,6	O13 --- O $\gamma_1$ -Thr150	2,8
O13 --- O $\gamma_1$ -Thr208	3,1	O13 --- O $\gamma_1$ -Thr208	2,6
O13 --- H <sub>2</sub> O	2,9	O13 --- H <sub>2</sub> O	2,6
O14 --- N-Thr150	2,6	O14 --- N-Thr150	3,1
O14 --- O $\gamma$ -Ser148	2,8	O14 --- O $\gamma$ -Ser148	2,7
O2 --- NO7-NAD	2,5	O2 --- N-Ser149	2,7
O2 --- N $\epsilon_2$ -His176	2,8		
O2P --- N $\theta_2$ -Arg231	2,7	O2P --- N $\theta_2$ -Arg231	2,9
O2P --- O $\gamma_1$ -Thr179	2,4	O2P --- O $\gamma_1$ -Thr179	2,6
O3P --- NO2*-NAD	3,0	O3P --- H <sub>2</sub> O	2,7
O3P --- H <sub>2</sub> O	2,8		
O4P --- N $\theta_1$ -Arg195	3,1	O4P --- NO2*-NAD	2,5
O4P --- H <sub>2</sub> O	3,1		

**Tableau III-12 : Facteurs d'occupation et d'agitation thermique moyens (facteurs B en Å<sup>2</sup>) du résidu Ser149 et des résidus de la boucle 206-212 de la structure du mutant C149S.**

		O	Q	A	C
<b>Ser149 (confA/B)</b>	Taux d'occupation	0,5/0,5	0,35/0,65	0,45/0,55	0,35/0,65
	Facteur B moyen (Å)	10,0/9,3	12,3/12,6	18,9/18,6	13,2/13,0
<b>Boucle 206-212* (confA/B)</b>	Segment en double conformation	206-210	207-212	206-213	210-213*
	Taux d'occupation	0,6/0,4	0,7/0,3	0,7/0,3	0,4/0,6*
	Facteur B moyen (Å)	16,8/12,6	22,1/23,0	26,0/24,2	23,7/22,8*

\* La région 206-209 du monomère C présente une seule conformation, équivalente à la conformation A délimitant le nouveau site Pi dans les autres sous-unités.

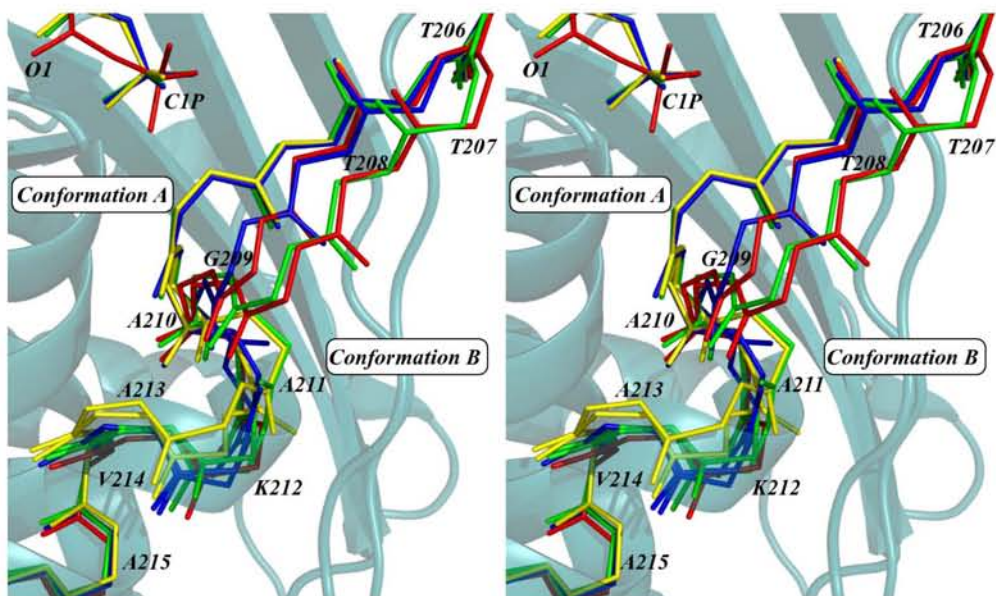
Dans cette conformation originale, sa chaîne latérale forme une liaison hydrogène avec la fonction carbonyle de la chaîne principale du résidu Asn313 située à une distance moyenne de 2,7 Å. En se basant sur les valeurs d'occupation estimées et sur l'observation des structures, il semblerait que la présence de la molécule de 1,3-DPG soit à l'origine du changement de conformation de la Ser149 en induisant une certaine gêne stérique.

#### A.3.2.4. Conformation de la boucle 206-212

La région 206-212 adopte une double conformation dans chacun des 4 monomères de la structure du mutant C149S (Tableau III-12). Les taux d'occupation ont, jusqu'à présent, été estimés en se basant à la fois sur l'observation des cartes de densité électronique et sur la comparaison de la valeur des facteurs B. Trois conformations-types de cette boucle peuvent être distinguées lors de la superposition des différentes sous-unités (Figure III-19). La conformation notée B dans les monomères O et A, est comparable à la conformation classique observée dans les structures précédentes de la *Bst*GAPDH (Skarzynski *et al.*, 1987 ; Didierjean *et al.*, 2003). La conformation notée A dans les monomères Q, A et C, est similaire à la nouvelle conformation observée dans les GAPDH de *T. maritima* et *L. mexicana* et pour la première fois, dans le cas de la GAPDH de *B. stearothermophilus*, pour la structure de la forme thioacylenzyme. Par rapport à la conformation classique (B), cette conformation présente un déplacement vers le résidu catalytique qui est maximal au niveau du résidu Gly209 (plus de 3,6 Å) et qui implique une déstabilisation du premier tour de l'hélice 211-217. Dans cette conformation, la boucle participe à la formation du nouveau site Pi au travers des résidus Thr208 et Gly209 qui interagissent avec le groupement C1P du 1,3-DPG. On note l'existence, dans les sous-unités O et Q, d'une conformation intermédiaire de la boucle. Le monomère O est le seul dans lequel la boucle est observée dans des conformations classique et intermédiaire mais pas "complètement déplacée" vers le résidu catalytique. Cette observation est vraisemblablement directement liée à la position particulière du groupement C1P du 1,3-DPG dans cette sous-unité par rapport aux autres dans lesquelles la boucle est majoritairement sous forme A (ou "complètement déplacée").

La structure du mutant C149A montre également des doubles conformations de la région 206-212. Toutefois, seules 7 sous-unités (O, Q, R, D, E, F et I) sur les 16 composant l'unité asymétrique du modèle présentent une double conformation. Dans les autres sous-unités, seule la conformation B, classique, est observée. Il est vraisemblable que la fixation du 1,3-DPG dans le site actif de l'enzyme soit à l'origine de la présence de la conformation A de la boucle 206-212 qui vient participer à la formation du nouveau site Pi. Le fait de ne pas toujours observer cette conformation





**Figure III-19 : Vue stéréoscopique de la superposition de la région 206-215 des différentes sous-unités du complexe ternaire C149S-1,3DPG-NAD<sup>+</sup>. Les sous-unités Q, A et C ont été superposées sur la sous-unité O. La chaîne principale du segment T206-A215 est représentée en bâtons et colorée en rouge, bleu, vert et jaune respectivement pour les sous-unités O, Q, A et C. Les deux conformations alternatives sont notées A et B. En haut à gauche apparaissent les molécules de 1,3-DPG, colorées de la même manière que le monomère dans lequel elles sont fixées.**

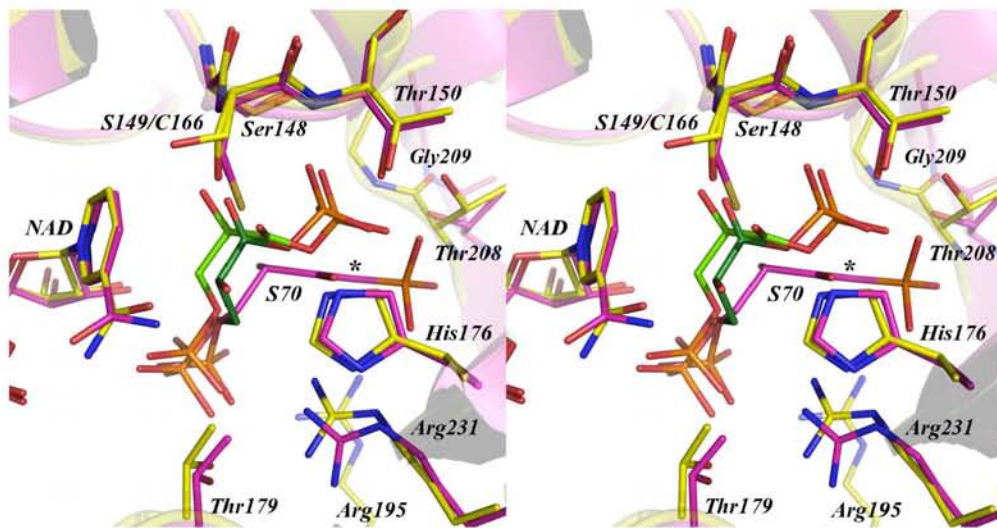
serait donc lié au faible taux d'occupation du 1,3-DPG.

### ***A.3.2.5. Comparaison avec les autres structures de GAPDH en complexe avec des substrats ou analogues de substrat***

#### ***A.3.2.5.1. Comparaison avec la structure de la GAPDH de *T. cruzi* en complexe avec un analogue du 1,3-DPG***

La structure de la GAPDH de *Trypanosoma cruzi* a été résolue en complexe avec un inhibiteur, analogue du 1,3-DPG, de type phosphonate (Ladame *et al.*, 2003, 1qxs). Le positionnement global du 1,3-DPG et de l'analogue, après superposition des deux structures sont relativement similaires (Figure III-20). En effet, leurs groupements C1P et C3P occupent, respectivement les sites Pi et Ps dans les deux cas. Toutefois, la position exacte du groupement C1P, notamment, est différente. Les 2 atomes de phosphore des groupements C1P sont ainsi distants de plus de 3 Å, alors que la distance séparant les deux atomes de phosphore des groupements C3P n'est que de 0,7 Å. Cette différence de localisation pourrait être expliquée par le fait que, dans la structure de *T. cruzi*, la boucle 206-212 (numérotation *B. stearothermophilus*), adopte la conformation B dite "classique" qui délimite le site Pi observé dans la plupart des structures de GAPDH (Figure III-20). La position du groupement C1P de l'analogue n'est toutefois pas équivalente à la position du site Pi "classique" observé dans la structure holoenzyme de la GAPDH de *B. stearothermophilus* mais est relativement proche (moins de 1,5 Å entre les deux atomes de P) de la position du groupement phosphate de l'inhibiteur covalent décrite par Castilho *et al.* (2003) pour l'enzyme de *T. cruzi*. Ainsi, le site Pi de la GAPDH de *T. cruzi* ne semble pas équivalent au site Pi des autres GAPDH ce qui rend les comparaisons particulièrement délicates. Par ailleurs, les atomes O1 et O2 dans les deux molécules occupent des positions complètement différentes, voire opposées.(Figure III-20).

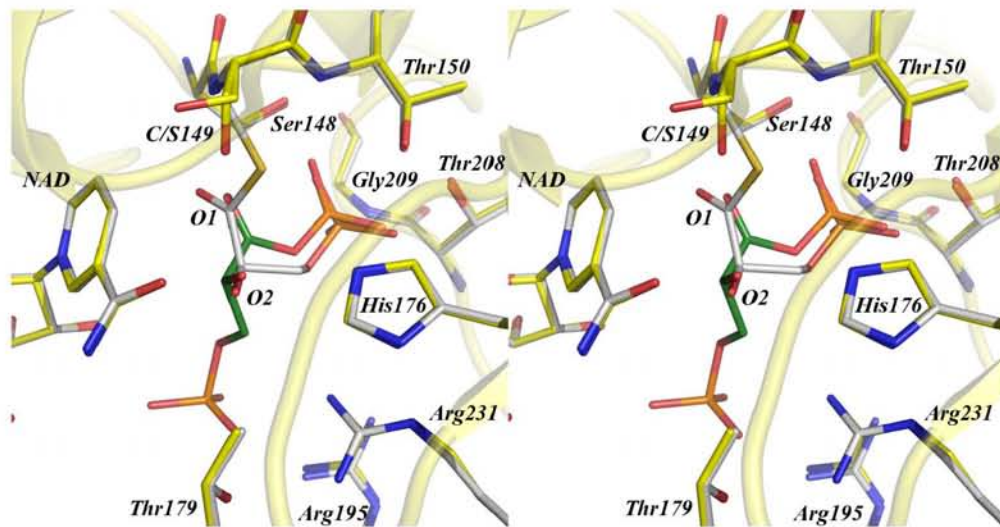
Il semble à ce stade difficile de savoir si ce sont les différences structurales existant entre le 1,3-DPG et l'analogue ou bien les différences structurales existant entre les deux enzymes qui sont à l'origine des différences de positionnement observées.



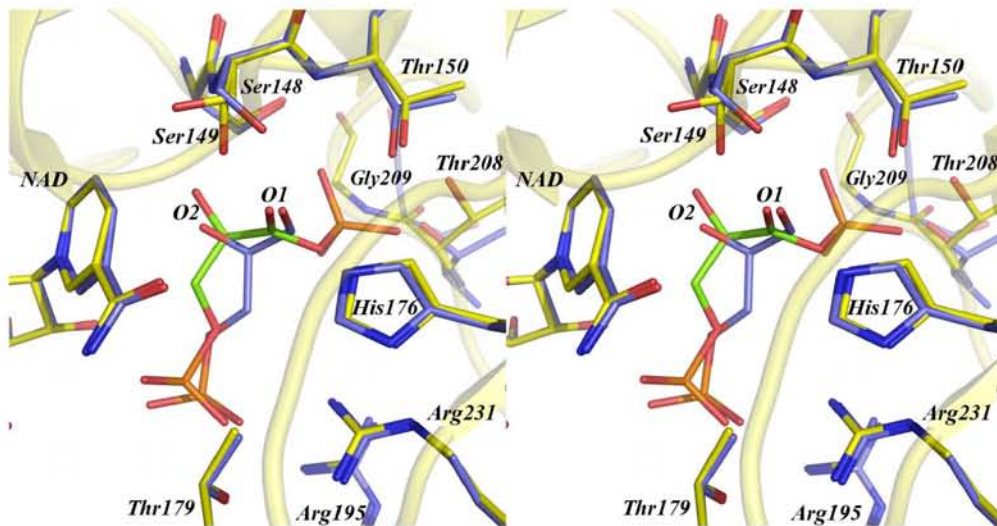
**Figure III-20 : Superposition du complexe ternaire C149S-1,3DPG-NAD<sup>+</sup> (monomère Q, en jaune) et de la structure de la GAPDH de Trypanosoma cruzi (en rose) en complexe avec un analogue du 1,3-diphosphoglycérate (S70).** Pour réaliser cette figure, le monomère A de la structure de T. cruzi (code pdb 1qxs, Ladame et al., 2003) a été superposé sur le monomère Q de notre structure (rmsd de 0,82 Å sur 281 atomes Ca). Les deux conformations, 1 et 2, du 1,3-DPG sont représentées en vert foncé et vert clair, respectivement. Les résidus importants pour la fixation du ligand sont annotés en utilisant la numérotation de la GAPDH de B. stearothermophilus (sauf pour le résidu "catalytique" qui a une double annotation (S149/C166)). L'atome de carbone qui différencie l'analogue de type phosphonate (S70) du 1,3-DPG est mis en évidence par un astérisque.

#### ***A.3.2.5.2. Comparaison avec les structures du complexe michaelien (C149S-G3P-NAD<sup>+</sup>) et de l'intermédiaire thioacylenzyme de la GAPDH de B. stearothermophilus***

Lorsque l'on superpose ce complexe ternaire à la structure du complexe ternaire avec le G3P (1nqo) ou à celle de l'intermédiaire thioacylenzyme (3cmc), on remarque immédiatement la ressemblance structurale existant entre les sites actifs des trois structures (Figure III-21 et III-22). Ainsi, à l'exception de la boucle 206-212, en conformation classique dans le complexe avec le G3P (Figure III-22), la superposition des chaînes latérales des résidus 148, 149, 150, 176, 179, 195 et 231 et des molécules de cofacteur est quasiment parfaite. De façon intéressante, la position du groupement C1P du 1,3-DPG est pratiquement identique à celle du groupement C3P de l'intermédiaire thioacylenzyme (Figure III-21). Les deux atomes de phosphore sont ainsi distants de seulement 0,5 Å et les réseaux d'interactions développés pour la formation du nouveau site Pi sont équivalents. Les atomes O1 et O2 du 1,3-DPG (conformation 1) et de l'intermédiaire thioacylenzyme occupent des positions très similaires, distantes de 0,7 et 0,6 Å, respectivement l'un de l'autre (Figure III-21). De ce fait, l'atome O1 du 1,3-DPG, pointe, comme dans le cas de l'intermédiaire thioacylenzyme, dans une direction opposée à l'His176. Il est également intéressant de préciser que des molécules d'eau occupent régulièrement les positions de ces atomes O1 et O2 dans les structures apo ou holoenzyme de la *Bst*GAPDH (2gd1, 1gd1) ou encore dans la structure holoenzyme du mutant C149S (1nq5). De la même manière, la superposition des deux complexes ternaires (avec le G3P ou le 1,3-DPG) du mutant C149S met en évidence la proximité structurale qui existe dans la localisation du site Ps stabilisant le groupement C3P des deux molécules (Figure III-22). La distance séparant les deux atomes de phosphore est de seulement 0,2 Å. La conformation du G3P peut, cette fois, être rapprochée de la conformation 2 de la molécule de 1,3-DPG. En effet, les atomes O1 et O2 des deux molécules ne sont distants que de respectivement 0,4 et 0,9 Å les uns des autres.



**Figure III-21 : Superposition du complexe ternaire C149S-1,3DPG-NAD<sup>+</sup> (en jaune) avec la structure de l'intermédiaire thioacylzyme (en blanc).** Pour réaliser la figure, le monomère O de la structure de l'intermédiaire thioacylzyme a été superposé sur le monomère Q du complexe ternaire (rmsd de 0,26 Å sur 300 atomes de Ca). Les résidus impliqués dans la fixation du 1,3-DPG (conformation 1, en vert) ou de l'intermédiaire réactionnel (en blanc, lié covalentement à la cystéine catalytique (C149)) sont annotés. Noter la proximité structurale existant entre les atomes O1 et O2 dans les deux ligands.



**Figure III-22 : Superposition des complexes ternaires C149S-1,3DPG-NAD<sup>+</sup> (en jaune) et C149S-G3P-NAD<sup>+</sup> (en bleu).** La sous-unité A de la structure Inqo (Didierjean et al., 2003) a été superposée à la sous-unité Q de notre modèle. Les résidus impliqués dans la fixation des ligands, 1,3-DPG (conformation 2, en vert) et G3P (en bleu) sont annotés. Noter la proximité structurale existant entre les atomes O1 et O2 dans les deux ligands.

## ***A.4. Structure d'un intermédiaire réactionnel acylenzyme-NAD<sup>+</sup>***

### ***A.4.1. Objectif initial et stratégie***

La réaction catalysée par la GAPDH est une réaction totalement réversible. Dans le sens "inverse" de la réaction (sens physiologique de la glucogénèse), le 1,3-DPG se fixe d'abord dans le site actif de l'enzyme en présence de cofacteur oxydé (NAD<sup>+</sup>). Une attaque de la cystéine catalytique intervient alors pour conduire à la libération de phosphate inorganique et à la formation d'un intermédiaire thioacylenzyme toujours en présence de NAD<sup>+</sup>. L'étape suivante consiste en l'échange de la molécule de cofacteur oxydé par un cofacteur réduit qui va ensuite, lors d'un transfert d'hydrure, être oxydé en cédant au substrat l'un des protons portés par le nicotinamide (en position C4) pour donner du G3P. Ainsi, en se plaçant en absence de cofacteur réduit (NADH) il est théoriquement possible de bloquer la réaction au stade thioacylenzyme-NAD<sup>+</sup>. Malgré les essais menés jusqu'à la résolution des structures correspondantes, il s'est toutefois avéré impossible d'obtenir un tel complexe en utilisant des cristaux de l'enzyme de type sauvage, sous forme holoenzyme, incubés en présence de 1,3-DPG. Néanmoins, en utilisant le mutant C149S qui présente une faible activité résiduelle mais une affinité pour les substrats et cofacteurs identiques à celle de l'enzyme sauvage, la structure d'un tel complexe a pu être obtenue. En effet, en utilisant les mêmes conditions que pour la structure du complexe ternaire C149S-1,3DPG-NAD<sup>+</sup>, mais en prolongeant les temps d'incubation des cristaux en présence du 1,3-DPG jusqu'à une vingtaine de minutes, il nous a été possible d'obtenir la structure d'un intermédiaire covalent de type acylenzyme (et non pas thioacylenzyme suite à la mutation de la cystéine catalytique en sérine) en présence de cofacteur oxydé, alors que la structure de l'intermédiaire thioacylenzyme précédemment présentée contenait du NADH.

### ***A.4.2. Présentation de la structure et qualité du modèle***

Contrairement au cas du trempage "court" en présence de 1,3-DPG d'un cristal holoenzyme du mutant C149S présenté précédemment, l'allongement du temps de trempage entraîne une réduction de symétrie du cristal qui passe du groupe d'espace *C 2* à un groupe d'espace *P 21*. La structure a été résolue par remplacement moléculaire à partir d'une des structures résolues lors des

tentatives de piégeage de l'intermédiaire thioacylenzyme à partir des cristaux de PEG et en utilisant le programme Molrep (Vagin and Teplyakov, 1997). L'unité asymétrique est composée de 4 tétramères indépendants (OPQR, ABCD, EFGH et IJKL), 16 molécules de cofacteur, 14 intermédiaires acylenzyme, 13 molécules de 1,3-DPG et de plus de 5500 molécules d'eau. L'affinement a été conduit avec les logiciels Refmac5 et Coot jusqu'à la convergence du modèle à des facteurs de R et  $R_{\text{free}}$  de respectivement 17,7 et 23,0 %. Les statistiques du jeu de données et de la structure sont présentées dans le Tableau III-13. Le diagramme de Ramachandran montre qu'en plus des résidus Asp186 et Val237, les résidus Ala210 et Ala211 appartenant à la boucle mobile 206-212 des sous-unités A et Q, G et H, notamment, sont situés en dehors des régions admises. Ces légères déviations ne remettent cependant pas en cause l'interprétation du modèle. Les faibles valeurs de rmsd obtenues lors de la superposition des différents monomères du modèle sur le monomère O, pris comme référence, (Tableau III-14) indiquent que leur structures sont équivalentes. On notera cependant, dans le cas des monomères G, J, K et L, que les valeurs de rmsd sont sensiblement plus élevées.

Une analyse plus fine des données cristallographiques de chacun des monomères révèle, comme dans le cas de la structure du mutant C149A, que certaines parties du modèle sont moins bien définies que d'autres. Ainsi, l'analyse des facteurs B moyens (Tableau III-15) montre très clairement que 2 des tétramères de l'unité asymétrique, OPQR et EFGH, ont des facteurs B moyens significativement plus faibles (respectivement 19,2 et 17,8 Å<sup>2</sup>) que ceux des deux autres (28,4 et 35,8 Å<sup>2</sup>). La Figure III-23 présente les 4 tétramères du modèle colorés en fonction de la valeur moyenne du facteur B de chaque résidu. La coloration rose observée dans les régions de surface des tétramères OPQR et EFGH correspond à l'augmentation normale du facteur B des régions exposée au solvant. En revanche, la coloration rose-rouge des tétramères ABCD et IJKL et, en particulier des sous-unités J et K, illustre clairement les facteurs B plus élevés de ces éléments du modèle. Ce phénomène a déjà été observé pour la structure du complexe ternaire C149A-1,3DPG-NAD<sup>+</sup> dont les paramètres cristallins sont identiques (Tableau III-7). La qualité des cartes de densité électronique est particulièrement bonne pour la plupart des sous-unités, mais décroît sensiblement dans les sous-unités plus "agitées", comme la sous-unité J en particulier qui présente un facteur B moyen de 47,1 Å<sup>2</sup> et pour laquelle les cartes de densité électronique sont de moindre qualité.

**Tableau III-13 : Statistiques du jeu de données de diffraction et d'affinement de la structure de l'intermédiaire acylenzyme/NAD<sup>+</sup> de la BstGAPDH.**

Données de diffraction	C149S
Groupe d'espace	$P 2_1$
Maille (Å)	$a = 120,2 ; b = 88,1 ; c = 265,7 ; \beta = 102,2^\circ$
Z	8
Limite de résolution (Å)	48,9 - 1,77
Couronne de plus haute résolution (Å)	1,87 - 1,77
Température (K)	100
Réflexions uniques	527207 (76730)
Complétude (%)	99,9 (100,0)
Redondance	3,4 (3,4)
R-merge <sup>a</sup> / Rmeas <sup>a</sup> (%)	8,9 (49,1) / 5,6 (31,4)
I/σ(I)	15,0 (4,0)
Affinement	
R <sup>b</sup> (%)	17,7
R <sub>free</sub> <sup>c</sup> (%)	23,0
<b>r.m.s.d. <sup>d</sup> par rapport à la géométrie standard</b>	
Longueur de liaisons (Å)	0,014
Angles de valence (°)	1,59
chiralité (°)	0,10
<b>Facteur B moyens (Å<sup>2</sup>)</b>	
Modèle	26,5
Atomes de la protéine	34,9
NAD	21,1
Intermédiaire covalent	21,8
1,3-DPG	29,4
Molécules d'eau	35,1

<sup>a</sup> Facteur R pour les intensités reliées par symétrie. <sup>a'</sup> Facteur R indépendant de la redondance (Diederich and Karplus, 1997)

<sup>b</sup> Facteur R cristallographique. Facteur R cristallographique calculé sur 5 % des réflexions choisies aléatoirement et non incluses dans l'affinement.

<sup>d</sup> Écart quadratique moyen.

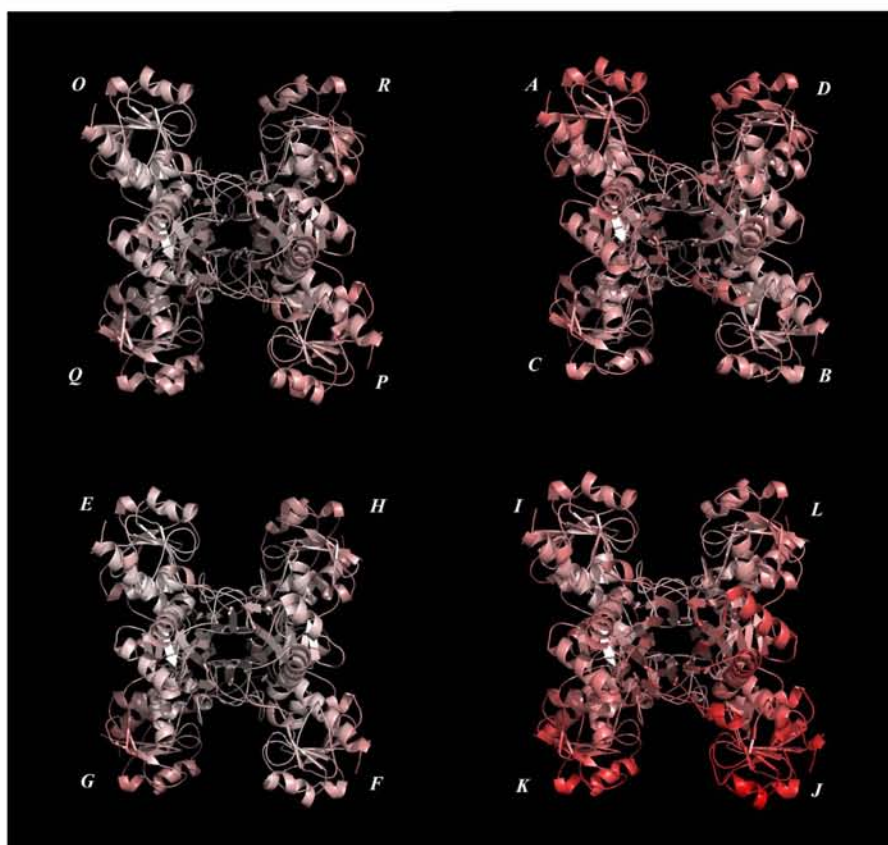


**Tableau III-14 : Valeur de rmsd résultant de la superposition des différents monomères sur le monomère O.** Les valeurs de rmsd ont été calculées sur 326 atomes de Ca (en omettant la boucle 206-212) et sont exprimées en Å.

	O	P	Q	R	A	B	C	D
rmsd	/	0,152	0,153	0,158	0,148	0,157	0,177	0,203
	E	F	G	H	I	J	K	L
rmsd	0,154	0,209	0,281	0,173	0,204	0,320	0,367	0,305

**Tableau III-15 : Étude des facteurs B moyens de la structure par sous-unité et par tétramère.**

Sous-unités	O	P	Q	R	A	B	C	D
B moyen (Å <sup>2</sup> )	16,3	21,8	19,1	20,0	31,6	25,9	27,2	29,0
		19,2				28,4		
Sous-unités	E	F	G	H	I	J	K	L
B moyen (Å <sup>2</sup> )	15,0	17,8	20,2	18,1	28,9	47,1	38,9	28,2
		17,8				35,8		



**Figure III-23 : Vue des 4 tétramères du modèle colorés en fonction de la valeur du facteur B.** La représentation a été réalisée avec le logiciel Pymol et un script dédié spécifiquement à ce type d'analyse. Les résidus ont ainsi été répartis équitablement en 50 groupes en fonction de leur valeur moyenne de facteur B. Les valeurs de B s'étendent de 4,6 (blanc) à 76,1 Å<sup>2</sup> (rouge).

### ***A.4.3. Conformation du site actif***

#### ***A.4.3.1. L'intermédiaire réactionnel acylenzyme***

Le signal observé dans les cartes de densité électronique dans la continuité du résidu "catalytique" Ser149 a été attribué à la présence d'une molécule de G3P liée de manière covalente à l'enzyme par une liaison ester. D'une qualité remarquable dans la plupart des sous-unités, les cartes de densité électronique ont permis de reconstruire dans 14 sites actifs sur 16 la molécule de substrat dans la continuité de la sérine en position catalytique. Les cartes de densité électronique de type "omit" confirment pleinement la présence de cet intermédiaire covalent (Figure III-24). Bien qu'un signal fort apparaisse au niveau des sites actifs des sous-unités G et J, celui-ci n'est pas suffisamment explicite pour modéliser l'intermédiaire avec une confiance raisonnable. Dans les autres sous-unités, 13 sur 14, l'intermédiaire acylenzyme a été modélisé dans une conformation notée  $\alpha$ , majoritaire. Néanmoins, le substrat modélisé dans le monomère K adopte une seconde conformation, notée  $\beta$ , qui est également appelée conformation minoritaire. Cette seconde conformation de l'intermédiaire acylenzyme est également rencontrée dans les monomères P, Q, R et L, en coexistence avec la conformation  $\alpha$  majoritaire. La différence majeure existant entre ces deux conformations, qui seront décrites et comparées plus précisément par la suite, concerne la position de l'atome O1 du substrat. Le taux d'occupation de l'intermédiaire réactionnel a été estimé à 0,5 en moyenne mais peut varier de 0,15 à 0,6 en fonction de la sous-unité ou de la conformation considérée et son facteur B moyen est de 21,8 Å<sup>2</sup> (Tableau III-16). Des affinements en cours réalisés par C. Vonrhein avec Buster-TNT (Bricogne and Irwin, 1996), devraient permettre d'obtenir de meilleures estimations de ces occupations.

La molécule de 3-phosphoglyceroyl est liée de façon covalente à l'enzyme au travers d'une liaison ester, plane, formée entre son atome C1 et l'atome O<sub>γ</sub> de la Ser149. Dans la conformation  $\alpha$  majoritaire (Tableau III-16), les atomes O1 et O2 de l'intermédiaire forment des liaisons hydrogène respectivement avec l'His176 et le NAD<sup>+</sup> (Figure III-25). Le groupement C3P est fixé dans le nouveau site Pi, en liaison hydrogène avec les résidus Ser148, Thr150, Thr208 et Gly209 ainsi que 3 molécules d'eau (Figure III-25 et Tableau III-17 pour les distances interatomiques). Dans la conformation  $\beta$  minoritaire, le réseau de liaisons hydrogène créé par le substrat est globalement identique à celui précédemment décrit pour la conformation  $\alpha$ , à ceci près que l'atome d'oxygène O1, dans une orientation inverse, est, cette fois, lié à l'atome d'azote de la chaîne principale en

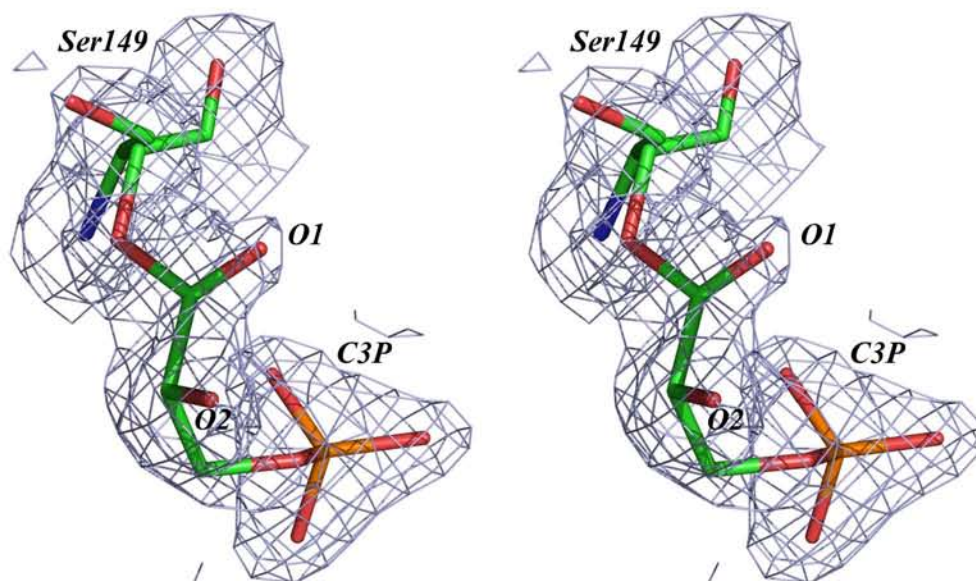
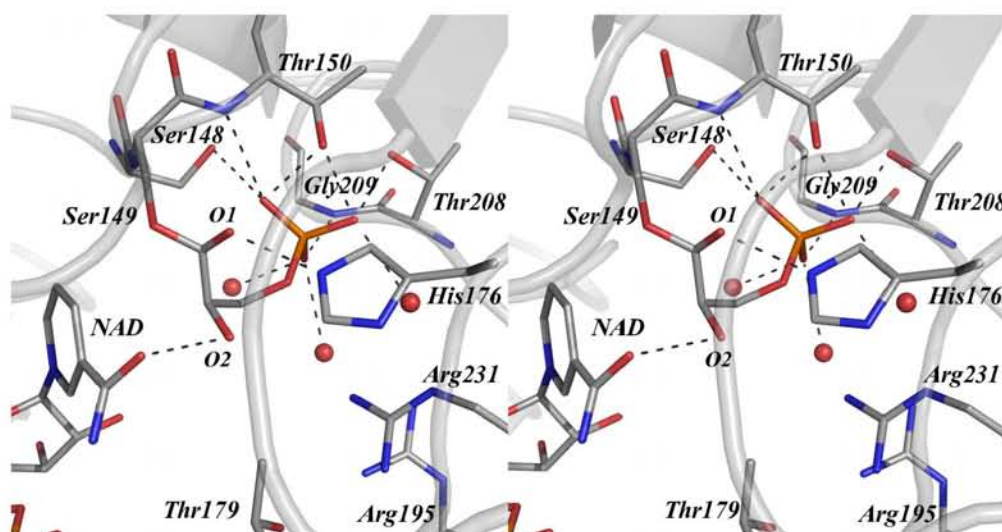


Figure III-24 : Vue stéréoscopique de l'intermédiaire acylenzyme de la GAPDH (monomère O). Le résidu Ser149, sous forme libre et liée, et la molécule de substrat (conformation  $\alpha$ ) sont représentés dans une carte de densité électronique omit de type 2mFo-DFc contourée à 1 sigma. Les atomes O1 et O2 et le groupement phosphate lié au carbone 3 du substrat (C3P) sont annotés.

Tableau III-16 : Facteurs d'occupation et facteurs B moyen des intermédiaires acylenzymes (AE1) et des molécules de 1,3-DPG (dpg) dans les différentes sous-unités du modèle. Lorsque 2 conformères existent, les taux d'occupation et valeurs de B moyen sont donnés pour chacune des conformations (conf  $\alpha/\beta$ )

		O	P	Q	R	A	B	C	D
A E 1	Facteur d'occupation	0,50	0,35/0,25	0,30/0,20	0,55/0,25	0,45	0,50	0,50	0,60
	Facteur B moyen	16,8	17,1/16,1	14,3/15,6	19,5/17,0	26,8	22,9	18,9	24,9
d p g	Facteur d'occupation	0,35	0,30	0,24	/	0,30	0,32	0,25	/
	Facteur B moyen	20,1	36,4	17,7	/	41,8	34,1	34,7	/
		E	F	G	H	I	J	K	L
A E 1	Facteur d'occupation	0,60	0,50	/	0,35	0,50	/	0,30	0,15/0,25
	Facteur B moyen	19,6	15,8	/	29,7	33,6	/	32,3	14,6/17,1
d p g	Facteur d'occupation	0,30	0,33	0,60	0,35	0,30	/	0,40	0,34
	Facteur B moyen	22,8	24,3	35,4	21,4	19,2	/	35,9	34,9



**Figure III-25 : Vue stéréoscopique de l'intermédiaire acylenzyme dans le site actif du monomère O.** Les résidus du site actif, l'intermédiaire acylenzyme (conformation  $\alpha$ ) et la molécule de  $\text{NAD}^+$  sont représentés sous forme de bâtons et annotés. Les molécules d'eau en interaction avec l'intermédiaire sont représentées sous forme de sphères. Les liaisons hydrogène stabilisant le substrat sont symbolisées par des pointillés. Noter la conformation *cis* de l'angle de torsion O1-C1-C2-O2. ( cf. Tableau III-17 pour les distances)

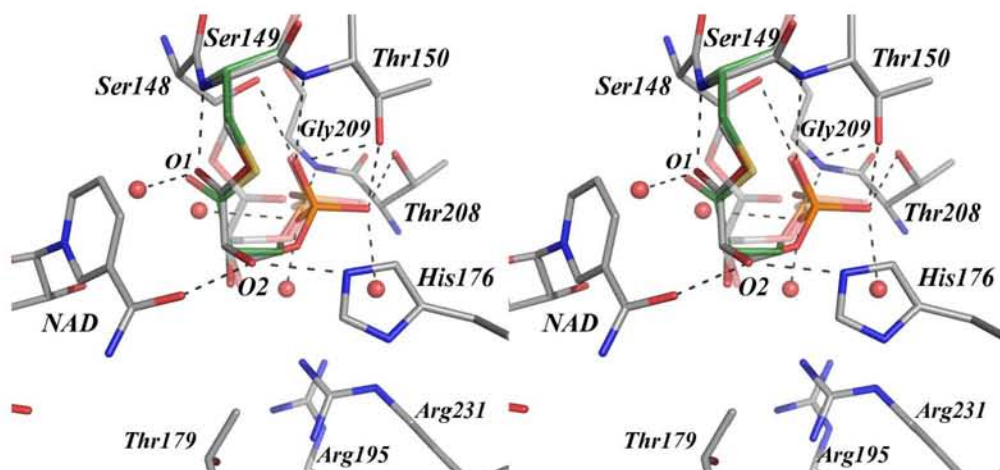
**Tableau III-17 : Interactions de l'intermédiaire (conformation  $\alpha$ ) avec la structure protéique.** (les distances entre atomes sont données en Å)

	O	P	Q	R	A	B	C	D	E	F	H	I	L
O1 --- His176(N $\epsilon_2$ )	2,6	2,4	2,5	2,5	2,5	2,6	2,4	2,6	2,4	2,5	2,2	2,3	2,3
O2 --- NAD(NO7)	3,1	3,2	3,6	3,2	3,1	3,2	2,9	3,3	3,2	2,9	4,1	3,4	3,5
O2P --- Thr150(O $\gamma_1$ )	2,6	2,7	3,1	2,6	2,8	2,6	2,5	2,5	2,5	2,4	2,9	2,8	2,7
O2P --- Thr208(O $\gamma_1$ )	2,7	2,6	2,5	2,6	2,9	2,6	2,6	2,8	2,7	2,6	2,3	3,0	2,8
O2P --- H <sub>2</sub> O	2,7	2,8	2,6	2,7	2,8	2,9	2,8	2,6	2,7	3,0	2,5	2,7	2,7
O3P --- Gly209(N)	2,6	2,6	2,5	2,6	2,7	2,6	2,6	2,7	2,7	2,8	3,1	2,4	2,0
O3P --- H <sub>2</sub> O	2,8	2,6	2,5	2,3	2,5	2,6	2,6	2,5	2,9	2,9	3,1	3,3	2,8
O3P --- H <sub>2</sub> O	2,7	2,5	2,6	2,9	2,7	2,6	2,7	2,6	2,8	2,8	2,3	2,5	/
O4P --- Ser148(O $\gamma$ )	2,9	2,5	2,8	2,5	2,8	2,5	2,7	2,7	3,1	2,7	2,7	2,9	2,8
O4P --- Thr150(N)	2,9	3,2	3,2	2,9	2,9	3,0	3,0	2,9	2,9	2,9	4,1	3,1	3,0
O4P --- Thr150(O $\gamma_1$ )	2,7	3,4	3,0	2,8	3,0	3,0	2,9	3,1	2,8	3,0	3,3	2,8	2,8

position 149 plutôt qu'au résidu His176 (Figure III-26 et Tableau III-18). La conformation  $\beta$  est équivalente à celle de l'intermédiaire thioacylenzyme caractérisé dans la première partie de ces travaux (Figure III-26). Seule la position de l'atome C3 diffère visiblement entre les deux intermédiaires. En revanche, la position de l'atome O1 comme celle du groupement C3P sont parfaitement identiques. Le passage de la conformation  $\alpha$  à la conformation  $\beta$  est obtenu par des rotations d'environ  $180^\circ$  autour des liaisons  $C\beta-O\gamma$  et C1-C2 et d'environ  $40^\circ$  autour des liaisons  $C\alpha-C\beta$  et  $O\gamma-C1$  qui permettent, sans modifier la position des atomes C2 et O2 et du groupement C3P, de faire pointer l'atome O1 de l'intermédiaire acylenzyme vers l'His176 (conformation  $\alpha$ ) ou vers l'azote de la chaîne principale en position 149 (conformation  $\beta$ ). D'un point de vue énergétique, la conformation  $\alpha$  avec un angle de torsion O1-C1-C2-O2 moyen de  $36^\circ$  (de  $16$  à  $59^\circ$ ) semble moins stable que la conformation  $\beta$  dans laquelle cet angle de torsion a une valeur moyenne de  $-140^\circ$ , plus proche d'une conformation *trans*.

Il est à noter que la sérine 149 n'est que partiellement acylée et que, en conséquence, celle-ci a été modélisée sous deux ou trois conformations alternatives, selon qu'une ou deux conformations de l'intermédiaire acylenzyme sont présentes dans le site actif. Les deux conformations acylées sont très proches l'une de l'autre puisqu'une simple rotation autour de la liaison  $C\alpha-C\beta$  ( $\chi_1$ ) de  $40$  à  $50^\circ$  existe entre les deux. Comme dans le cas du complexe ternaire C149S-1,3DPG-NAD<sup>+</sup>, la conformation non acylée de la Ser149 pointe vers l'arrière du site actif et son groupement hydroxyle forme une liaison hydrogène avec la fonction carbonyle de la chaîne principale du résidu Asn313.

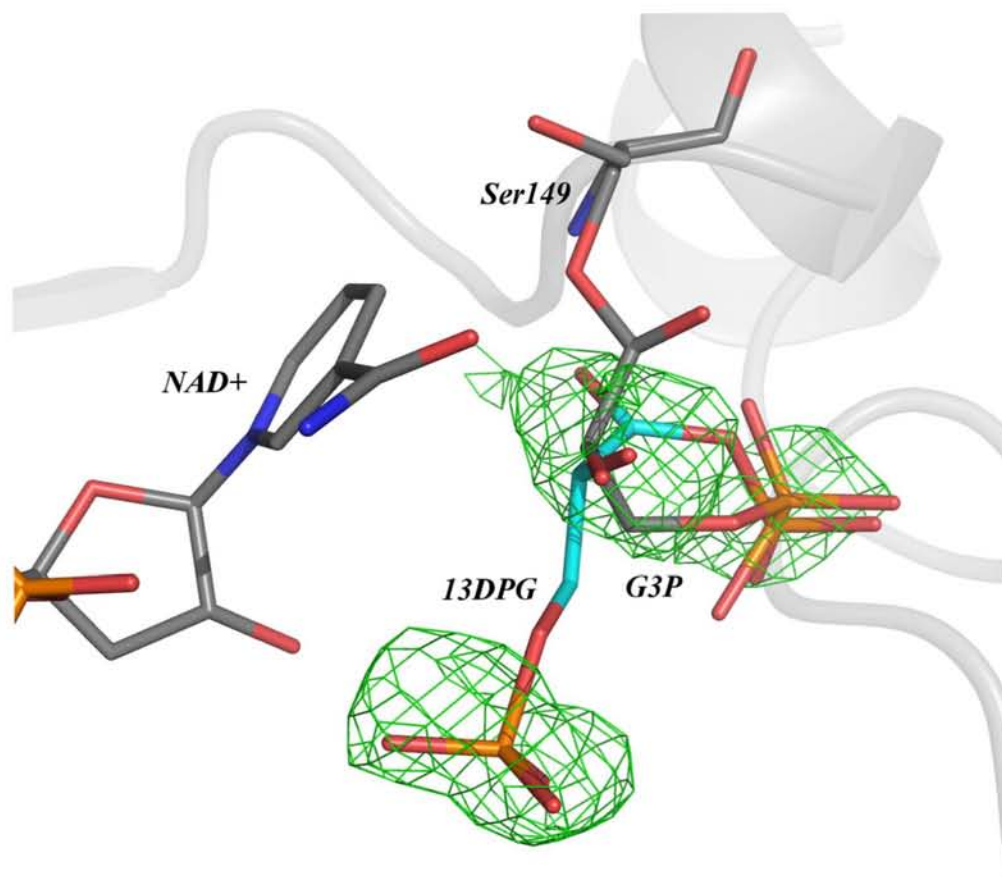
L'analyse des cartes de densité électronique et notamment des cartes Fourier différence révèle cependant que la seule présence de l'intermédiaire acylenzyme ne suffit pas à expliquer le signal de densité présent dans le site actif (Figure III-27). En effet, un fort signal résiduel (3 sigma) reste visible au niveau du site Ps et à proximité du résidu catalytique. Lorsque l'on superpose à cette structure le complexe ternaire C149S-1,3DPG-NAD<sup>+</sup> (Figure III-27), la position de la molécule de 1,3-DPG (en conformation 1) semble pouvoir expliquer le signal de densité positive observé, à la fois dans les sites Pi et Ps avec les deux groupements phosphate, et aussi à proximité de la Ser149, avec les atomes C1 et O1. Ainsi 13 molécules de 1,3-DPG ont été modélisées dans la structure en occupation partielle. Comme dans le cas de l'intermédiaire acylenzyme, le facteur d'occupation de ces molécules a été estimé en se basant sur l'analyse des facteurs B et des cartes de densité électronique. Le taux d'occupation moyen de la molécule de 1,3-DPG, estimé à 0,34, est inférieur à celui de l'intermédiaire acylenzyme et son facteur B moyen est de  $29,4 \text{ \AA}^2$  (Tableau III-16). Comme dans le cas des complexes ternaires C149A/S-1,3DPG-NAD<sup>+</sup>, la molécule de 1,3-DPG n'est pas complètement définie dans les cartes de densité électronique avec un défaut de signal au niveau de



**Figure III-26 : Vue stéréoscopique de l'intermédiaire acylenzyme dans le site actif du monomère P.** Les résidus du site actif, l'intermédiaire acylenzyme (conformation  $\beta$ ) et la molécule de  $\text{NAD}^+$  sont représentés sous forme de bâtons et annotés. Les molécules d'eau en interaction avec l'intermédiaire sont représentées sous forme de sphères. Les liaisons hydrogène stabilisant le substrat sont symbolisées par des pointillés. Noter la conformation trans de l'angle de torsion O1-C1-C2-O2. L'intermédiaire thioacylenzyme en liaison covalente avec la Cys149 qui présente une conformation proche de l'acylenzyme  $\beta$  est également représenté sous forme de bâtons en vert (3mc) et la conformation  $\alpha$  de l'acylenzyme en transparence (cf. Tableau III-18 pour les distances)

**Tableau III-18 : Interactions de l'intermédiaire (conformation  $\beta$ ) avec la structure protéique.** (les distances entre atomes sont données en Å)

	P	Q	R	K	L
O1 --- Ser149(N)	2,5	2,2	2,5	2,1	2,7
O1 --- H <sub>2</sub> O	2,7	2,6	2,6	3,0	2,6
O2 --- NAD(NO7)	2,8	3,0	3,0	3,7	3,5
O2 --- His176(N $\epsilon_2$ )	2,6	3,0	2,6	3,4	3,3
O1P --- H <sub>2</sub> O	/	/	/	/	3,1
O2P --- Thr150(O $\gamma_1$ )	2,4	2,5	2,5	2,9	2,6
O2P --- Thr208(O $\gamma_1$ )	2,9	3,1	2,6	2,9	2,6
O2P --- H <sub>2</sub> O	2,6	2,6	2,6	2,8	2,8
O3P --- Gly209(N)	3,1	3,1	2,8	/	2,2
O3P --- H <sub>2</sub> O	3,2	3,4	2,8	2,6	2,9
O3P --- H <sub>2</sub> O	2,5	2,3	2,6	2,4	/
O4P --- Ser148(O $\gamma$ )	3,1	3,1	2,9	3,4	2,9
O4P --- Thr150(N)	2,9	3,4	3,0	/	3,2
O4P --- Thr150(O $\gamma_1$ )	2,6	3,1	2,7	/	2,7



**Figure III-27 :** *Vue du site actif de la sous-unité O (gris) dans lequel est lié l'intermédiaire acylenzyme (conformation  $\alpha$ ). La sérine en position catalytique (Ser149), l'intermédiaire et la molécule de cofacteur de cette structure sont représentés en mode bâton et colorés en gris. En bleu, la molécule de 1,3-DPG (conformation 1) de la sous-unité O de la structure du complexe ternaire C149S-1,3DPG-NAD<sup>+</sup> après superposition des deux structures. La carte de densité électronique Fourier différence de type mFo-DFc est contournée à 3 sigma.*

l'atome C3. Les interactions créées entre la protéine et le 1,3-DPG sont similaires à celles décrites pour les complexes ternaires en partie A.3.2.3.2.

#### *A.4.3.2. Conformation de la boucle 206-212*

Comme dans le cas des structures présentées précédemment, la boucle 206-212 a été modélisée sous deux conformations alternatives dans 13 des 16 sous-unités (une seule conformation dans les sous-unités B, J et K). Les facteurs d'occupation relatifs de chacune des deux conformations ont été estimés sur la base de l'analyse des facteurs B et des cartes de densité électronique Fourier différence (Tableau III-19). Notées A et B, les deux conformations sont équivalentes aux conformations qui ont été décrites dans la structure de l'intermédiaire thioacylenzyme et qui correspondent respectivement aux conformations dites "nouvelle" (nouveau site Pi) et "classique" (site Pi classique) (non montré). Dans la plupart des sous-unités, la conformation majoritaire est la conformation délimitant le nouveau site Pi (taux d'occupation moyen estimé à 0,64) dans lequel sont fixés les groupements C3P et C1P de l'intermédiaire acylenzyme et du 1,3-DPG, respectivement. Dans le cas des sous-unités H, K et L où la conformation dite "classique" est majoritaire (respectivement 0,65, 1 et 0,7 de taux d'occupation) la position du nouveau site Pi est plus variable, à l'image des atomes de phosphore des groupements C3P et C1P respectivement de l'intermédiaire et du 1,3-DPG, qui peuvent s'écarter de 0,7 à 1,8 Å dans le cas du monomère K.

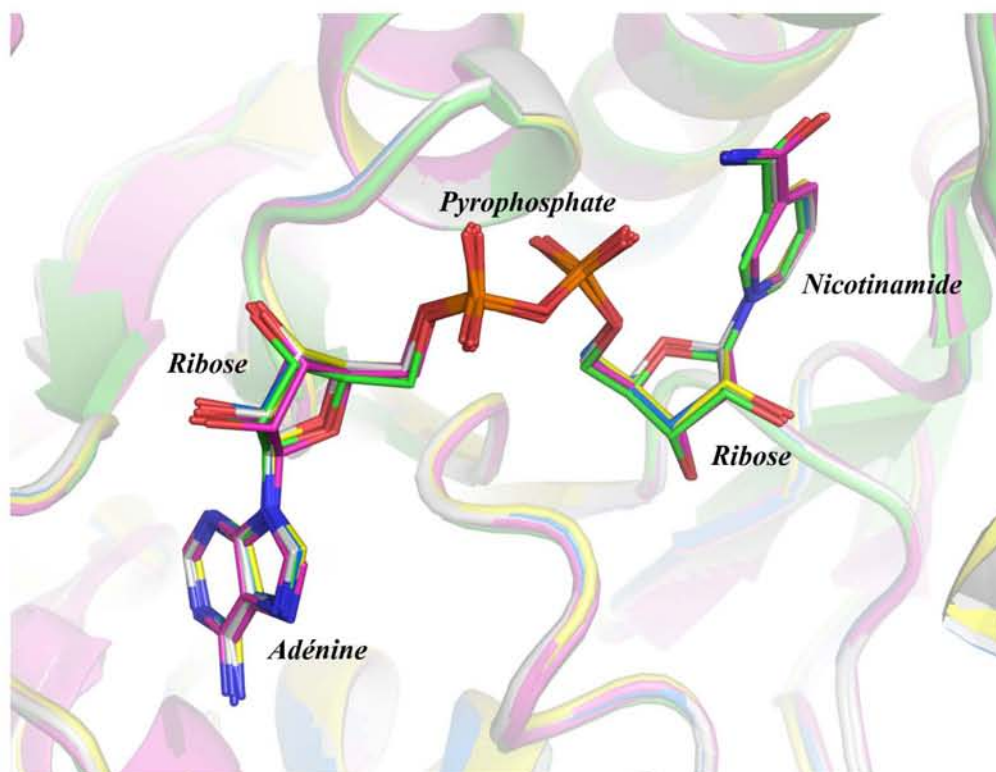
#### *A.4.3.3. La molécule de cofacteur*

Chaque sous-unité du modèle contient une molécule de cofacteur NAD<sup>+</sup> qui a été modélisée avec un taux d'occupation de 1. Leur facteur B moyen est variable en fonction de la sous-unité considérée (de 10,5 à 45,5 Å<sup>2</sup>) ce qui est à rapprocher des précédentes remarques sur l'analyse des facteurs B des différentes sous-unités. Toutefois, le facteur B moyen des molécules de cofacteur à 21,1 Å<sup>2</sup> est inférieur au facteur B moyen du modèle (26,5 Å<sup>2</sup>) et le facteur B moyen de la molécule de cofacteur de chaque monomère est, dans tous les cas, inférieur à celui de la sous-unité correspondante. La conformation du cofacteur tout comme le réseau de liaisons hydrogène créé avec la protéine sont identiques dans l'ensemble des sous-unités du modèle et, comme l'illustre la Figure III-28, sont également en tous points comparables à ceux décrits pour la structure holoenzyme, pour les complexes ternaires C149A/S-G3P-NAD<sup>+</sup> et à ceux observés dans les structures du thioacylenzyme (NADH) et des complexes ternaires C149A/S-1,3DPG-NAD<sup>+</sup>.



**Tableau III-19 : Paramètres cristallographiques de la boucle 206-212 de la structure de l'acylenzyme-NAD<sup>+</sup>. Deux valeurs sont reportées et correspondent aux conformations A et B, dites nouvelle et classique, respectivement. Les bornes de la région modélisée en double conformation sont indiquées avec la lettre correspondant à la sous-unité.**

Sous-unités	O	P	Q	R	A	B	C	D
	207-213	206-213	207-213	206-213	206-214		206-212	206-212
<b>Taux d'occupation estimé</b>	0,65 0,35	0,65 0,35	0,65 0,35	0,65 0,35	0,65 0,35	1	0,60 0,40	0,7 0,3
<b>Facteurs B</b>	18,3 18,1	22,2 16,9	22,8 20,9	16,8 15,0	33,5 31,2	35,0	25,9 27,6	29,6 27,2
Sous-unités	E	F	G	H	I	J	K	L
	206-213	206-212	206-212	206-212	206-213	207-213		207-213
<b>Taux d'occupation estimé</b>	0,60 0,40	0,65 0,35	0,65 0,35	0,35 0,65	0,60 0,45	1	1	0,3 0,7
<b>Facteurs B</b>	16,5 15,2	17,4 15,9	17,0 17,6	16,1 16,4	32,4 31,1	49,8	38,9	22,1 23,5



**Figure III-28 : Superposition des molécules de cofacteur des différentes formes de la GAPDH de B. stearothermophilus. La sous-unité O de la structure holoenzyme (en vert) (Skarzynski et al., 1987 : 1gd1) a été prise comme référence pour superposer les sous-unités O des complexes ternaires C149A/S-G3P-NAD<sup>+</sup> (en rose) (Didierjean et al., 2003 : Inqa, Inqo), de la structure du thioacylenzyme (en jaune), des complexes ternaires C149A/S-1,3DPG-NAD<sup>+</sup> (en blanc) et de l'intermédiaire acylenzyme (en bleu). Les différents groupements chimiques composant le NAD sont annotés.**

## ***PARTIE B : DISCUSSION***

### ***L'intermédiaire thioacylenzyme-NADH : structure et implication pour la catalyse***

Cette étude avait pour objectif premier l'obtention de la structure cristallographique de l'intermédiaire thioacylenzyme de la GAPDH. La stratégie développée pour piéger cet intermédiaire dans les cristaux s'est appuyée sur la forte altération de la vitesse de l'étape de désacylation en absence de phosphate. En effet, dans ces conditions, le taux d'hydrolyse de l'intermédiaire est de  $10^{-2} \text{ s}^{-1}$  seulement, soit environ  $6 \cdot 10^{-4}$  fois moins rapide que l'étape d'acylation (Michels *et al.*, 1996), et autorise l'accumulation de l'intermédiaire. En dépit de l'occupation partielle de cet intermédiaire dans le modèle final, la qualité des données de diffraction et de la structure résolue nous a permis de modéliser deux états distincts de l'enzyme, coexistant au sein de chacun des monomères de l'enzyme. L'état le plus peuplé correspond à la structure de l'intermédiaire thioacylenzyme lié covalamment à la cystéine catalytique, alors que l'autre correspond à un état où la cystéine catalytique est libre et où un anion sulfate est fixé dans le site de reconnaissance Pi. Chacun de ces états est associé à une conformation particulière de la boucle 206-212, à partir de laquelle les résidus invariants Thr208 et Gly209 interagissent directement soit avec l'intermédiaire thioacylenzyme soit avec l'anion sulfate.

Le groupement phosphate (C3P) de l'intermédiaire réactionnel est localisé au niveau du nouveau site Pi. Ce résultat est en accord avec la localisation de l'intermédiaire hémithioacétal (*sp3*)

obtenu avec l'apoGAPDH d'*E. coli* (Yun *et al.*, 2000). Les deux intermédiaires se superposent étroitement (Figure III-12), sauf au niveau de l'atome C1, vraisemblablement en raison de la différence d'hybridation de cet atome entre les deux intermédiaires (*sp*<sup>2</sup> ou *sp*<sup>3</sup>). Par ailleurs, les structures considérées comme représentatives d'un complexe michaelien productif entre l'enzyme, le NAD et le G3P, montrent que le site Ps constitue le site de fixation initial du groupement C3P dans le site actif de l'enzyme (Didierjean *et al.*, 2003). L'ensemble de ces résultats suggère que, même en présence d'un site Ps complet (présence du ribose du cofacteur pour interagir avec le site Ps, une condition qui n'était pas remplie dans le cas de la structure de l'hémithiacétal obtenue à partir de la forme apo de la GAPDH d'*E. coli*), le groupement C3P a migré du site Ps vers le nouveau site Pi au cours de la réaction d'oxydoréduction. Le moment exact auquel ce mouvement a lieu demeure incertain. L'hypothèse selon laquelle la formation de la liaison covalente entre la cystéine catalytique et le G3P conduirait au déplacement du C3P vers le nouveau site Pi ne peut être totalement écartée. Néanmoins, il semble plus probable que l'ensemble de la première étape de la réaction se déroule alors que le groupement C3P est positionné dans le site Ps et que celui-ci ne bascule qu'à l'issue du transfert d'hydrure. Cette hypothèse présente notamment l'avantage de maintenir une interaction étroite entre le substrat et le cofacteur lors du transfert d'hydrure, qui pourrait être favorable à la réalisation de cette étape.

En plus de la localisation du groupement phosphate du substrat, l'orientation de son atome d'oxygène O1, qui pointe à l'opposé de l'His176 dans une conformation trans énergétiquement favorable vis-à-vis de l'atome O2, est inattendue. En effet, il est admis que pour catalyser efficacement un transfert d'hydrure, le site actif d'une enzyme doit présenter soit un site oxyanion soit un catalyseur basique capable de stabiliser la charge négative développée lors de la formation de l'état de transition tétraédrique. Or, les études menées sur la GAPDH ont explicitement démontré que l'His176 jouait le rôle de catalyseur basique (Soukri *et al.*, 1989 ; Talfournier *et al.*, 1998), et aucun site oxyanion n'est identifiable à partir de la structure. De ce fait, le transfert d'hydrure a obligatoirement eu lieu alors que l'atome O1 était bien orienté par rapport à l'histidine catalytique. Dans la structure de l'intermédiaire thioacylenzyme présentée ici, représentative d'un état consécutif au transfert d'hydrure, l'atome O1 adopte une conformation différente, qui résulte dans la perte de l'alignement des atomes N<sub>ε</sub><sub>2</sub> de l'His176, O1 et C1 du substrat et C4 du noyau pyridinium.

Il apparaît ainsi que la réaction d'oxydoréduction implique d'importants changements conformationnels du substrat, parmi lesquels le changement d'orientation de l'atome O1 vis-à-vis de l'His176 et le déplacement du groupement C3P du substrat depuis le site Ps vers le nouveau site Pi. La possibilité que ces deux mouvements surviennent de manière concertée ou non demeure

incertaine. Toutefois, le fait que ces deux événements semblent devoir survenir à la même étape, à savoir après le transfert d'hydrure, semble *a priori* indiquer que ces changements conformationnels sont concertés. Il faut noter qu'à l'exception de la boucle 206-212, la conformation du site actif ne montre aucun changement notable entre la structure du complexe michaelien et celle de l'intermédiaire thioacylenzyme, et que seul le substrat voit sa conformation modifiée au cours de l'étape d'oxydoréduction. Le passage d'une conformation à l'autre implique principalement une rotation d'environ 180° du plan O1-C1-C2-O2, ainsi qu'une rotation autour de la liaison C2-C3 qui achève le positionnement du groupement C3P dans le nouveau site Pi.

Dans le mécanisme catalytique, l'étape consécutive à la formation de l'intermédiaire thioacylenzyme correspond à la libération de la molécule de NADH (étape cinétiquement déterminante) et à l'entrée d'une nouvelle molécule de NAD<sup>+</sup>, qui favorise la réaction de phosphorylation (Trentham, 1971a). Les événements conduisant à la libération du cofacteur restent indéterminés. Cependant, la perte de l'interaction entre l'atome O2' du ribose et le groupement C3P qui survient lors du déplacement du phosphate du substrat depuis le site Ps jusqu'au nouveau site Pi, pourrait constituer le premier signal conduisant à la dissociation de la molécule de NADH. Dans ce contexte, des calculs réalisés sur le serveur web PISA (Krissinel and Henrick, 2007 ; [http://www.ebi.ac.uk/msd-srv/prot\\_int/pistart.html](http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html)) montrent que la surface accessible au solvant de la molécule de cofacteur passe de 13,9 Å<sup>2</sup> dans la structure du complexe michaelien dans laquelle le G3P est fixé dans le site Ps (1nqo) à 33,4 Å<sup>2</sup> dans la structure de l'intermédiaire thioacylenzyme (3cmc). Parallèlement, il convient de rappeler qu'il est impossible de distinguer les deux formes de la molécule de cofacteur à partir des cartes de densité électronique et ce, alors que NADH et NAD<sup>+</sup> coexistent dans les cristaux (réaction incomplète). Ainsi, mis à part le point précédemment évoqué, il n'existe aucun signe d'une quelconque déstabilisation du complexe enzyme-cofacteur qui pourrait être interprété comme la première étape vers la libération du NADH.

La fixation d'une nouvelle molécule de NAD<sup>+</sup> conduit vraisemblablement ensuite aux changements conformationnels nécessaires à la réalisation de la seconde étape de la réaction (phosphorylation). En effet, considérant que le résidu His176 joue potentiellement le même rôle que lors de l'oxydoréduction, mais aussi que le phosphate inorganique attaque vraisemblablement l'intermédiaire thioacylenzyme depuis le nouveau site Pi, alors la conséquence de la fixation de la molécule de cofacteur oxydé devrait être la réorientation correcte de l'atome O1 vis-à-vis de l'His176 et le basculement du groupement C3P depuis le nouveau site Pi vers le site Ps.

## ***Les complexes michaeliens (GAPDH-1,3-DPG-NAD<sup>+</sup>) : structures, identification du site de phosphorylation et caractérisation du flip-flop***

Afin de compléter la collection des "images" des différentes étapes du mécanisme catalytique de la GAPDH et aussi dans le but de confirmer l'implication du nouveau site Pi dans l'étape de phosphorylation, il nous a paru nécessaire de caractériser le complexe ternaire formé entre la GAPDH le NAD<sup>+</sup> et le produit de la réaction (1,3-DPG). En utilisant la même démarche que celle utilisée dans la résolution du complexe michaelien avec le G3P (Didierjean *et al.*, 2003), à savoir la diffusion du ligand dans des cristaux de mutants inactif ou faiblement actif (C149A/S), cette étude a permis d'obtenir les deux premiers complexes de la GAPDH avec son produit de réaction et le NAD<sup>+</sup>. Le mécanisme catalytique de la GAPDH étant entièrement réversible, ce complexe correspond également au complexe de type michaelien (enzyme-substrat) de la réaction en sens inverse.

Bien que l'occupation de la molécule de 1,3-DPG dans les structures ne soit pas complète, elle atteint des valeurs acceptables qui font de la molécule de ligand l'espèce majoritaire présente dans le site actif du mutant C149S. La qualité des cartes de densité électronique, notamment dans le cas du mutant C149S (1,66 Å de résolution), tout comme la convergence des observations réalisées sur chacun des deux mutants (C149A et C149S), nous permettent de considérer avec une bonne confiance la structure de ces complexes.

Dans ces structures où le 1,3-DPG est lié de façon non covalente à l'enzyme, les groupements C1P et C3P interagissent respectivement avec le nouveau site Pi et avec le site Ps. Le positionnement du groupement C1P dans le nouveau site Pi confirme que ce site de reconnaissance anionique est bien le site à partir duquel a lieu l'attaque nucléophile du phosphate inorganique (phosphorylation). Le groupement C3P occupe, lui, le site de reconnaissance anionique Ps. L'analyse des différentes structures de la GAPDH de *B. stearothermophilus* replacées dans le déroulement du mécanisme catalytique (*i.e.* complexe michaelien avec le G3P, intermédiaire thioacylenzyme et complexe ternaire avec le 1,3-DPG) permet de conclure que le groupement C3P opère un mouvement de "va-et-vient" ou "flip-flop" lors de la catalyse. En effet, le groupement C3P du substrat se fixe d'abord dans le site Ps (complexe michaelien) bascule ensuite vers le nouveau site Pi (intermédiaire thioacylenzyme), pour revenir dans le site Ps pour la phosphorylation (complexes ternaires avec le 1,3-DPG).

Deux conformations du 1,3-DPG, notées 1 et 2, sont observées dans la structure du

complexe ternaire (GAPDH-1,3DPG-NAD<sup>+</sup>). Elles se distinguent l'une de l'autre principalement au travers de la position des atomes O1 et O2 de la chaîne carbonée du ligand sans que les positions des groupements phosphate soient modifiées. Dans la conformation majoritaire (conformation 1), l'atome O1 pointe vers l'atome d'azote de la chaîne principale du résidu 149, alors que dans l'autre (conformation 2) celui-ci pointe vers l'His176. Bien qu'aucune étude cinétique n'ait pu démontrer le rôle de catalyseur acide/base de l'His176 dans la seconde étape de la réaction, ce rôle est toutefois communément accepté, au vu de la ressemblance des deux états de transition tétraédriques. Il semble donc clair que la conformation minoritaire du 1,3-DPG (conformation 2) est la plus favorable pour subir l'attaque nucléophile de la cystéine catalytique avec l'assistance de l'His176. Dans le complexe michaelien, il n'est pas exclu que le 1,3-DPG oscille entre ces deux conformations. D'un point de vue énergétique, la conformation 1 présente une conformation *trans* pour les atomes O1 et O2 et apparaît donc être la plus stable. Ceci pourrait expliquer pourquoi la conformation 1 *a priori* non compétente pour la catalyse a pu être accumulée au détriment de l'autre dans nos conditions d'expériences.

### ***L'intermédiaire acylenzyme-NAD<sup>+</sup> : influence de l'état redox du cofacteur sur la conformation du substrat et lien avec la catalyse***

L'influence de l'état d'oxydation du cofacteur c'est-à-dire de l'étape d'échange de cofacteur sur la conformation de l'intermédiaire thioacylenzyme est alors apparue comme étant la principale inconnue à résoudre pour compléter l'étude structurale du mécanisme de la GAPDH. D'après le schéma du mécanisme catalytique de l'enzyme, il semblait possible, dans le sens inverse de la réaction (c'est à dire en utilisant le 1,3-DPG en tant que substrat), d'accumuler un intermédiaire thioacylenzyme de la GAPDH en présence de NAD<sup>+</sup>. En effet, le NAD<sup>+</sup> doit être présent pour réaliser la première étape de la réaction (dans le sens 2 comme dans le sens 1). Dans le sens 2, la première étape de la réaction conduit à la formation d'un intermédiaire thioacylenzyme et à la libération d'une molécule de phosphate inorganique. Une autre étape est alors nécessaire avant le transfert d'hydrure et consiste en l'échange du cofacteur NAD<sup>+</sup> par du NADH. En absence de NADH dans le milieu, l'intermédiaire thioacylenzyme peut donc être théoriquement piégé. L'utilisation du substrat naturel et de cristaux de l'enzyme de type sauvage n'a pu conduire à l'observation de cet intermédiaire, pour des raisons probablement liées au nombre restreint de tentatives, et surtout par l'absence d'informations provenant de la microspectrophotométrie, méthode qui s'était avérée déterminante dans l'obtention de la structure de l'intermédiaire thioacylenzyme à partir du G3P. Néanmoins, un trempage de longue durée en présence de 1,3-DPG

(22 min) des cristaux du mutant C149S a conduit à l'accumulation d'un intermédiaire acylenzyme en présence de  $\text{NAD}^+$ , dont la structure a été exploitée en comparaison à celle de l'intermédiaire thioacylenzyme précédemment obtenu.

Le groupement C3P de l'intermédiaire acylenzyme- $\text{NAD}^+$  est localisé au niveau du nouveau site Pi. Le fait que le groupement C3P du 1,3-DPG ait été observé dans le site Ps dans les complexes ternaires (C149S/A-1,3DPG- $\text{NAD}^+$ ) montre qu'un basculement de ce groupement phosphate vers le nouveau site Pi a lieu à la suite de la libération du phosphate inorganique. Ceci suggère donc que le mouvement de "va-et-vient" du groupement C3P a lieu également dans le sens 2 de la réaction. De plus, le fait que les intermédiaires covalents thioacylenzyme- $\text{NADH}$  et acylenzyme- $\text{NAD}^+$  présentent tous deux leur groupement C3P dans le nouveau site Pi confirme l'importance du basculement du substrat dans ce site pour la réalisation de l'échange de cofacteur. En effet, dans les deux cas, l'étape consécutive à l'obtention de l'intermédiaire covalent correspond à l'échange de la molécule de cofacteur, et la libération de celui-ci est certainement facilitée par la perte de l'interaction entre l'atome O2' du ribose et le groupement C3P qui survient lors du déplacement du phosphate du substrat depuis le site Ps jusqu'au nouveau site Pi (voir ci-dessus). Ce résultat constitue un point important de notre étude. En effet, jusqu'à présent, la littérature s'était limitée à proposer l'existence de cette étape d'échange de cofacteur sur la base d'observations cinétiques montrant que la présence du  $\text{NAD}^+$  augmentait la vitesse de phosphorolyse tout comme celle de la déphosphorylation réductive. Pour la première fois, cette étude montre que les deux phénomènes, flip-flop et échange de cofacteur, semblent étroitement liés sur le plan structural. D'un côté, le flip-flop semble nécessaire pour que l'échange de cofacteur puisse avoir lieu pour la raison évoquée ci-dessus. Toutefois, s'il est évident que l'échange de cofacteur est essentiel dans le sens 2 de la réaction pour la réalisation du transfert d'hydrure, la façon dont le  $\text{NAD}^+$  accélère la phosphorolyse (sens 1) reste incomprise. En ce sens, les premières hypothèses émises à la suite de l'analyse de la structure de l'intermédiaire thioacylenzyme- $\text{NADH}$ , nous avaient conduits à attribuer à l'entrée du  $\text{NAD}^+$  un rôle déterminant pour la suite de la réaction. De fait, compte tenu de la conformation de l'intermédiaire thioacylenzyme obtenu à partir de G3P, il est nécessaire préalablement à la phosphorolyse que le groupement C3P du substrat libère le nouveau site Pi, mais aussi fasse basculer l'atome O1 en direction de l'His176. L'intermédiaire acylenzyme- $\text{NAD}^+$ , dans sa conformation  $\alpha$  majoritaire, présente effectivement un atome O1 idéalement orienté vis-à-vis de l'His176 qui tend à confirmer le rôle, jusqu'à présent non démontré, de l'His176 dans cette étape de la réaction. De plus, le positionnement différent de l'atome O1 dans cet intermédiaire contenant du  $\text{NAD}^+$ , par rapport à l'intermédiaire thioacylenzyme- $\text{NADH}$  est également en accord avec notre

première hypothèse concernant l'influence de l'état d'oxydation du cofacteur sur l'orientation de cet atome. Toutefois, l'observation d'une conformation  $\beta$  minoritaire de l'intermédiaire acylenzyme-NAD<sup>+</sup>, équivalente à la conformation de l'intermédiaire thioacylenzyme-NADH (*i.e.* avec un atome O1 dans une direction opposée à l'His176) est troublante. La présence des deux conformations au sein de la même structure (comme dans les complexes avec le 1,3-DPG) met clairement en évidence la liberté conformationnelle de l'atome O1 du substrat. D'un point de vue énergétique, il semble évident que cette conformation  $\beta$  dans laquelle les atomes O1 et O2 adoptent une conformation *trans*, est plus stable que la conformation  $\alpha$  de l'acylenzyme dont les atomes O1 et O2 sont plus proches l'un de l'autre (angle de torsion O1-C1-C2-O2 moyen de seulement 36°). Bien que moins stable, la conformation  $\alpha$  (O1 orienté vers l'histidine) représente une conformation dont l'existence est nécessaire pour la catalyse, ce qui est moins évident pour l'autre conformation. Ainsi, il n'est pas exclu que la conformation  $\beta$ , équivalente à la conformation du thioacylenzyme, ne soit qu'un artéfact de l'accumulation de l'intermédiaire. En solution, l'existence potentielle de cette conformation serait vraisemblablement conditionnée par le rapport des constantes de vitesse représentatives de l'isomérisation du O1 d'une part, et de la libération du NADH d'autre part. Lors de l'analyse de la structure de l'intermédiaire thioacylenzyme, nous avons également proposé que cette conformation, qui rompt l'alignement entre les atomes N $\epsilon$ 2 de l'His176, O1 et C1 du substrat et C4 du nicotinamide essentiel à la réalisation du transfert d'hydrure, puisse présenter un avantage catalytique réel. En effet, l'étape limitante de la réaction est l'étape immédiatement consécutive à la formation du thioacylenzyme, à savoir la libération de la molécule de cofacteur réduite. En conclusion, il apparaît pour l'heure hasardeux de vouloir conclure sur l'existence de la conformation  $\beta$  de l'intermédiaire covalent dans le déroulement normal de la réaction.

### ***Proposition d'un mécanisme***

Le G3P se fixe dans le site actif de l'enzyme avec son groupement C3P dans le site Ps et son atome O1 orienté vers l'His176. La réaction commence par l'attaque nucléophile de la Cys149 sur le carbone aldéhydique C1 du substrat et conduit à la formation d'un intermédiaire tétraédrique transitoire stabilisé par l'His176. Le transfert d'hydrure à partir de l'atome C1 vers l'atome C4 du NAD<sup>+</sup> est ensuite réalisé grâce à l'assistance basique de l'histidine catalytique. A cette étape de la réaction, le groupement C3P du substrat est toujours vraisemblablement localisé au niveau du site Ps, en interaction étroite avec la molécule de cofacteur.

L'intermédiaire thioacylenzyme subit des réarrangements conformationnels, au cours desquels son groupement C3P bascule vers le nouveau site Pi et son atome O1 change



probablement d'orientation de telle sorte qu'il ne pointe plus en direction de l'His176. Ces réarrangements conduisent à l'intermédiaire thioacylenzyme-NADH dont la structure a été obtenue. L'étape suivante consiste en l'échange de la molécule de cofacteur (sortie de NADH et entrée de  $\text{NAD}^+$ ). Cette étape est facilitée par la perte des interactions entre le cofacteur et le substrat à la suite du basculement du groupement C3P de ce dernier vers le nouveau site Pi. La fixation d'une nouvelle molécule de  $\text{NAD}^+$  favorise ensuite la phosphoryse probablement en induisant les réarrangements conformationnels du substrat nécessaires au déroulement de cette étape (c'est-à-dire le retour du groupement C3P vers le site Ps et éventuellement le retournement de l'atome O1 en direction de l'His176). Enfin, le phosphate inorganique se fixe et attaque l'intermédiaire thioacylenzyme à partir du nouveau site Pi, conduisant à la libération du produit (1,3-DPG). A la fin, l'enzyme retrouve une forme holoenzyme, prête à fixer une nouvelle molécule de G3P.



**CHAPITRE IV : ETUDE STRUCTURALE  
ET FONCTIONNELLE DE L'ERYTHROSE-  
4-PHOSPHATE DÉSHYDROGÉNASE  
D'*ESCHERICHIA COLI***

# *PARTIE A : RÉSULTATS*

## *A.1. Contexte et principaux résultats*

Comme exposé dans l'introduction de ce manuscrit, la seule structure d'E4PDH connue au début de ces travaux était celle de l'E4PDH d'*Escherichia coli* résolue par Brice Kauffmann (2000) au LCM3B. D'une résolution modeste (2,7 Å), cette structure renferme des zones mal définies, notamment au niveau de la boucle S (181-204) dont la région 178-191 a été modélisée par une chaîne de polyalanine. Par ailleurs, l'absence de tout ligand pouvant mimer le phosphate au niveau des sites de reconnaissance anionique potentiels de l'enzyme n'a pas permis l'analyse des résidus participant à leur formation. De plus, cette structure a révélé la présence d'un pont disulfure entre la cystéine 153 et la cystéine 311. Ce pont n'étant pas détecté lors de dosages de thiols en solution (Boschi-Muller, communication personnelle), il s'est vraisemblablement formé dans le cristal suite à l'oxydation des cystéines, et pourrait engendrer des contraintes dans une zone proche du site de fixation du nicotinamide. Enfin, l'E4PDH se caractérise par une relativement faible affinité vis-à-vis du NAD<sup>+</sup> par rapport à la GAPDH avec laquelle elle partage pourtant une forte identité de séquence (valeur de K<sub>D</sub> égale à 0,8 mM, soit 800 fois supérieure à celle de la *Bst*GAPDH, Boschi-Muller *et al.*, 1997). Jusqu'à présent, cette faible affinité n'a pu être expliquée ni par l'analyse de la séquence, ni par l'analyse de la structure d'E4PDH obtenue sous la forme apoenzyme (Kauffmann, 2000).

L'un des premiers objectifs de ces travaux, débutés en DEA, a donc été d'obtenir de nouvelles conditions de cristallisation de l'E4PDH, de façon à améliorer la résolution de la structure, mais aussi dans le but d'obtenir des structures renfermant un ou plusieurs ligands : anion phosphate (ou analogue), cofacteur et substrat.

Le criblage d'une centaine de conditions de cristallisation a permis d'isoler 6 conditions, dont une contenant des ions phosphates et donc potentiellement intéressante pour cette étude. L'optimisation de cette condition a conduit à l'obtention reproductible de cristaux de l'E4PDH d'*E. coli* et à la collecte de données de diffraction sur la ligne de lumière FIP (BM30A) de l'ESRF à Grenoble (Tableau IV-1), données utilisées pour la résolution d'une structure (APO1) à 2,30 Å présentant un anion phosphate au niveau du site actif de l'enzyme. Dans le même temps, dans le but d'obtenir des structures de complexes binaires enzyme-substrat et enzyme-cofacteur, des conditions de co-cristallisation de l'enzyme avec le NAD et ou l'E4P ont été recherchées, ceci en utilisant une enzyme inactive dans laquelle le résidu cystéine catalytique avait été remplacé par un résidu alanine. Afin d'éliminer la possibilité de formation du pont disulfure artéfactuel entre les résidus Cys153 et Cys311, une mutation supplémentaire a été introduite, correspondant à la substitution de la cystéine en position 153 par une sérine. Dans le cadre de la recherche de conditions de co-cristallisation de l'enzyme en présence du substrat, une nouvelle forme cristalline de l'enzyme, de type quadratique, a été obtenue. Malgré l'absence de ligand dans les différentes structures résolues, cette nouvelle forme cristalline, obtenue avec le double mutant C149A-C153S, révèle pour la première fois une région 178-191 définie dans les cartes de densité électronique. Ce manuscrit présente l'analyse de l'une de ces structures (APO2), résolue avec des données de diffraction collectées à 2,38 Å de résolution sur la ligne de lumière ID14eh3 de l'ESRF (Tableau IV-1) en utilisant la méthode du remplacement moléculaire avec le modèle APO1. Enfin, l'obtention d'une structure de l'E4PDH sous forme holoenzyme, c'est-à-dire en complexe avec le cofacteur, s'étant avérée impossible, nous avons finalement décidé de résoudre la structure de l'enzyme en complexe covalent avec un inhibiteur, le 3-(chloroacétyl) pyridine adénine dinucléotide (3-CAPAD). En effet, ce marqueur d'affinité est un analogue du cofacteur NAD dans lequel la fonction carboxamide en C3 du noyau pyridine est "remplacée" par une fonction chlorométhylcétone. La cystéine catalytique attaque l'inhibiteur sur le carbone électropositif conduisant à une inactivation irréversible (départ du chlore) (Biellmann *et al.*, 1974). L'E4PDH préalablement inactivée par traitement au 3-CAPAD a été cristallisée et la structure du complexe covalent, notée ECAP, a été résolue par remplacement moléculaire à partir de la structure de la forme apoenzyme (APO1). La structure dérive d'un jeu de

**Tableau IV-1 : Statistiques de collecte, d'affinement et qualité géométrique des structures apoenzymes de l'E4PDH d'E. coli (APO1 et APO2) et de la structure du complexe avec l'inhibiteur 3-CAPAD (ECAP).**

Statistiques de collecte	APO1	APO2	ECAP
<b>Ligne de lumière</b>	FIP - BM30A ESRF	ID14eh3 ESRF	ID14eh4 ESRF
<b>Longueur d'onde (Å)</b>	0,9797	0,9310	0,9395
<b>Groupe d'espace</b>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 4 <sub>1</sub> 2 <sub>1</sub> 2	<i>P</i> 1
<b>Paramètres de maille (Å)</b>	<i>a</i> = 86,2 ; <i>b</i> = 110,7 ; <i>c</i> = 137,9	<i>a</i> = 134,8 ; <i>b</i> = 134,8 ; <i>c</i> = 246,1	<i>a</i> = 96,0 ; <i>b</i> = 111,5 ; <i>c</i> = 135,2 <i>α</i> = 88,7° ; <i>β</i> = 83,9° ; <i>γ</i> = 88,4°
<b>Limite de résolution (Å) (dernière couronne)*</b>	30 - 2,30 (2,42 - 2,30)	60 - 2,38 (2,51 - 2,38)	60 - 2,90 (3,06 - 2,9)
<b>I/σ(I)</b>	24,4 (5,6)*	31 (6,3)*	9,7 (2,1)*
<b>Rmerge (%)</b>	5,0 (39,0)* <i>Rmerge</i> 5,5 (43,1)* <i>Rmeas</i>	7,3 (70,3)* <i>Rmeas</i> 2,4 (29,2)* <i>Rp.i.m.</i>	10,4 (58,0)* <i>Rmerge</i> 14,1 (78,6)* <i>Rmeas</i>
<b>Réflexions uniques</b>	59159 (8270)*	91 158 (12 989)*	117 678 (17 594)*
<b>Complétude</b>	99,6 (99,8)*	99,7 (98,4)*	95,3 (96,1)*
<b>Redondance</b>	5,6 (5,5)	10,3 (5,9)	2,1 (2,1)
Statistiques d'affinement			
Facteur R (%)	20,3	17,0	24,2
Facteur R <sub>free</sub> (%)	26,2	22,6	30,8
B moyen modèle (Å <sup>2</sup> )	35,9	34,6	45,7
B moyen phosphate (Å <sup>2</sup> )	25,0	/	/
B moyen 3-CAPAD (Å <sup>2</sup> )	/	/	65
B moyen sulfate (Å <sup>2</sup> )	/	/	79,7
B moyen molécules d'eau (Å <sup>2</sup> )	38,1	40,4	24,5
Déviation par rapport à la géométrie idéale			
Longueurs de liaison (Å)	0,012	0,024	0,010
Angles de valence (°)	1,5	2,1	1,3
Angles de torsion (°)	18,7	19,2	20,8
Chiralité (°)	0,10	0,16	0,09

*Rmeas* = Facteur R indépendant de la redondance (Diederichs and Karplus, 1997)

*Rp.i.m.* = Facteur indépendant de la redondance et indicatif de la précision de la mesure moyenne (Weiss, 2001)

données à 2,9 Å de résolution collecté sur la ligne de lumière ID14eh4 de l'ESRF (Tableau IV-1). Cette structure confirme effectivement la présence de la molécule de 3-CAPAD dans le site de fixation du cofacteur et a permis l'analyse des interactions potentielles entre l'enzyme et le cofacteur.

Dans la suite de ce manuscrit, ces trois structures de l'*EcoE4PDH* ont été analysées et comparées les unes aux autres et les observations réalisées confrontées aux connaissances approfondies disponibles pour les GAPDH. Ces deux enzymes partagent en effet plus de 40 % d'identité de séquence. Bien que notre modèle d'E4PDH soit issu d'*E. coli*, c'est la GAPDH de *B. stearothermophilus* qui a été principalement utilisée comme référence, puisque cette dernière présente une plus forte identité de séquence avec l'E4PDH d'*E. coli* que la GAPDH d'*E. coli* elle-même (43,6 contre 40 % d'identité). Les travaux d'enzymologie réalisés par l'équipe du Professeur Branlant sur l'E4PDH d'*E. coli* ont d'ailleurs également utilisé la GAPDH de *B. stearothermophilus* comme référence pour les comparaisons (Boschi-Muller *et al.*, 1997).

## ***A.2. Présentation des structures et qualité des modèles***

### ***A.2.1. Structure cristallographique de la forme apoenzyme de l'E4PDH d'E. coli (APO1)***

La structure est composée d'un tétramère physiologique d'E4PDH (chaîne A à D) qui contient les résidus 0 à 333, quatre anions phosphates et environ 700 molécules d'eau. La structure, résolue par remplacement moléculaire à partir de la structure d'E4PDH obtenue par Kauffmann (2000), a été affinée jusqu'à la convergence des facteurs R et R<sub>free</sub> à des valeurs de 20,4 et 26,8 %, respectivement (Tableau IV-1). Le diagramme de Ramachandran du modèle révèle l'existence de 17 résidus dans des régions non favorables du diagramme. Mis à part le résidu Asn186 qui, bien que relativement bien défini dans les cartes de densité électronique, est systématiquement en dehors des régions favorables dans les quatre sous-unités du modèle, les autres résidus concernés appartiennent principalement à des régions exposées au solvant, relativement mal définies. La qualité de ces cartes a cependant permis de modéliser sans ambiguïté la chaîne principale sur pratiquement toute sa longueur ainsi que la plupart des chaînes latérales. Comme pour la structure d'E4PDH déjà connue (Kauffmann, 2000), la boucle S (région s'étendant des résidus 181 à 204 de la séquence) n'est pas complètement définie dans les cartes de densité électronique. L'absence totale de signal de densité

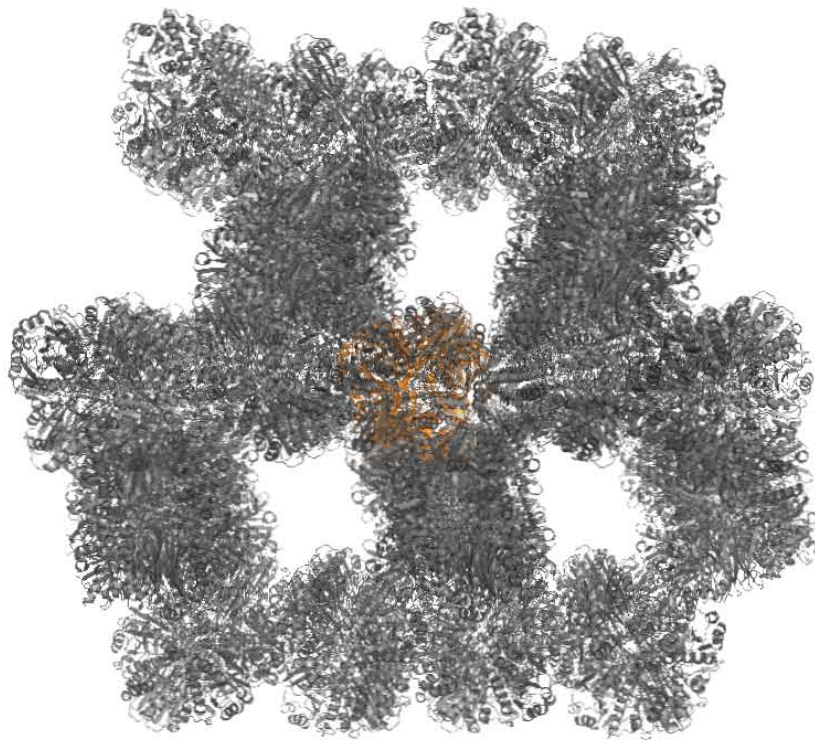
électronique pour les résidus 180 à 191 suggère fortement que cette boucle est particulièrement agitée et non structurée. Les résidus 181-189 n'ont donc pas été modélisés.

### ***A.2.2. Nouvelle forme cristalline de l'E4PDH d'E. coli sous forme apoenzyme (APO2)***

Le modèle est composé d'un tétramère d'E4PDH (chaîne A à D ; résidus 0 à 334), de 5 molécules d'éthylène glycol et d'une molécule de polyéthylène glycol provenant de la condition de cristallisation, de 6 molécules de glycérol provenant du traitement de cryoprotection des cristaux, et d'environ un millier de molécules d'eau. Le modèle, affiné jusqu'à des valeurs de 17,0 et 22,6 % de R et  $R_{\text{free}}$ , respectivement, présente une qualité géométrique satisfaisante (Tableau IV-1) malgré la présence de 7 résidus non glycine dans les zones défavorables du diagramme de Ramachandran. Ces résidus appartiennent à des zones relativement mal définies dans les cartes de densité électronique (boucle S du monomère A) et/ou exposée au solvant (Arg61A, Asn221B). Le modèle est bien défini dans les cartes de densité électronique y compris la boucle S (181-204) des sous-unités B et C qui a pu être entièrement reconstruite et dont la conformation est observée pour la première fois. Dans les sous-unités A et D, le signal n'est, en revanche, pas suffisamment clair pour reconstruire complètement cette boucle. Ainsi les régions 186-191 du monomère A et 181-183 du monomère D n'ont pas été modélisées.

L'analyse de l'empilement cristallin de cette nouvelle forme cristalline quadratique donne un coefficient de Matthews de 3,76 et un taux de solvant de 67,3 % qui, comparés aux valeurs obtenues pour la forme monoclinique précédemment décrite (avec des valeurs de 2,21 et 44,4 %, respectivement) révèle un empilement globalement moins compact de cette forme quadratique. En effet, de larges canaux de solvant existent dans cet empilement (Figure IV-1). L'analyse et la comparaison des contacts intermoléculaires existant entre les différents tétramères dans ces deux empilements cristallins, révèlent que si les surfaces de contact avec les tétramères voisins sont similaires (égales à 12400 et 12480 Å<sup>2</sup>, respectivement pour les formes monoclinique et quadratique) leur répartition est différente. En effet dans le cas de la forme monoclinique, cette surface de contact est continue alors que dans le cas de la forme quadratique, les contacts inter-tétramères se répartissent en 3 régions isolées.





**Figure IV-1 : Vue de l'empilement cristallin de la forme quadratique de l'E4PDH d'E. coli (APO2). Le tétramère d'E4PDH constituant la maille élémentaire du cristal est représenté en orange, avec, en gris, les molécules voisines, reconstituées par symétrie et formant l'empilement cristallin du modèle APO2. Cette forme cristalline se caractérise par l'existence de larges canaux de solvant.**

### ***A.2.3. Structure cristallographique du complexe entre l'E4PDH d'E. coli et l'inhibiteur 3-(chloroacétyl) pyridine adénine dinucléotide (ECAP)***

Quatre tétramères indépendants d'E4PDH composent la maille triclinique du modèle final qui contient également 22 anions sulfates (dont 16 dans le site actif de l'enzyme) et près de 700 molécules d'eau. Les monomères sont notés de A à D pour le premier tétramère, de E à H pour le second, de I à L pour le troisième et enfin de M à P pour le dernier. Justifiée par la faible résolution des données de diffraction, l'application de contraintes strictes de géométrie au niveau de l'affinement grâce au facteur de pondération (géométrie/données de diffraction) et l'utilisation de contraintes de symétrie non cristallographique entre les différentes sous-unités ont permis d'obtenir un modèle de bonne qualité géométrique (Tableau IV-1) présentant des valeurs de R et  $R_{\text{free}}$  de 24,2 et 30,8 %, respectivement. Seuls 27 résidus d'acides aminés non glycine sur les 5115 composant le modèle sont situés à l'extérieur des zones autorisées du diagramme de Ramachandran. Ces résidus appartiennent essentiellement à des régions mal définies dans les cartes de densité électronique (correspondant notamment aux résidus 47-51 et 181-204 (boucle S)) qui seront discutées plus tard dans ce manuscrit. Il faut également noter que, comme dans les autres structures d'E4PDH, certains résidus n'ont pas pu être modélisés du fait de l'absence de signal dans les cartes de densité électronique. Il s'agit des résidus 179 à 188 des sous-unités A, E, N et P, des résidus 179 à 190 des sous-unités B, F, G, J, L, M et O, des résidus 179 à 191 des sous-unités C, D et K et de 179 à 192 ou 196, respectivement pour les sous-unités I et H. Ces régions correspondent à la partie N-terminale de la boucle S (181-204) qui a déjà été décrite comme étant non structurée dans la structure de l'E4PDH sous forme apoenzyme (APO1), mais qui a pu être observée dans la structure du double mutant (APO2).

## ***A.3. Analyse des structures tertiaire et quaternaire de l'E4PDH d'E. coli***

### ***A.3.1. Analyse topologique de la structure***

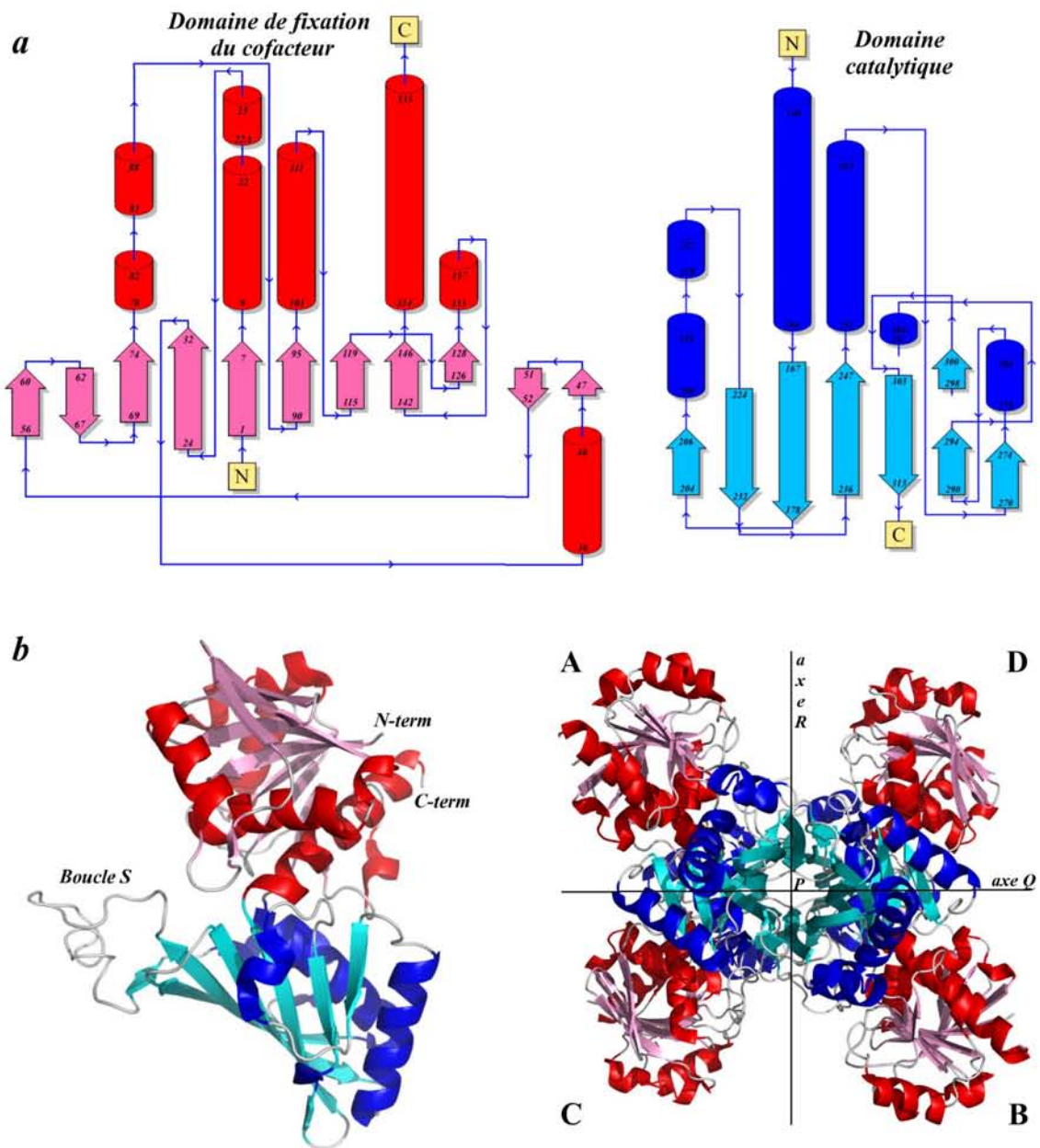
L'analyse de ces structures confirme que la structure de l'E4PDH est très proche structurellement des GAPDH phosphorylantes. L'E4PDH est un homotétramère dont chacun des

monomères est composé de deux domaines. Le domaine de fixation du cofacteur est composé des résidus 0 à 147 qui adoptent le repliement caractéristique de Rossmann (Rossmann *et al.*, 1974) associés à la partie C-terminale (résidus 314-334), structurée en une hélice  $\alpha$  qui vient se plaquer sur le feuillet central (Figure IV-2). Composé des résidus 148 à 313, le second domaine, dit domaine catalytique, présente un repliement de type  $\alpha/\beta$ , constitué d'un feuillet  $\beta$  central à 8 brins et de 6 hélices  $\alpha$  qui sont concentrées sur une seule face du feuillet, l'autre face intervenant dans une interaction inter-monomères (Figure IV-2).

### ***A.3.2. Comparaison de la structure des monomères d'E4PDH***

Au sein d'un même modèle, les superpositions des différentes sous-unités considérées deux à deux donnent des écarts-types (rmsd) sur les distances entre carbones  $\alpha$  équivalents de 0,66 Å pour la structure APO1, 0,62 Å pour la structure APO2 et une valeur de 0,34 Å pour la structure ECAP (Tableau IV-2). Les très faibles valeurs obtenues pour la superposition des Ca des monomères du modèle ECAP sont dues à l'utilisation de contraintes de symétrie non cristallographique lors de l'affinement de la structure. Concernant les structures APO1 et APO2, l'analyse des profils de déviation a tout d'abord permis de mettre en évidence l'existence de boucles adoptant des conformations variables selon le monomère considéré. Ces régions (listées dans le Tableau IV-2), qui sont parfois bien définies dans les cartes de densité électronique, présentent en effet des déviations importantes dues soit à l'empilement cristallin soit à l'existence de conformations alternatives dans les différentes sous-unités. Les cas les plus intéressants seront discutés spécifiquement dans le paragraphe A.5. Ces régions ont alors été écartées de l'analyse, et les rmsd réévalués (Tableau IV-2) sont seulement de 0,40 Å et 0,37 Å respectivement pour les modèles APO1 et APO2. Ces faibles valeurs indiquent que les trois modèles (APO1, APO2 et ECAP) sont des homotétramères ayant une symétrie 222 quasi parfaite. Sauf mention contraire, une seule sous-unité pour chaque modèle sera prise en compte dans les analyses.

Lorsque l'on compare les différentes structures, les superpositions (en excluant toujours les régions décrites dans le Tableau IV-2) des carbones  $\alpha$  des monomères d'une structure avec ceux d'une autre structure donnent des rmsd moyens de 0,58 Å entre les monomères des structures APO1 et APO2, de 0,37 Å entre les monomères des structures APO1 et ECAP et de 0,42 Å entre les monomères des structures APO2 et ECAP. Même si l'une des valeurs (0,58 Å) est un peu plus



**Figure IV-2 : Présentation de la structure de l'E4PDH d'E. coli.** (a) Schéma topologique des deux domaines constituant le monomère d'E4PDH. Le domaine de fixation du cofacteur adopte un repliement caractéristique de Rossmann et le domaine catalytique un repliement  $\alpha/\beta$ . (b) A gauche, le monomère d'E4PDH est représenté sous forme de rubans et flèches avec les deux domaines colorés en rouge et bleu en respectant les couleurs utilisées dans le schéma topologique. A droite, le tétramère physiologique d'E4PDH est représenté avec la même légende de couleur que précédemment. Les axes de symétries internes de l'enzyme, nommés P Q et R par analogie aux GAPDH, sont indiqués et permettent à partir d'une sous-unité A choisie arbitrairement de définir les sous-unités B C et D, respectivement.

**Tableau IV-2 : Valeurs de rmsd résultant des superpositions des atomes de Ca des sous-unités composant chaque modèle. Pour les modèles APO1 et APO2, le tableau donne les valeurs de déviation moyenne (en Å.) résultant de la superposition de toutes les unités considérées deux à deux. Pour le modèle ECAP, comportant 16 monomères, seule la valeur moyenne et les valeurs extrêmes sont reportées .**

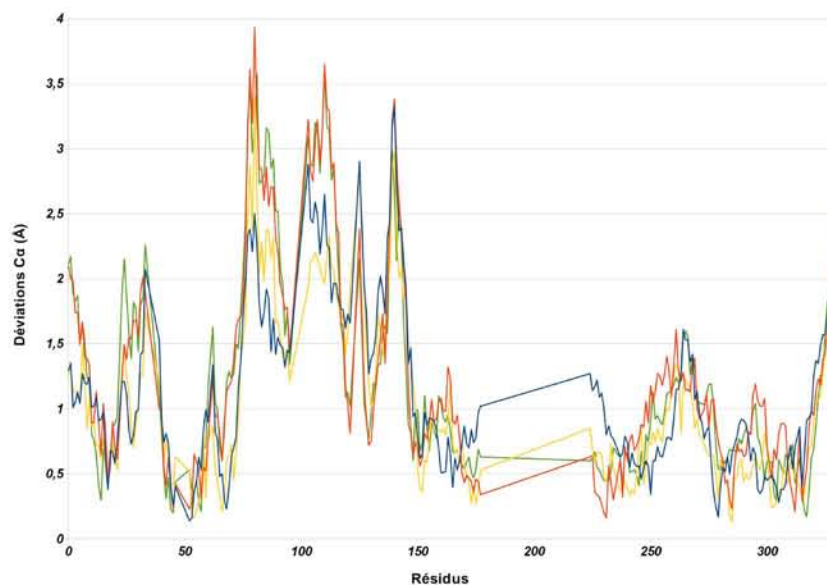
	<i>APO1</i>				<i>APO2</i>				<i>ECAP</i>
		B	C	D		B	C	D	
<i>Superposition brute</i>	A	0,47	0,64	0,68	A	0,65	0,50	0,57	0,34 (de 0,17 à 0,59)
	B	/	0,78	0,65	B	/	0,63	0,81	
	C	/	/	0,74	C	/	/	0,58	
<i>Superposition "fine"</i>		B	C	D		B	C	D	/
	A	0,42	0,30	0,43	A	0,47	0,36	0,31	
	B	/	0,48	0,31	B	/	0,37	0,45	
<i>Régions exclues des superpositions "fines"</i>	Boucle 34-38								/
	Boucle 47-51								
	Résidu Arg61								
	Boucle 96-102								
	Région 178-223 (incluant la boucle S incomplète dans la majorité des modèles)								
	Boucle 235-236								
	Résidu 301								
	Résidu 313								
Extrémité C-terminale (330-334)									

importante que les autres, ces valeurs de rmsd sont voisines des valeurs résultant des comparaisons entre monomères d'une même structure et suggèrent qu'il n'existe pas de différence majeure entre les monomères des modèles.

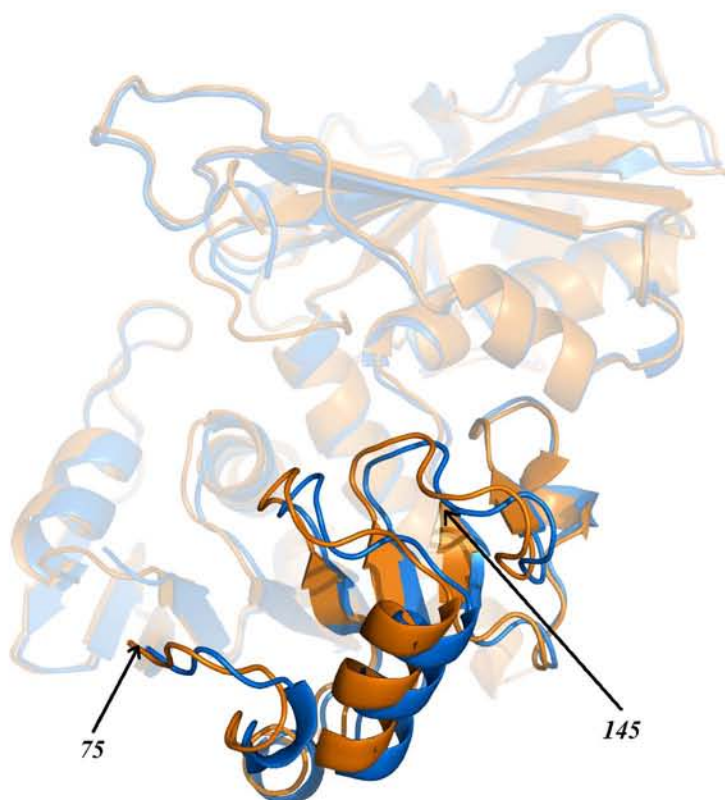
### ***A.3.3. Structure du tétramère de l'E4PDH d'E. coli***

Les structures tétramériques des trois modèles résolus ont été superposées en excluant toujours les régions décrites dans le Tableau IV-2. Dans chaque comparaison (APO1/APO2 ; APO1/ECAP ; APO2/ECAP), 4 superpositions ont été réalisées pour faire correspondre tour à tour le monomère A de la structure superposée aux monomères A, B, C et D de la structure de référence. Les rmsd moyens sont de 1,39, 0,72 et 1,05 Å respectivement pour les superpositions des tétramères APO1 sur APO2, APO1 sur ECAP et APO2 sur ECAP. Ainsi, il existe plus de différence structurale entre les deux structures sous forme apoenzyme qu'entre les structures sous forme apoenzyme et la structure du complexe avec le 3-CAPAD. Lorsque l'on superpose les tétramères APO1 et APO2 en considérant uniquement les domaines de fixation du cofacteur ou les domaines catalytique des quatre sous-unités, les rmsd moyens prennent alors des valeurs de 1,69 et 0,80 Å, respectivement (à comparer avec 1,39Å pour l'ensemble du tétramère). Ces valeurs indiquent que les différences entre les deux structures se situent essentiellement au niveau des domaines de fixation du cofacteur alors que les domaines catalytiques, constituant le coeur du tétramère (Figure IV-2), peuvent être considérés comme étant structurellement équivalents. De manière intéressante, le profil de déviation obtenu lors la superposition des tétramères APO1 et APO2 complets ressemble étroitement au profil obtenu lors de la comparaison des tétramères apoenzyme et holoenzyme de la *Bst*GAPDH, code pdb 2gd1 et 1gd1, respectivement (Figure IV-3, Skarzynski and Wonacott, 1988). Ainsi, les zones variant entre les deux structures concernent majoritairement, dans les cas de l'E4PDH comme de la GAPDH, le domaine de fixation du cofacteur et plus particulièrement la région constituée des résidus 75-145 (Figure IV-4). C'est cette région qui présente le mouvement de plus grande amplitude lors de la transition apo-holo chez la GAPDH de *B. stearothermophilus* (Skarzynski and Wonacott, 1988). Ces superpositions montrent que, dans le cas de l'E4PDH comme dans celui de la GAPDH, les domaines de fixation du cofacteur au sein du tétramère présentent une certaine mobilité par rapport aux domaines catalytiques.

La comparaison des structures des tétramères des trois formes de l'E4PDH obtenues avec le tétramère apoGAPDH de *B. stearothermophilus* (Skarzynski and Wonacott, 1988) conduit à l'obtention de rmsd moyens de 2,63, 1,72 et 2,31 Å, respectivement pour les formes APO1, APO2



**Figure IV-3 : Profils de déviations moyennes de la position des atomes de Ca découlant de la superposition des tétramères APO1 et APO2 de l'E4PDH d'E. coli. Les courbes en bleu, rouge, jaune et vert représentent la superposition des sous-unités A sur A, B sur B, C sur C et D sur D.**



**Figure IV-4 : Mise en évidence de la variabilité conformationnelle de la région 75-145 entre les formes APO1 et APO2 de l'E4PDH. Pour réaliser la figure, les tétramères APO1 (bleu) et APO2 (orange) ont été superposés. La région 75-145 qui présente la plus grande déviation entre les deux structures (cf. Figure IV-3) est représentée en couleur pleine au sein du monomère représenté en transparence.**

et ECAP. Leurs comparaisons au tétramère d'holoGAPDH conduisent à des valeurs de rmsd moyens plus élevées de 3,40, 2,30 et 3,02 Å, respectivement. Ainsi, les trois structures d'E4PDH obtenues sont toutes plus proches de la forme apoenzyme de la *Bst*GAPDH que de sa forme holoenzyme.

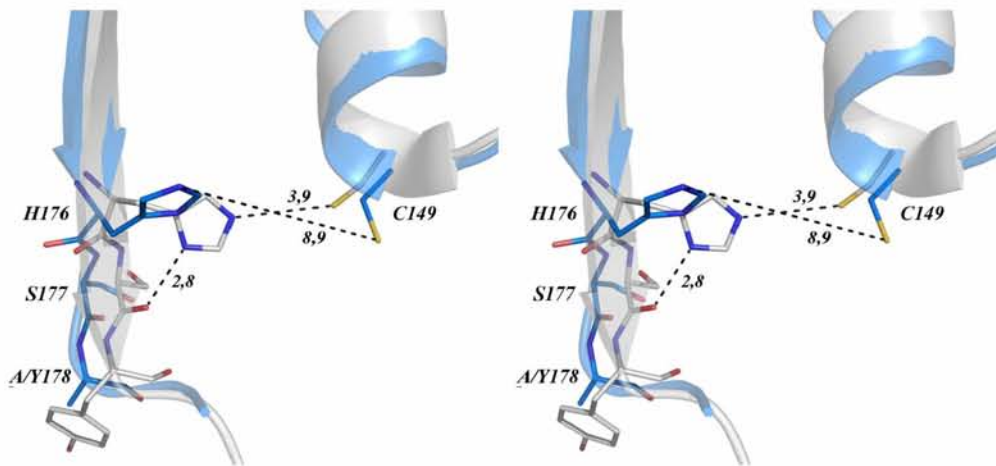
## ***A.4. Analyse des sites fonctionnels de l'E4PDH d'E. coli***

### ***A.4.1. Conformation du site actif***

Trois structures de l'E4PDH ont été obtenues et présentent trois états distincts du site actif de l'enzyme. Le modèle APO1 correspondant à l'enzyme de type sauvage contient les deux résidus catalytiques, Cys149 et His176 et a donc été principalement utilisé pour décrire la conformation du site actif de l'E4PDH d'*E. coli*. En effet, le modèle APO2 correspond à la structure d'un double mutant de l'enzyme dans lequel la cystéine catalytique, d'une part, a été mutée en résidu alanine, et le résidu Cys153, d'autre part, a été muté en sérine afin d'empêcher la formation d'un pont disulfure Cys153-Cys311 (*cf.* en A.1.). Enfin, le modèle en complexe avec le 3-CAPAD, obtenu avec l'enzyme de type sauvage, présente la cystéine catalytique engagée dans une liaison covalente avec l'inhibiteur. Il n'est donc pas représentatif d'un état physiologique, et n'a pas été exploité pour l'analyse du site actif. Dans toutes ces structures, les résidus catalytiques (Cys149 et His176) ou en position catalytique (Ala149 dans APO2), ainsi que leur environnement proche, sont parfaitement définis dans les cartes de densité électronique.

Dans le modèle APO1, les conformations de la cystéine et de l'histidine catalytique sont comparables dans les différentes sous-unités (excepté dans la sous-unité A où une rotation de 90° autour de la rotule  $\chi_1$  de la Cys149 amène l'atome S $\gamma$  de sa chaîne latérale vers le centre du site actif (non montré)). La géométrie du site actif est équivalente à celle observée dans la structure de l'E4PDH d'*E. coli* résolue à 2,7 Å (Kauffmann, 2000). La cystéine et l'histidine catalytiques ne sont pas orientées l'une vers l'autre et la distance séparant leurs atomes respectifs, S $\gamma$  et N $\epsilon_2$ , varie entre 6,9 et 8,8 Å (Figure IV-5). Par rapport à la GAPDH, dans laquelle la cystéine et l'histidine catalytiques sont déjà parfaitement orientées l'une par rapport à l'autre dans la forme apoenzyme (Figure IV-5) le site actif de l'E4PDH ne semble pas dans une conformation adéquate pour la formation d'une paire d'ions, comme c'est le cas dans la GAPDH. Dans la GAPDH, la chaîne latérale de l'His176 est en conformation *trans*, et la formation d'une liaison hydrogène entre le





**Figure IV-5 : Comparaison de la conformation des résidus catalytiques dans le site actif des structures APO1 de l'E4PDH d'*E. coli* et apoGAPDH de *B. stearotherophilus* (Skarzynski and Wonacott, 1988 ; 2gd1). La sous-unité O de la structure apoenzyme de la GAPDH (en blanc) a été superposée à la sous-unité B de la structure de l'E4PDH (en bleu). Les résidus catalytiques Cys149 et His176 de la GAPDH et de l'E4PDH ainsi que les résidus en position 177 et 178 ont été représentés sous forme de bâtons et annotés. Les distances (en Å) séparant les atomes S $\gamma$  et Ne $_2$  respectivement de la Cys149 et de l'His176 dans les deux enzymes, matérialisées par des pointillés, ont été indiquées. La liaison hydrogène existant dans la GAPDH entre les résidus 176 et 177 est également représentée par des pointillés.**

groupement carbonyle du résidu Ser177 (O) et le cycle imidazole de l'His176 (N<sub>δ1</sub>) empêche la rotation autour de  $\chi_2$  en fixant le plan du noyau imidazole. Dans la structure APO1 de l'E4PDH, la chaîne latérale de l'His176 est dans une conformation totalement différente, gauche (+) équivalente dans les quatre sous-unités du modèle. Cette orientation différente de la chaîne latérale de l'His176 pourrait s'expliquer par l'absence d'interaction avec la Ser177. En effet, bien que le résidu Ser177 soit également présent dans l'E4PDH, la conformation de la chaîne principale autour de ce résidu est très différente de celle rencontrée dans la GAPDH, et ne permet pas à la Ser177 d'interagir avec l'His 176 dans l'E4PDH (Figure IV-5). Ceci est sans doute lié au fait que ce résidu appartient à l'extrémité N-terminale de la boucle S (181-204) qui est particulièrement flexible dans l'E4PDH et non définie au delà du résidu Met179 dans ce modèle.

L'analyse de la conformation du site actif du modèle APO2 (double mutant C149A/C153S) montre que la mutation de la cystéine catalytique comme de la Cys153, située dans l'environnement proche de la première, n'induit pas de changement conformationnel notable sur le site actif. En effet, seul un léger décalage (<1 Å) de l'hélice 149-165 est observé, vraisemblablement en relation avec l'absence du pont disulfure dans le double mutant par rapport à l'enzyme de type sauvage. La conformation du résidu His176 est équivalente à celle précédemment décrite pour le modèle APO1.

## ***A.4.2. Le site de fixation du cofacteur***

Le site de fixation du cofacteur a été analysé principalement sur la base du modèle ECAP dans lequel les interactions enzyme/inhibiteur ont été exploitées pour apporter des informations sur les interactions possibles entre l'enzyme et le cofacteur.

### ***A.4.2.1. Le complexe E4PDH - 3(chloroacétyl) pyridine adénine dinucléotide (ECAP)***

Dans la majorité des sous-unités du modèle ECAP (14 sur 16 au total), un signal fort de densité positive présent dans les cartes Fourier différence au niveau de la poche de fixation du cofacteur a permis de modéliser la molécule de 3-(chloroacétyl) pyridine adénine dinucléotide (3-CAPAD). Notre analyse s'est focalisée sur les 8 sous-unités (A, C, E, G, J, L, N et P) pour lesquelles la qualité des cartes de densité électronique nous autorise à considérer le modèle avec confiance. En effet, la densité électronique est de moindre qualité dans les 6 autres monomères où l'inhibiteur a été modélisé (B, F, H, I, K et M), et dans les deux sous-unités D et O, le signal de densité électronique, n'est observable qu'au niveau du pyrophosphate, et n'a pu permettre de

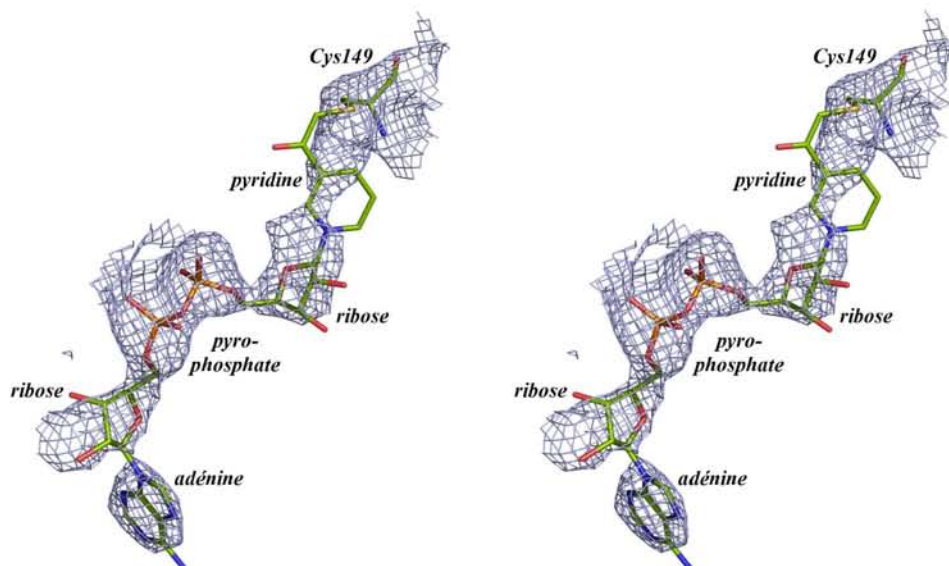
modéliser la molécule de 3-CAPAD. Dans chaque tétramère (4 dans l'unité asymétrique), la molécule de 3-CAPAD est bien définie et donc bien fixée dans deux monomères sur quatre. Les deux monomères en question sont dans tous les cas des sous-unités reliées par l'axe binaire  $Q$ .

La Figure IV-6 présente la molécule de 3-CAPAD, en liaison avec la cystéine catalytique, dans une carte de densité électronique de type "omit". Le signal de densité observé dans la continuité de la cystéine catalytique (Cys149) implique clairement ce résidu dans une liaison covalente avec l'inhibiteur. Cependant, le signal est trop faible et ambigu pour positionner explicitement le noyau pyridine, qui adopte plusieurs positions dans les différents monomères (Figure IV-7). Les groupements ribose adjacent au noyau pyridine, pyrophosphate et adénosine sont beaucoup mieux définis dans la densité électronique. Ceci indique que cette partie de la molécule de 3-CAPAD, bien stabilisée dans la structure, est potentiellement représentative du site de fixation du cofacteur physiologique. L'analogue de cofacteur développe 5 à 6 liaisons hydrogène avec la protéine : le groupement pyrophosphate avec la chaîne principale des résidus Arg10 et Ile11, les riboses adjacents à l'adénine et au nicotinamide avec les chaînes latérales respectives des résidus Glu32 et Ser119, et dans certaines sous-unités la chaîne latérale de la Tyr99 avec l'adénine. Selon la sous-unité considérée, une à trois molécules d'eau interagissent également avec le 3-CAPAD (Figure IV-8, Tableau IV-3 pour les distances interatomiques). Des interactions de van der Waals, de part et d'autre du noyau adénine impliquent par ailleurs les chaînes latérales des résidus Ile33 et Thr96.

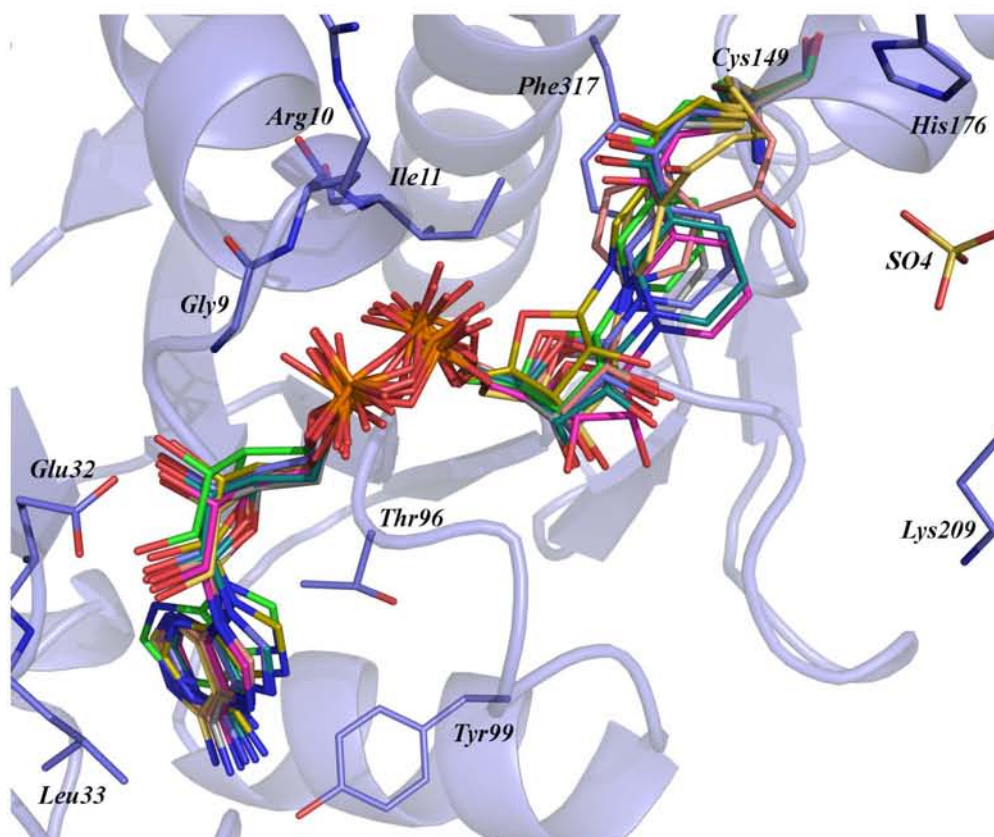
La comparaison de la poche de fixation du cofacteur de ce complexe avec les poches de fixation des formes APO1 et APO2 ne met en évidence aucune modification notable de la conformation des résidus, seules les chaînes latérales des résidus Ile33 et Thr96 semblent adopter une conformation plus stable, vraisemblablement suite à la fixation de l'adénine.

#### *A.4.2.2. Comparaison avec le site de fixation du cofacteur de la GAPDH de *B. stearothermophilus**

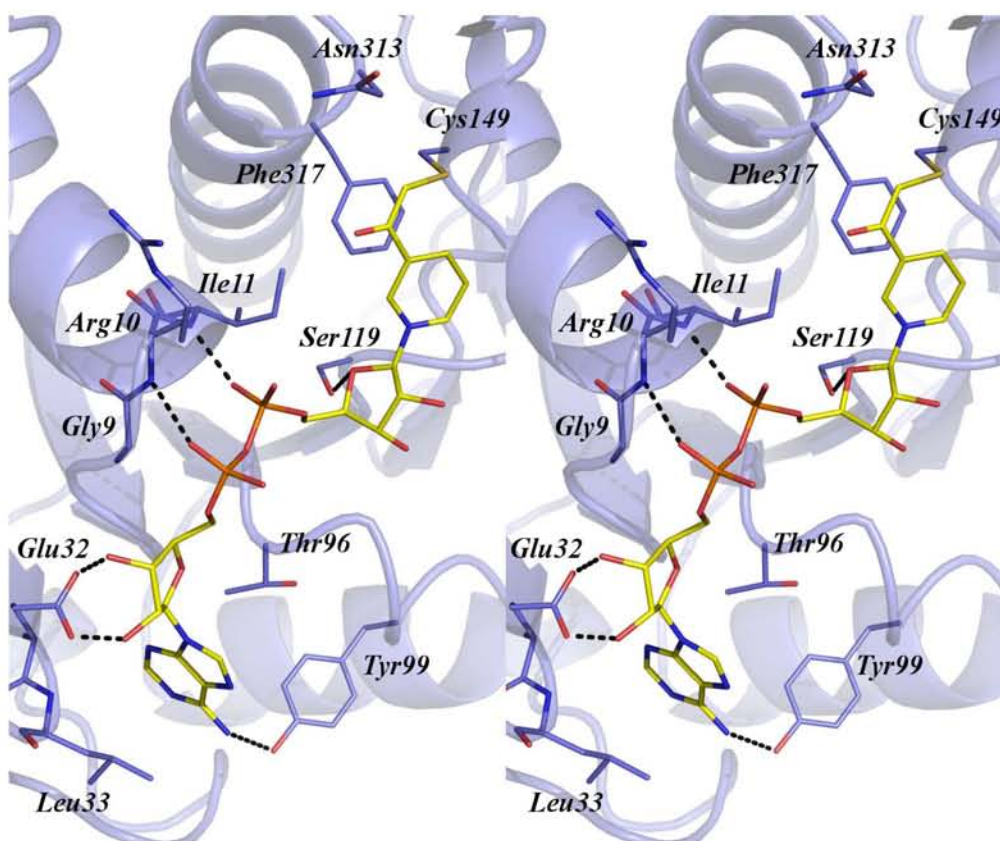
Mis à part le noyau pyridine, la position et la conformation de la molécule de 3-CAPAD, après superposition du monomère à la GAPDH sous forme holoenzyme (Skarzynski *et al.*, 1987, 1gd1), sont globalement très similaires à celles de la molécule de NAD<sup>+</sup> fixée dans la GAPDH (Figure IV-9). En effet, les atomes de l'inhibiteur et du NAD se superposent avec un rmsd de 0,68 Å alors que cette valeur tombe à 0,25 Å lorsque les atomes du noyau pyridine sont écartés du calcul.



**Figure IV-6 :** *Vue stéréoscopique de la molécule de 3-CAPAD en liaison covalente avec le résidu Cys149 (sous-unité G). La molécule de 3-CAPAD et le résidu Cys149 ont été retirés du modèle qui a ensuite été soumis à un affinement avec Refmac5. La carte de densité électronique ( $2mF_o-DF_c$ ) "omit" qui résulte de cet affinement est représentée à un niveau de contour de 1 sigma.*



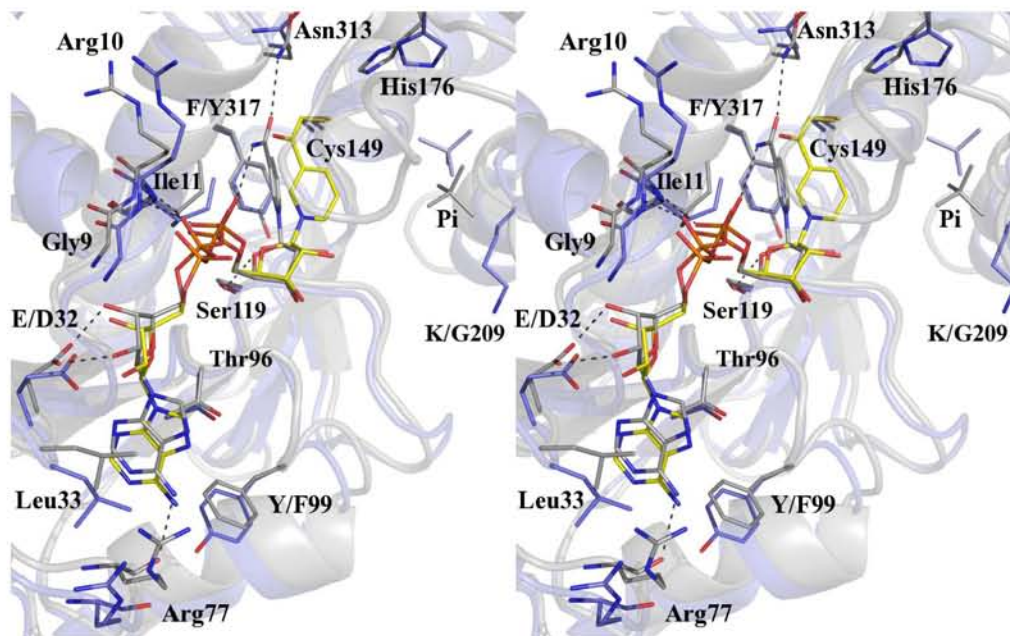
**Figure IV-7 :** *Superposition des molécules de 3-CAPAD. Les monomères A (vert), C (fuchsia), E (rose), G (bleu), J (vert), L (jaune), N (blanc) et P (or) ont été superposés sur la base de leur Ca.. Les 8 molécules de 3-CAPAD sont représentées dans la sous-unité G (en bleu) dont les résidus du site de fixation du cofacteur et du site actif, représentés sous forme de bâtons, ont été annotés.*



**Figure IV-8 :** *Vue stéréoscopique du site de fixation de la molécule de 3-CAPAD. Le monomère G (en bleu) a été représenté en mode ruban avec en mode bâton les résidus du site de fixation du cofacteur. Les liaisons hydrogène sont représentées par des pointillés.*

**Tableau IV-3 :** *Distances interatomiques des liaisons hydrogène entre le 3-CAPAD et l'E4PDH.*

	A	C	E	G	J	L	N	P
O2A --- Arg10(N)	2,65	3,0	2,9	2,85	2,6	2,8	3,1	3,0
O2N --- Ile11(N)	2,9	3,3	3,0	2,5	2,8	3,0	3,0	2,8
O2B --- Glu32(Oδ1)	2,4	2,9	2,8	2,3	2,5	2,9	2,6	/
O3B --- Glu32 (Oδ2)	/	3,3	2,8	2,7	2,9	2,8	2,8	2,6
O4D --- Ser119(Oγ)	/	3,2	3,2	3,3	/	3,3	2,9	/
N6A --- Tyr99(Oθ)	3,8	3,8	3,5	3,6	3,3	3,8	3,8	3,1
	2,9 (O1A)	3,2 (O3B)	3,1 (O1A)	3,4 (O2A)	2,6 (O3)		2,5 (O2B)	2,9 (O3A)
Molécule(s) d'eau (atome 3-CAPAD)	2,7 (O3B)	2,8 (O2A)					2,4 (O1A)	3,4 (O1N)
							2,8 (O7N)	

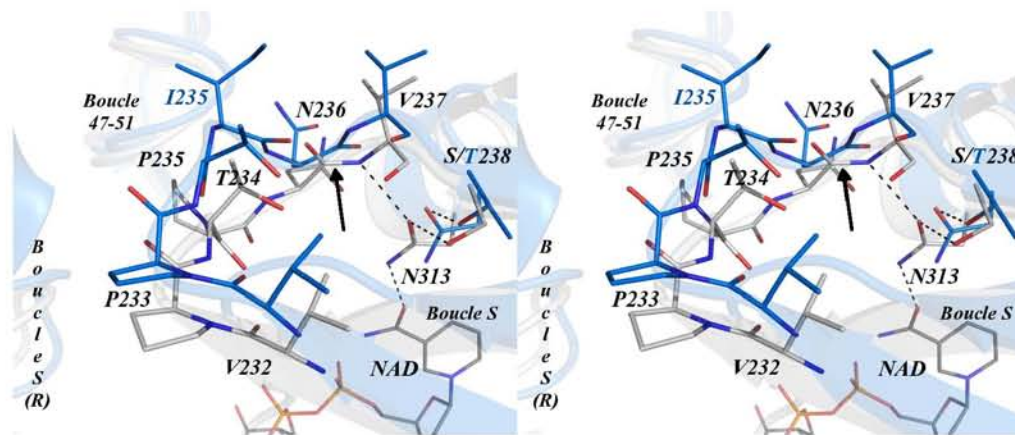


**Figure IV-9 : Comparaison des sites de fixation du 3-CAPAD et du NAD.** Le monomère G de la structure du complexe ECAP (bleu, avec le 3-CAPAD en jaune) a été superposé au monomère Q de la BstGAPDH sous forme holoenzyme (Igd1, blanc). Les résidus du site actif comme ceux impliqués dans la reconnaissance et la fixation de la molécule de cofacteur sont représentés sous forme de bâtons et annotés. Le site Pi (sulfate) de la GAPDH est indiqué. Lorsque les résidus de l'E4PDH et de la GAPDH ne sont pas les mêmes, le code à 1 lettre a été utilisé en donnant en premier le résidu de la séquence de l'E4PDH. Les liaisons hydrogène formées entre le NAD<sup>+</sup> et la GAPDH sont indiquées en pointillés.

Étant donné que la poche de fixation du noyau pyridine se superpose à celle observée dans la GAPDH de *B. stearrowthermophilus*, la non-superposition des noyaux pyridine peut s'expliquer par les contraintes imposées par la liaison covalente formée entre la cystéine catalytique et le noyau acétyl pyridine. En effet, les résidus Ile11 et Phe317 (Tyr317 dans la séquence de la GAPDH), qui stabilisent le noyau pyridinium dans la GAPDH, sont dans des conformations identiques dans les deux enzymes.

Le réseau de liaisons hydrogène stabilisant le NAD dans la structure de la GAPDH est presque entièrement retrouvé dans le complexe E4PDH/3-CAPAD. Ainsi, la chaîne principale des résidus Arg10 et Ile11, respectivement en position 4 et 5 du motif G/AXG/AXXG caractéristique du repliement de Rossmann (Rossmann *et al.*, 1974), forme deux liaisons hydrogène avec le groupement pyrophosphate du 3-CAPAD. Comme dans le cas de la GAPDH, la chaîne latérale du résidu Ser119 forme une liaison hydrogène avec le ribose adjacent au nicotinamide.

Le nicotinamide, dans les structures de GAPDH, est stabilisé au travers d'une liaison hydrogène avec le résidu Asn313. Le positionnement adéquat de la chaîne latérale de ce résidu est obtenu grâce à l'établissement de deux liaisons hydrogène avec la chaîne latérale de la Ser238 (O $\gamma$ ) et avec l'atome d'azote de la chaîne principale de la Val237 (Figure IV-10). Dans le cas de l'E4PDH, la chaîne latérale de la Thr238, joue un rôle équivalent à celui de la Ser238 des GAPDH. En revanche, une certaine variabilité conformationnelle existe autour du résidu en position 237 (Figure IV-10). La Val237 n'adopte une conformation équivalente à celle rencontrée chez les GAPDH (*i.e.* en dehors des régions admises du diagramme de Ramachandran (Duée *et al.*, 1996)) que dans un seul monomère parmi les 24 observés dans les différentes structures d'E4PDH résolues au cours de cette étude, où elle interagit avec le résidu Asn313, lui assurant un positionnement idéal à la fixation de la fonction carboxamide du nicotinamide. Dans les 23 autres monomères, la conformation de la boucle 233-237 est différente de celle rencontrée chez les GAPDH, et confère à la Val237 une conformation autorisée dans le diagramme de Ramachandran. Cette conformation, qui inverse l'orientation de la liaison peptidique 236-237 (Figure IV-10), est incompatible avec la stabilisation de l'Asn313, dont la chaîne latérale montre une conformation moins stable, en relation avec le signal plus diffus de sa densité électronique. Il est intéressant de noter également que cette boucle 233-237 est située au coeur du tétramère, dans l'environnement direct des boucles S et 47-51 (Figure IV-10), dont l'absence de structuration et/ou la mobilité ont déjà été évoquées précédemment et seront développées ci-dessous. La variabilité conformationnelle de la boucle 233-237 est facilitée par le remplacement du résidu Pro235 des GAPDH par une isoleucine invariante dans les E4PDH (Figure IV-10).



**Figure IV-10 : Comparaison de la conformation des boucles 233-237 de la BstGAPDH (blanc) et de l'E4PDH d'E. coli (bleu).** Les résidus de cette boucle, les résidus 313 et 238 ainsi que le NAD fixé dans la GAPDH sont représentés sous forme de bâtons. Les boucles mobiles environnantes, boucle S et 47-51 de la même sous-unité ainsi que la boucle S de la sous-unité reliée par l'axe binaire R sont également représentées sous forme de ruban. Noter l'inversion de l'orientation de la liaison peptidique 236-237 indiquée par la flèche.



La chaîne latérale acide du résidu Asp32 de la GAPDH est remplacée, dans les séquences d'E4PDH, par un groupement propionate conservé. La chaîne latérale de ce résidu Glu32 et celle du résidu Asp32 stabilisent le ribose adjacent à l'adénine au travers de deux liaisons hydrogène respectivement avec les atomes d'oxygène 2' et 3'. Il est important de préciser que le résidu Asp32 n'a pas la capacité de stabiliser le ribose dans la forme apoenzyme de la GAPDH et que la transition apo-holo est nécessaire afin que ce résidu participe à la fixation du cofacteur. Le groupement carbonyle en position 77 (Arg), qui vient stabiliser le cycle adénine dans la structure de la GAPDH à la suite de la transition apo-holo, est dans ECAP trop éloigné pour former une telle liaison. Cependant, cette interaction du cycle de l'adénine semble être remplacée, dans le complexe ECAP, par une interaction avec la chaîne latérale du résidu Tyr99 (Figure IV-8 et IV-9). Ce résidu, majoritairement conservé dans les séquences d'E4PDH, correspond, dans les séquences de la GAPDH, à une phénylalanine qui ne peut développer une telle interaction. La région 179-190 de l'E4PDH n'étant pas définie dans les cartes de densité électronique de cette structure, aucune liaison équivalente à celle formée entre le résidu Asn180 de la GAPDH (His180 dans la séquence de l'E4PDH) et le pyrophosphate du NAD n'a pu être identifiée dans la structure du complexe E4PDH/3-CAPAD. Enfin, dans la GAPDH, 8 molécules d'eau, stabilisées par des liaisons hydrogène avec la protéine, participent à la fixation du cofacteur (Figure I-5). Dans la structure de l'E4PDH, ces molécules d'eau ne sont pas observées, vraisemblablement en raison de la faible résolution de la structure.

### ***A.4.3. Les sites de reconnaissance anionique***

Deux structures ont été obtenues en présence d'un anion susceptible d'occuper un site de reconnaissance anionique de l'E4PDH. Il s'agit de la structure APO1 cristallisée en présence de phosphate et de la structure ECAP obtenue en présence de sulfate. Dans les deux cas, des pics de densité positive observés dans les cartes Fourier différence nous ont conduits à y modéliser les anions contenus dans le milieu de cristallisation.

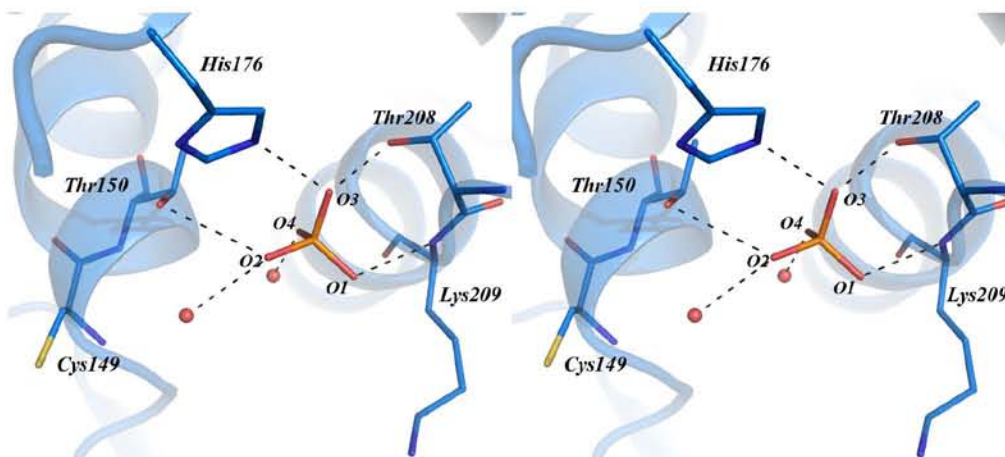
#### ***A.4.3.1. Site anionique du modèle APO1***

Un ion phosphate a été modélisé dans chacune des quatre sous-unités du modèle APO1. Le facteur B moyen de ces anions est de 25,0 Å<sup>2</sup> (à comparer au B moyen du modèle à 35,9 Å<sup>2</sup>) et leur facteur d'occupation estimé à 0,5 dans les 4 sous-unités du modèle. La position des anions phosphate et les interactions qu'ils développent avec la chaîne protéique sont équivalents dans les

différentes sous-unités. Ce site de reconnaissance anionique implique les chaînes latérales des résidus Thr150, His176 et Thr208, l'atome d'azote de la chaîne principale du résidu Lys209 et une à trois molécules d'eau en fonction du monomère considéré (Figure IV-11 et Tableau IV-4 pour les distances interatomiques). Par analogie structurale, on peut rapprocher la position de ce site de reconnaissance anionique avec les sites appelés Pi ou nouveau Pi des GAPDH phosphorylantes. Même si la localisation exacte du site n'est pas parfaitement identique à ceux rencontrés dans les GAPDH (position de l'atome de phosphore à environ 3,0 et 2,5 Å, respectivement, de la position caractéristique des sites Pi et nouveau Pi), il est néanmoins très proche de ceux-ci et implique globalement les mêmes résidus d'acides aminés (Figure IV-12). Les différences majeures concernent en premier lieu la participation, dans l'E4PDH, du résidu His176 dont la conformation particulière a déjà été décrite (*cf.* en A.4.1.). La lysine en position 209, conservée dans les séquences d'E4PDH, qui correspond à une glycine, elle-même invariante, dans les GAPDH, ne gêne en aucun cas la participation de l'atome d'azote de sa chaîne principale à la formation du site de reconnaissance anionique, sa chaîne latérale pointant en direction du solvant. La position du site de reconnaissance anionique assimilé à l'atome de phosphore central est située à une distance de 6,9, 8,3, 7,5 et 7,9 Å de l'atome S $\gamma$  de la Cys149 dans les sous unité A, B, C et D, respectivement. Pour comparaison, les distances S $\gamma$ -phosphore, caractéristiques des sites Pi et nouveau Pi de la GAPDH, sont respectivement de 6,9 et 4,3 Å.

#### *A.4.3.2. Site anionique du modèle ECAP*

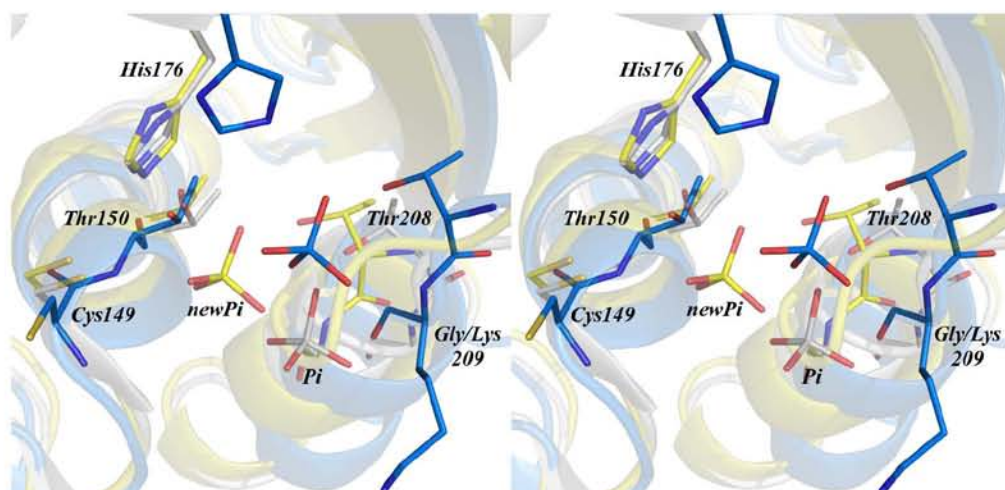
Les 16 sous-unités du modèle 3-CAPAD contiennent, au niveau de leur site actif, un ion sulfate dont l'occupation a été fixée à 1 et dont le B moyen est de 79,7 Å<sup>2</sup>. Cette valeur assez élevée, comparée au B moyen du modèle à 45,7 Å<sup>2</sup>, semble indiquer soit que l'occupation de ces anions est surestimée soit que l'anion est assez agité, les deux possibilités n'étant pas mutuellement exclusives. Même si la résolution des données de diffraction limite l'utilisation des cartes Fourier différence, aucun signal de densité négative n'apparaît néanmoins dans celles-ci au niveau des anions et leurs taux d'occupation ont ainsi été conservés à une valeur de 1. La position du site est pratiquement équivalente dans les différentes sous-unités avec un écart maximum de 1 Å entre la position des atomes de phosphore après superposition des monomères. La position du site tout comme les résidus participant à sa formation sont totalement équivalents à ceux décrits dans le cas du modèle APO1. Ceci indique, par ailleurs, que contrairement à l'hypothèse émise dans le cas de la GAPDH (Kim *et al.*, 1995), la nature de l'anion (sulfate ou phosphate) n'influe pas sur la position du site de reconnaissance de l'E4PDH.



**Figure IV-11 : Vue du réseau de liaisons hydrogène du site de reconnaissance anionique de l'E4PDH apoenzyme (APO1 monomère B).** Les résidus catalytiques ou impliqués dans la formation du site de reconnaissance anionique, comme les atomes du phosphate inorganique sont annotés et les liaisons hydrogène symbolisées par des pointillés.

**Tableau IV-4 : Distances interatomiques des liaisons hydrogène du site de reconnaissance anionique de l'E4PDH (APO1).**

(en Å)	A	B	C	D
O1 --- Lys209(N)	2,7	3,0	2,9	/
O2 --- Thr150(O $\gamma_1$ )	/	3,6	3,5	2,75
O2 --- H <sub>2</sub> O	/	3,4	2,5	2,4 - 2,5
O3 --- Thr208(O $\gamma_1$ )	2,9	2,7	2,5	2,7
O3 --- His176(N $\epsilon_2$ )	3,4	3,1	3,5	3,5
O4 --- H <sub>2</sub> O	2,3	2,5	/	3,6
				Lys209(N)



**Figure IV-12 : Comparaison de la position des sites de reconnaissance anionique Pi et "new Pi" de la GAPDH avec le site identifié dans la structure APO1 de l'E4PDH d'E. coli.** Les structures de la GAPDH de *T. maritima* (1hdg, jaune) et de *B. stearothermophilus* (1gd1, blanc) ont été superposées au modèle APO1(bleu). Les résidus impliqués dans la formation des sites de reconnaissance anionique sont annotés.

### *A.4.3.3. Recherche d'un site équivalent au site Ps des GAPDH*

Dans la forme apoenzyme de la GAPDH de *B. stearothermophilus*, le site Ps est partiellement occupé par un ion sulfate qui est stabilisé par l'intermédiaire des chaînes latérales des résidus Thr179 et Arg231. Dans la forme holoenzyme, où le site Ps est totalement occupé par un ion sulfate, la molécule de cofacteur participe également à la formation de ce site par la fonction 2'OH du ribose du nicotinamide.

Aucune des structures d'E4PDH obtenues au cours de ces travaux ne présente un anion qui pourrait occuper une localisation équivalente au site Ps des GAPDH, conduisant à la conclusion qu'il n'existe dans cette enzyme, aucun autre site « préformé » susceptible d'accueillir un groupement phosphate. Tout d'abord, le fait qu'une méthionine soit présente en position 179 dans l'E4PDH (au lieu d'une thréonine dans la GAPDH) exclut la participation de ce résidu à la formation d'un site de reconnaissance anionique. L'étude structurale présentée ici révèle qu'aucun autre résidu ne se substitue à la Met179 pour participer à la fixation d'un groupement phosphate. Cependant, la conformation très variable de ce résidu, liée à la grande flexibilité de la boucle S, ne semble pas non plus faire de ce résidu un antidéterminant de la formation d'un site Ps. En ce qui concerne la position 231, occupée par une arginine ou une lysine dans les séquences d'E4PDH, l'examen des 3 modèles de l'E4PDH montre que la chaîne latérale du résidu Arg231 n'est généralement pas définie par les cartes de densité électronique. Par ailleurs, lorsqu'elle est définie, celle-ci pointe le plus souvent dans une direction incompatible avec une participation éventuelle à la formation d'un site Ps équivalent à celui des GAPDH. La flexibilité de cette chaîne latérale est très probablement liée à l'absence de structuration de la boucle S (*cf.* en A.5.1.). En effet, la conformation de l'Arg231, dans la GAPDH, est maintenue grâce à une interaction avec le résidu Gln182 de la boucle S. Bien que le résidu Gln182 soit conservé dans la séquence de l'E4PDH, l'absence de structuration de la boucle S ne permet pas une telle interaction.

## ***A.5. Les boucles de structure variable***

### ***A.5.1. Conformation de la boucle S (181-204)***

#### ***A.5.1.1. Préambule***

Comme précédemment évoqué lors de la présentation des structures, la boucle S (181-204) de l'E4PDH n'est généralement pas complètement définie dans les cartes de densité électronique. Ainsi, dans le modèle APO1, les résidus 181 à 189 n'ont pas été modélisés. C'est également le cas pour les résidus 186-191 du monomère A et 181-183 du monomère D dans le modèle APO2. En revanche, pour les sous-unités B et C, la densité électronique était d'une qualité suffisante pour reconstruire la totalité de la boucle et cette structure offre donc pour la première fois la possibilité d'étudier cette région particulièrement agitée de l'E4PDH d'*E. coli*. Dans la structure ECAP, la boucle n'est pas complètement définie et la région non modélisée s'étend selon les monomères des résidus 179 à 190/192, voire jusqu'à la position 196 pour la sous-unité H.

#### ***A.5.1.2. Conformation de la boucle S (181-204) dans le modèle APO2***

Dans les GAPDH, la conformation des boucles S (181-204) est identique dans toutes les sous-unités du tétramère. Cette boucle est située majoritairement à l'interface entre deux monomères reliés par l'axe binaire R (c'est-à-dire soit à l'interface des monomères A et D soit à l'interface des monomères B et C). A cette interface, les deux boucles S, très proches de l'axe binaire R, interagissent ensemble *via* des interactions de van der Waals (entre les deux résidus Ile185 notamment et les deux résidus Ala200 dans le cas de la GAPDH de *B. stearothermophilus* ou entre les deux résidus Thr185 et Ser200 dans le cas de la GAPDH d'*E. coli*). De plus dans la forme holoenzyme de la *Bst*GAPDH, la chaîne latérale des résidus Asn180 établit une liaison hydrogène avec les groupements carbonyle des résidus Asp186.

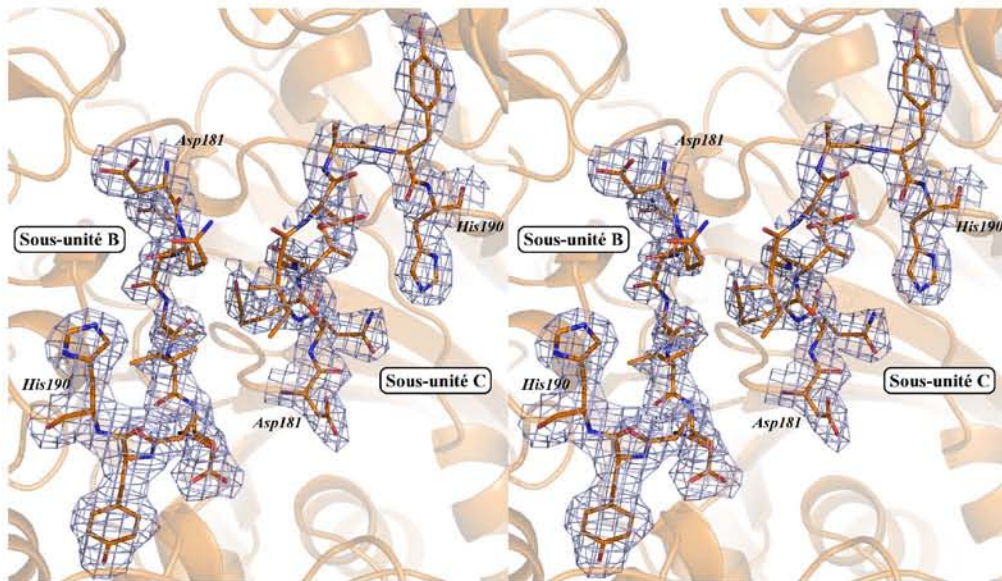
Dans le cas de l'E4PDH, la superposition des quatre sous-unités du modèle APO2 montre que la région 192-204 (C-terminale) de la boucle S adopte une conformation identique dans les monomères B, C et D. Elle adopte cependant une conformation différente dans la sous-unité A, en

relation directe avec la conformation particulière de la boucle 47-51 adjacente du monomère D qui sera discutée ci-dessous (Figure IV-13). La partie N-terminale de la boucle S (179-191) des monomères B et C, entièrement visible dans les cartes de densité électronique (Figure IV-14), adopte, en revanche, une conformation variable en fonction du monomère considéré (Figure IV-13).

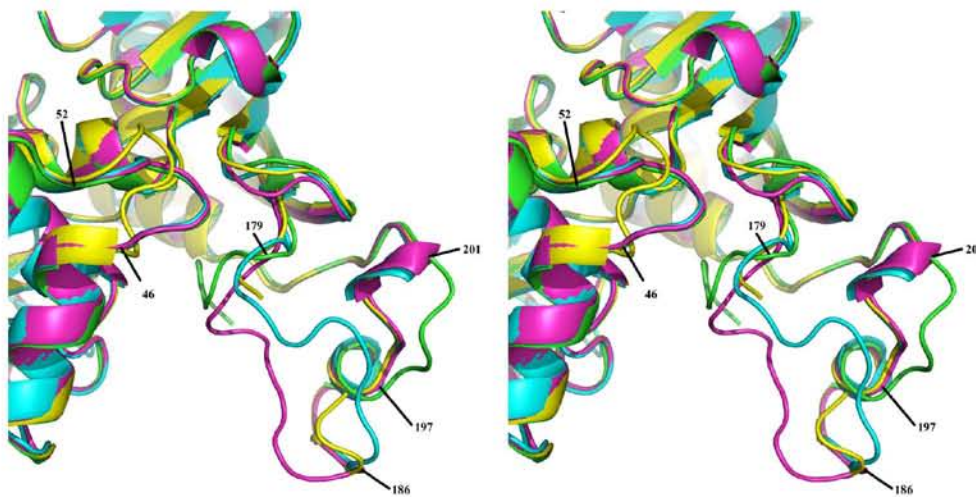
Lorsque l'on compare la structure de la boucle S de l'E4PDH avec celle de la GAPDH, il apparaît que, dans leur partie C-terminale, les boucles S des deux enzymes sont structurellement très proches (Figures IV-15 et IV-16). En effet, la distance séparant les atomes de carbone  $\alpha$  équivalents des boucles S de l'E4PDH et de la *Bst*GAPDH après superposition des tétramères est inférieure à 3 Å sur toute la partie C-terminale de la boucle. Par contre, la région N-terminale de la boucle semble beaucoup plus variable, avec des distances entre atomes de  $C\alpha$  équivalents allant à plus de 9 Å au niveau du résidu Leu185 (monomère C) (Figure IV-16). Il est intéressant de noter que, dans les GAPDH, c'est à travers cette région N-terminale (179-191) que les boucles S interagissent à l'interface de deux monomères. L'analyse de la séquence de l'E4PDH indique la présence, en position 185, d'un résidu leucine qui pourrait parfaitement établir le même type d'interaction de van der Waals que ces homologues Ile ou Thr rencontrés respectivement dans les séquences de *Bst*GAPDH et *Eco*GAPDH. La position 180, occupée par un résidu asparagine dans la *Bst*GAPDH et qui interagit avec le groupement carbonyle du résidu 186, est occupée dans la séquence de l'E4PDH d'*E. coli* par un résidu histidine. Ce résidu pourrait jouer un rôle d'antidéterminant stérique en prévenant le rapprochement des boucles S à l'interface entre les monomères. Néanmoins, cette position n'est pas strictement conservée dans les E4PDH puisque l'alignement de 58 séquences révèle la présence de 59 % d'asparagine et 41% d'histidine en cette position. Il serait intéressant d'étudier la structuration de la boucle S dans une E4PDH possédant une asparagine en position 180.

### ***A.5.2. La boucle 96-102***

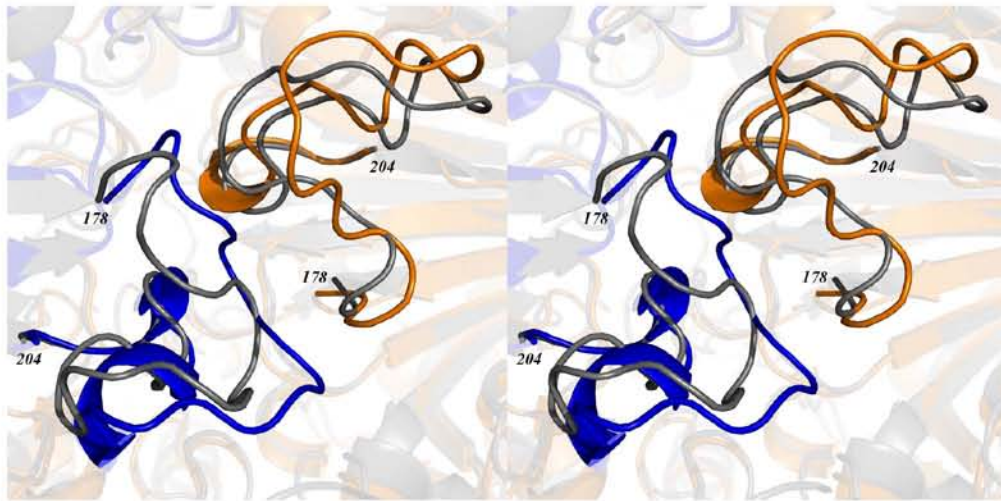
La boucle 96-102 présente dans la plupart des structures résolues dans cette étude (23 monomères sur 24) une conformation similaire à celle observée dans les structures de GAPDH. Dans cette conformation, elle participe à la stabilisation de l'adénine du cofacteur *via* la chaîne latérale du résidu Tyr99. Néanmoins, une conformation particulière est observée dans la sous-unité C du modèle APO1. Cette conformation ne semble cependant pas avoir un taux d'occupation de 1, des signaux résiduels dans les cartes différences (Fo-Fc) suggérant la présence de la conformation observée dans les autres monomères. En raison de la résolution moyenne de la structure, seule cette conformation particulière, majoritaire, a été modélisée. Elle se caractérise par un déplacement de la



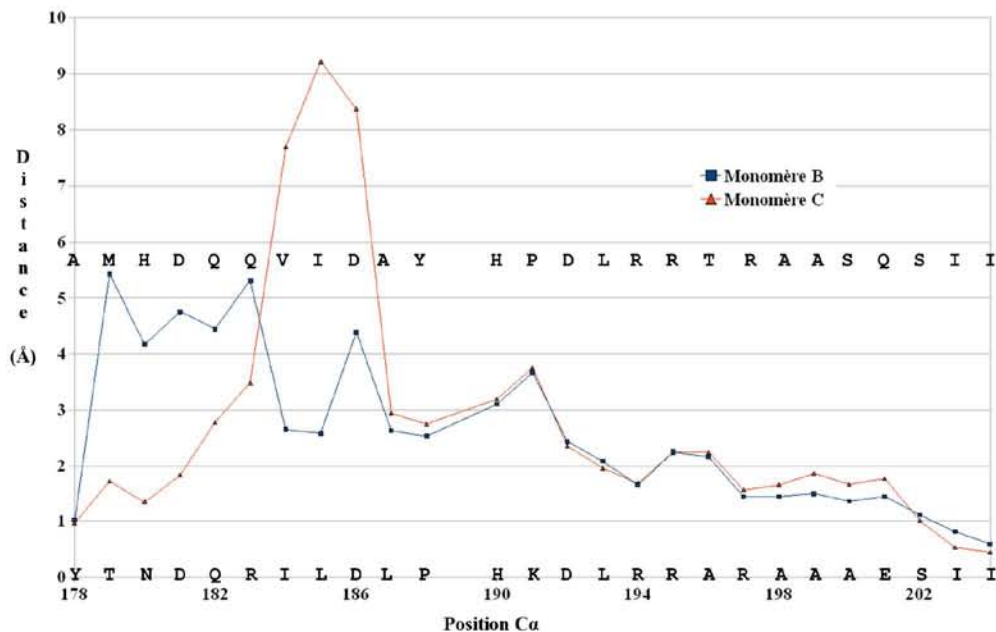
**Figure IV-13 :** *Vue stéréoscopique des boucles 181-190 des sous-unités B et C. Les boucles, de séquence DQQVIDAYH, sont représentées dans une carte de densité électronique de type 2mFo-DFc contournée à 1σ.*



**Figure IV-14 :** *Vue stéréoscopique de la superposition des boucles S (181-204) des quatre sous-unités du modèle APO2. Les monomères A, B, C et D sont respectivement représentés en vert, bleu, rose et jaune. Les positions 46, 52, 179, 186, 192 et 201 sont indiquées par des flèches.*



**Figure IV-15 :** *Vue stéréoscopique de la superposition des boucles S de l'E4PDH d'E. coli avec celles de la forme apoenzyme de la BstGAPDH (2gd1, Skarzynski and Wonacott, 1988). Le tétramère de la GAPDH (en gris) a été superposé sur le modèle APO2. Les régions 178-204 des sous-unités B et C de l'E4PDH sont colorées en bleu et orange, respectivement.*



**Figure IV-16 :** *Graphique représentant l'écart (en Å) entre la position des Ca des boucles S de l'E4PDH par rapport à la boucle S de la BstGAPDH sous forme apoenzyme (2gd1). Pour les mesures, les monomères B et C du modèle ont été superposés sur la sous-unité O de la GAPDH. Les séquences primaires des boucles S de la GAPDH et de l'E4PDH ont été indiquées selon le code à une lettre sur l'axe des abscisses et dans la zone de tracé, respectivement.*



chaîne principale d'une amplitude maximale d'environ 5 Å au niveau du résidu Tyr99 en direction du site actif (Figure IV-17). L'orientation de la chaîne latérale de la Tyr99 est complètement inversée pour venir former une liaison hydrogène avec la chaîne latérale de la Lys209. Ce mouvement crée un encombrement stérique vis-à-vis de la chaîne latérale de l'His120 dont la conformation est modifiée (rotation autour de la liaison C $\alpha$ -C $\beta$ ). Le noyau imidazole se trouve alors dans l'environnement du site actif et sa distance avec la cystéine catalytique passe de 11,9 à 6,1 Å (distance N $\epsilon_2$ -S $\gamma$ ) (Figure IV-17).

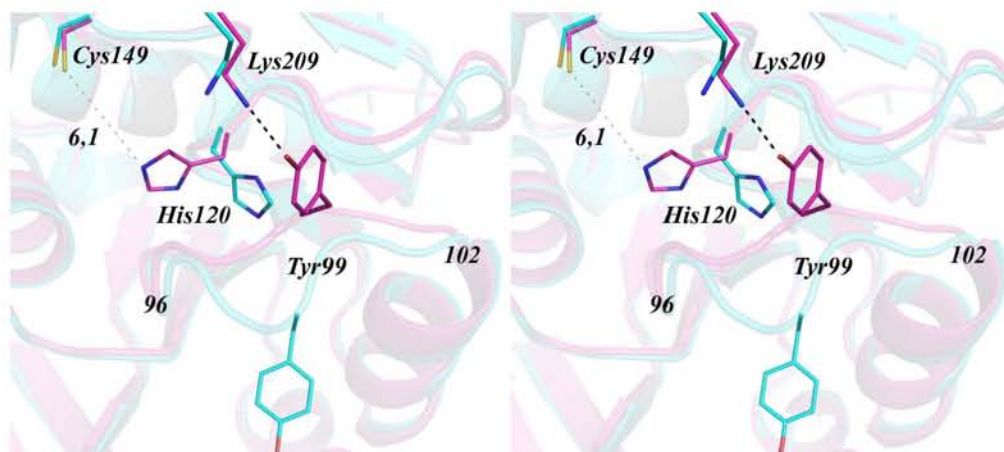
### ***A.5.3. La boucle 48-52***

La boucle 48-52 est l'une des boucles dont la conformation apparaît être variable en fonction des structures, mais aussi en fonction des sous-unités au sein d'une même structure. Ainsi, on peut distinguer deux conformations types (Figure IV-18) :

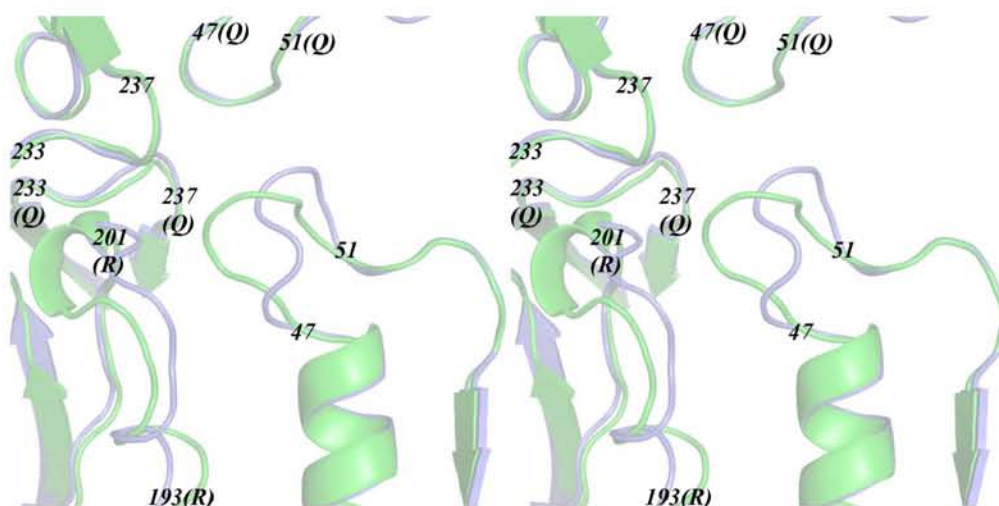
- La première, observée dans la majorité des monomères quelle que soit la structure, interagit avec la partie C-terminale de la boucle S du monomère relié par l'axe *R*. Les conformations de cette boucle 48-52 et de la boucle S sont équivalentes à celles observées dans les structures de GAPDH.
- La seconde conformation, très minoritaire, n'est rencontrée que dans les sous-unités D des modèles APO1 et APO2. Dans ces deux monomères, la conformation de la partie C-terminale de la boucle S est, elle aussi, différente de celle observée dans les GAPDH. Le mouvement de ces deux boucles semble concerté, puisque qu'elles interagissent ensemble quelles que soient leurs conformations.

## ***A.6. Construction d'un modèle de l'intermédiaire hémithioacétal entre l'E4PDH et son substrat (E4P).***

Un modèle de l'intermédiaire hémithioacétal de l'E4PDH a été construit à partir des informations recueillies dans les 3 structures résolues au cours de ces travaux. En utilisant la structure du modèle ECAP, une molécule de cofacteur NAD a tout d'abord été positionnée dans le site de fixation du cofacteur par superposition à l'analogue 3-CAPAD présent dans le modèle. Ensuite, une molécule de substrat (E4P) a été construite dans le site actif de l'enzyme en liaison covalente avec la cystéine catalytique (Cys149), pour mimer un intermédiaire hémithioacétal. Deux hypothèses ont été retenues pour cette construction : (i) le groupement phosphate du substrat occupe



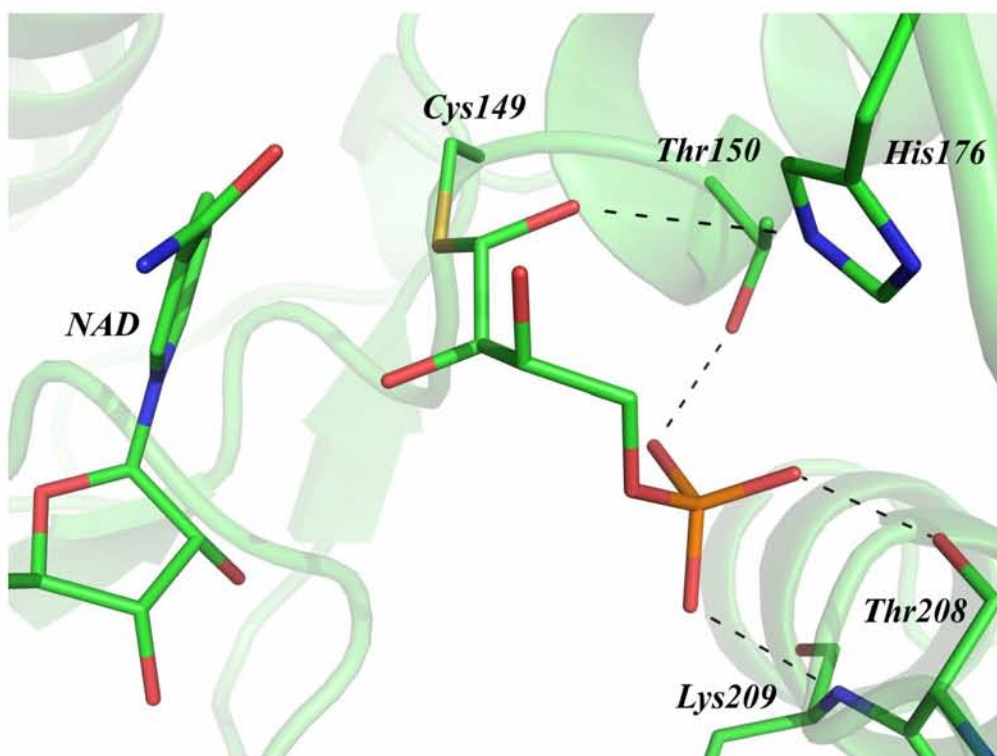
**Figure IV-17 :** Comparaison des conformations de la boucle 96-102 dans les monomères B (bleu) et C (rose). Les sous-unités ont été superposées en utilisant les atomes de Ca. Les extrémités de la boucle 96-102 sont symbolisées par les chiffres correspondants. Les chaînes latérales des résidus Cys149, His176, His120 et Tyr99 sont représentées sous forme de bâtons et annotées. La liaison hydrogène existant entre les chaînes latérales de la Tyr99 et la Lys209 du monomère C et la distance séparant les atomes  $S_{\gamma}$  de la Cys149 et  $N_{\epsilon_2}$  de l'His120 sont matérialisées respectivement par des pointillés noirs et gris.



**Figure IV-18 :** Vue stéréoscopique du coeur du tétramère d'E4PDH. La structure tétramérique APO1 (vert) a été superposée à elle-même, en faisant correspondre aux sous-unités A, B, C et D les sous-unités D, C, B et A respectivement (bleu). Les bornes des boucles mobiles du coeur du tétramère sont indiquées par des chiffres avec, entre parenthèse, une lettre correspondant à l'axe binaire au travers duquel les sous-unités sont reliées. Ainsi, la boucle 47-51 verte, située en bas à droite, appartient à la sous-unité A et interagit avec la boucle 193-201 de la sous-unité reliée par l'axe binaire R, c'est à dire, appartenant à la sous-unité D. De la même manière, la boucle 47-51 bleue, en bas à droite qui appartient à la sous-unité D interagit avec la boucle 193-201 du monomère A.

le site Pi identifié dans les modèles APO1 et ECAP et (ii) l'histidine catalytique (His176) joue, comme dans le cas de la GAPDH, le rôle d'un catalyseur acide/base et forme une liaison hydrogène avec l'atome O1 du substrat lors de la formation de l'intermédiaire tétraédrique. Pour ce faire une rotation de 180° du noyau imidazole de l'His176 est nécessaire pour faire pointer l'atome NE2 vers la cystéine catalytique. Ce modèle a enfin été soumis à une procédure d'idéalisation géométrique par le logiciel Refmac5 jusqu'à obtenir une conformation stable.

Le groupement phosphate du substrat, localisé dans le site de reconnaissance anionique Pi, forme des interactions avec les chaînes latérales des résidus Thr150 et Thr208 et avec la chaîne principale de la Lys209 (Figure IV-19). L'atome O1 du substrat est situé à distance de l'His176 pour former une liaison hydrogène. Ainsi, selon ce modèle, la géométrie originale du site actif de l'E4PDH par rapport à celui de la GAPDH, est compatible avec l'intervention de l'His176 en tant que catalyseur acide/base.



*Figure IV-19 : Modèle de l'intermédiaire hémithioacétal de l'E4PDH.*

## ***PARTIE B : DISCUSSION***

L'E4PDH catalyse l'oxydation irréversible du *D*-érythrose-4-phosphate (*D*-E4P) en 4-phosphoérythronate (4PE) en présence de NAD et d'une molécule d'eau. Cette enzyme présentant de nombreuses similitudes avec la GAPDH (repliement, nature des résidus catalytiques, du cofacteur, substrats structurellement très proches...), son étude structurale constitue une opportunité d'établir un parallèle intéressant entre ces deux enzymes, dans le but de comprendre comment de subtiles différences structurales pouvaient être à l'origine des propriétés spécifiques de chacune de ces deux enzymes. Les études biochimiques réalisées sur l'E4PDH d'*E.coli* ont notamment révélé que, comparée aux GAPDH, cette enzyme présentait une affinité relativement faible vis-à-vis de son cofacteur, que ni les expériences de mutagenèse dirigée, ni l'étude cristallographique de la structure apoenzyme n'avaient permis d'expliquer (Boschi-Muller *et al.*, 1997 ; Kauffmann, 2000). L'unique structure connue au début de cette étude, qui présentait des zones non définies et, ne renfermait aucun ligand, n'avait pas non plus permis d'analyser le mode de fixation du cofacteur ou la présence éventuelle de sites de reconnaissance anionique. Pour tenter de répondre à ces interrogations nous avons résolu des structures de l'E4PDH sous forme de différents complexes.

Trois structures de l'E4PDH d'*Escherichia coli* ont été résolues au cours de cette étude. Deux formes apoenzymes, APO1 et APO2, dont les cristaux ont été obtenus dans des conditions originales, ont permis d'apporter des informations sur l'existence d'un site anionique dans le site actif de l'E4PDH (APO1) et sur la structuration de la boucle S (APO2). La résolution de la structure cristallographique d'un complexe covalent (ECAP) entre l'E4PDH et un analogue du cofacteur, nous

a permis d'étudier le mode de fixation du cofacteur dans l'E4PDH et, par comparaison aux données disponibles pour les GAPDH, de proposer des hypothèses pour expliquer la faible affinité de cette enzyme pour le NAD. L'analyse de ces données structurales nous a enfin permis d'identifier le résidu qui intervient vraisemblablement dans l'activation de la molécule d'eau lors de l'étape de désacylation, et de proposer un mécanisme réactionnel.

### ***Fixation de la molécule de cofacteur : affinité, transition apo-holo et coopérativité***

L'E4PDH se caractérise par une affinité très faible pour son cofacteur : en comparaison à la GAPDH, la valeur de la constante de dissociation ( $K_D$ ) déterminée pour le  $\text{NAD}^+$  est en effet 800 fois plus élevée. Cette faible affinité explique probablement pourquoi l'enzyme n'a jamais pu être isolée sous forme holo en solution, et également pourquoi, malgré les nombreux essais de cocristallisation effectués (concentrations en  $\text{NAD}^+$  variant de 1,5 à 10 mM) nous n'avons jamais pu obtenir une structure du complexe binaire. Face à cet échec, nous avons recherché une solution alternative, qui consistait à utiliser un marqueur d'affinité, analogue de  $\text{NAD}^+$ , le 3-(chloroacétyl) pyridine adénine dinucléotide (3-CAPAD). Cette stratégie offrait l'avantage de permettre de contrôler aisément la fixation de l'analogue à l'enzyme avant de réaliser les essais de cristallisation, puisque le 3-CAPAD réagit avec la cystéine catalytique et conduit à l'inactivation irréversible de l'enzyme.

#### ***- Transition apo holo***

Les études structurales des GAPDH de *B. stearothermophilus* et d'*E. coli* (Skarzynski *et al.*, 1987 ; Skarzynski and Wonacott, 1988 ; Duée *et al.*, 1996 ; Yun *et al.*, 2000) ont permis de mettre en évidence une série de changements conformationnels associée à la fixation du cofacteur appelée transition apo-holo. Cette transition peut se résumer par une rotation du domaine de fixation du cofacteur par rapport au domaine catalytique qui conduit à la fermeture du site de fixation du coenzyme et à l'optimisation des interactions entre l'enzyme et son ligand. Il était naturel de penser que le même type de transition pouvait avoir lieu dans l'E4PDH, compte tenu de ses similarités structurales avec les GAPDH. Malgré la mise en évidence d'une certaine liberté conformationnelle de la structure de l'E4PDH qui rappelle celle de la structure de la GAPDH lors de la transition apo - holo, les structures APO1, APO2 et celle du complexe (ECAP) avec l'analogue de cofacteur sont toutes plus proches de la forme apo que de la forme holo de la GAPDH. La première raison qui pourrait expliquer l'absence de transition apo-holo dans le complexe ECAP pourrait être liée au fait que le complexe a été obtenu avec un analogue et non le cofacteur naturel. A cet égard, l'équipe du

Professeur Wonacott (Biesecker and Wonacott, 1977 ; Skarzynski and Wonacott, 1988) a décrit la fixation du cofacteur comme étant séquentielle, de la partie adénosine jusqu'au pyridinium et que la fixation de ce dernier était essentielle à la réalisation complète de la transition. Dans notre cas, la liaison covalente formée entre la cystéine catalytique et le pyridinium modifié nuit à son positionnement idéal et pourrait ainsi compromettre la réalisation de la transition apo - holo. Néanmoins, l'ensemble de nos résultats (voir ci-dessous) ne semble pas indiquer qu'une transition équivalente à celle décrite pour la GAPDH soit nécessaire au bon fonctionnement de l'E4PDH. On ne peut toutefois pas exclure que la fixation du cofacteur naturel induise des réarrangements, même mineurs, de la structure qui n'ont pas pu être mis en évidence dans cette étude.

### ***- Affinité du cofacteur***

L'analyse, dans la structure du complexe ECAP, des interactions entre l'enzyme et la molécule de 3-CAPAD révèle que la plupart des interactions observées dans le cas des GAPDH sont retrouvées dans le cas de cette enzyme. En effet, sur les 8 liaisons hydrogène directes formées entre la GAPDH et le NAD (Figure I-5), 5 sont équivalentes dans le complexe E4PDH/3-CAPAD, les trois interactions non retrouvées impliquant les résidus en position 77, 180 et 313.

L'interaction provenant de la chaîne principale en position 77 qui stabilise le cycle de l'adénine (NH<sub>2</sub>) dans la GAPDH à la suite de la transition apo – holo n'est pas retrouvée dans le complexe E4PDH/3-CAPAD. Les deux partenaires sont situés à une distance trop importante pour interagir car une transition semblable à celle des GAPDH n'est pas observée dans l'E4PDH lors de la fixation du 3-CAPAD. Dans la GAPDH, la fixation du cofacteur implique également le rapprochement de la chaîne latérale de l'Asp32 qui vient stabiliser les deux fonctions hydroxyle du ribose de l'adénine. Dans l'E4PDH, la position 32 est occupée par un glutamate invariant dont la chaîne latérale, plus longue, permet à ce dernier de jouer un rôle équivalent à l'Asp32, malgré l'absence de transition apo – holo. De plus, bien que l'Arg77 ne puisse contribuer à la stabilisation de l'adénine, une interaction originale par rapport à la GAPDH existe dans l'E4PDH. Elle est formée grâce à la présence d'une tyrosine invariante en position 99 dont la fonction hydroxyle interagit avec le groupement NH<sub>2</sub> de l'adénine. Cette tyrosine se substitue au résidu phénylalanine généralement rencontré dans les séquences de GAPDH. Ainsi, l'E4PDH semble avoir développé un réseau d'interactions original pour stabiliser l'adénosine sans qu'une transition apo-holo du même type que celle de la GAPDH soit nécessaire. De plus, le fait que le site de fixation du cofacteur de l'E4PDH soit plus exposé au solvant que dans le cas de la GAPDH peut en partie expliquer la faible affinité de l'enzyme pour son cofacteur.

La seconde interaction manquante implique le résidu en position 180. Dans la *Bst*GAPDH, ce résidu est une asparagine qui stabilise le groupement pyrophosphate ( $O_{2A}$ ) *via* sa chaîne latérale ( $N\delta_2$ ). Dans l'E4PDH d'*E. coli*, ce résidu est une histidine qui appartient à la partie non structurée de la boucle S. Il convient cependant de préciser que ce résidu n'est pas strictement conservé dans les séquences d'E4PDH puisque la position 180 est occupée soit par une histidine soit par une asparagine. La position 180 n'est d'ailleurs pas conservée non plus dans les séquences de GAPDH. Une alanine ou une glycine, deux résidus ne pouvant jouer un rôle équivalent à celui de l'asparagine de l'enzyme de *B. stearothermophilus*, sont préférentiellement rencontrés sans qu'aucun effet notable sur l'affinité de l'enzyme n'ait été mentionné (Corbier *et al.*, 1990b). De fait, le résidu 180 ne semble pas déterminant pour la fixation du cofacteur dans les GAPDH et ne devrait pas l'être non plus dans les E4PDH.

Enfin, la position 313 est occupée par une asparagine invariante dans les GAPDH comme dans les E4PDH. Dans les GAPDH, l'Asn313 ( $N\delta_2$ ) forme une liaison hydrogène avec le carboxamide du cofacteur ( $O_{7N}$ ). L'absence d'interaction entre le 3-CAPAD en l'Asn313 est probablement une conséquence directe du fait que, dans l'analogue, la fonction carboxamide est remplacée par une fonction chlorométhylcétone. Néanmoins, bien que la position du résidu Asn313 soit équivalente dans les structures de l'E4PDH et de la GAPDH, la plupart des structures résolues au cours de cette étude montrent que la densité électronique de la chaîne latérale de l'asparagine est diffuse, car non stabilisée par la chaîne principale du résidu Val237. Ce dernier adopte une conformation originale par rapport aux GAPDH et appartient à la boucle 233-237 qui se trouve à proximité des boucles 47-51 et S, peu structurées dans l'E4PDH. La mobilité de ces régions a pour conséquence un manque de stabilisation de la chaîne latérale de l'Asn313 qui ne peut ainsi jouer pleinement son rôle dans la stabilisation du cofacteur. Dans les GAPDH, dont les structures sont connues sous forme apo ou holo, la Val237 est systématiquement située à l'extérieur des zones autorisées du diagramme de Ramachandran (Duée *et al.*, 1996). La chaîne principale de la Val237 dans sa conformation particulière ( $\phi, \psi$  ;  $\sim 80^\circ, \sim 130^\circ$ ) participe donc indirectement à la fixation du cofacteur en positionnant la chaîne latérale de l'Asn313 pour stabiliser le nicotinamide. Dans les structures de l'E4PDH, la chaîne principale Val237 ( $\sim -100^\circ, \sim 130^\circ$ ) se trouve dans une zone autorisée du diagramme de Ramachandran et ne peut stabiliser l'Asn313 dans cette conformation.

En résumé, deux arguments peuvent être avancés pour expliquer la faible affinité de l'E4PDH pour son cofacteur. D'une part l'absence de transition apo-holo similaire à celle de la GAPDH dont l'une des conséquences est de fermer le site de fixation autour de la molécule de cofacteur pour former un complexe stable. D'autre part, la mobilité de boucles situées au coeur du



tétramère (boucles S, 47-41, 233-237) perturbe le positionnement exact de l'Asn313 déterminant dans la fixation du nicotinamide qui est le groupement accepteur d'électrons lors de l'acte catalytique.

### **- Coopérativité**

De façon générale, les GAPDH se caractérisent par des propriétés de coopérativité vis-à-vis de la fixation du cofacteur, coopérativité qui se révèle être soit positive soit négative suivant la source de l'enzyme ou les conditions physico-chimiques (Conway and Koshland, 1968 ; Trentham, 1968 ; De Vijlder *et al.*, 1969 ; Seydoux *et al.*, 1973 ; Allen and Harris, 1975 ; Berni *et al.*, 1979). Dans le cas de la GAPDH de *B. stearothermophilus*, une coopérativité négative a été observée (Allen and Harris, 1975). Cette enzyme tétramérique peut être décrite comme un dimère de dimères (types O-P et Q-R) avec dans chacun, deux sites du NAD reliés par une coopérativité négative. Ainsi, la fixation de la première molécule de NAD sur l'un des monomères (par exemple O) du dimère diminue l'affinité de l'autre sous-unité (P). Jusqu'à ce jour, aucune information n'était disponible concernant l'E4PDH. L'analyse de la structure renfermant le 3-CAPAD montre que, à l'intérieur de chacun des quatre tétramères, la molécule de 3-CAPAD est bien définie dans deux monomères sur quatre. Les deux monomères en question sont dans tous les cas des sous-unités du type O et Q (ou du type P et R), c'est-à-dire reliés par l'axe binaire *Q*. Ce résultat semble indiquer que l'E4PDH d'*E.coli* présente également des propriétés de coopérativité négative pour le cofacteur. Il est intéressant de noter que des résultats semblables ont été obtenus par Cowan-Jacob *et al.* (2003) dans l'étude cristallographique de la GAPDH du muscle de lapin. La structure révèle 2 molécules de NAD fixées par tétramère dans deux sous-unités reliées par l'axe binaire *Q*. Les études réalisées par l'équipe du Professeur G. Branlant ont permis de montrer que la coopérativité s'exprimait au travers de l'interface O-P plutôt que l'interface O-R constituée majoritairement par la boucle S (Clermont *et al.*, 1993 ; Roitel *et al.*, 1999 et 2003). Nos résultats cristallographiques obtenus sur le complexe ECAP et les études physico-chimiques réalisées sur la GAPDH sembleraient indiquer que l'E4PDH peut également être décrite sous la forme d'un dimère de dimère et que chaque dimère (O-P ou Q-R) présente une coopérativité négative vis-à-vis de la fixation du cofacteur. La détermination des constantes d'affinités intrinsèques des sous-unités pour le cofacteur permettrait de confirmer cette hypothèse.

## ***Analyse fonctionnelle du site actif de l'E4PDH***

### ***- Conformation du site actif***

Comme dans la première structure d'E4PDH résolue (Kauffmann, 2000), les résidus catalytiques (Cys149 et His176) adoptent (dans le modèle APO1) des conformations différentes de celles observées dans la GAPDH. La différence la plus notable concerne l'histidine catalytique dont la chaîne latérale, au lieu de pointer en direction de la cystéine comme dans la GAPDH, est dirigée vers le site de reconnaissance anionique dans l'ensemble des structures résolues. La conformation des résidus catalytiques exclut toute possibilité de formation d'une paire d'ions dans ces structures de type apo. Il semble donc, que l'His176 ne contribue pas à augmenter la nucléophilie de la Cys149, par abaissement de son pKa dans l'E4PDH, comme elle le fait dans la GAPDH (Soukri *et al.*, 1989 ; Talfournier *et al.*, 1998). Il n'est cependant pas exclu que la fixation du substrat et/ou du cofacteur puisse conduire à une réorganisation du site actif. L'analyse des résultats de modélisation du substrat dans le site actif de l'E4PDH, révèle que, malgré sa conformation, l'histidine catalytique est capable de créer une liaison hydrogène avec l'atome O1 de l'E4P, ce qui lui permettrait de jouer un rôle de catalyseur acide/base équivalent à l'His176 des GAPDH en accord avec les résultats de mutagenèse dirigée (Boschi-Muller *et al.*, 1997). L'ensemble de ces observations tend à suggérer que, si l'His176 doit activer la cystéine catalytique, une réorganisation du site actif est nécessaire, mais que l'His176 dans la conformation observée est compétente pour jouer son rôle acide/base dans l'acte catalytique.

### ***- Site de reconnaissance anionique***

Deux des structures de l'E4PDH résolues au cours de cette étude, APO1 et ECAP, sont issues de cristaux obtenus en présence respectivement de phosphate et de sulfate. Dans ces deux modèles, le site actif de l'enzyme s'est révélé contenir l'un de ces anions dans un site, structuralement analogue au site de reconnaissance anionique Pi des GAPDH. Dans l'E4PDH, ce site appelé Pi par analogie aux GAPDH occupe une position distante en moyenne de 7,1 Å de l'atome Sy de la cystéine catalytique, lorsque celle-ci est correctement orientée pour la catalyse. Seul site de reconnaissance anionique fonctionnel identifié du site actif de l'E4PDH, ce site Pi constitue donc vraisemblablement le site de fixation du groupement phosphate (C4P) du substrat. D'autre part, l'étape de désacylation correspondant à une hydrolyse et non pas à une phosphorolyse comme dans le cas de la GAPDH, il est très probable que le substrat demeure fixé dans ce site tout au long de la réaction.

L'absence de tout autre site de reconnaissance anionique fonctionnel dans le site actif de l'E4PDH devient particulièrement intéressante au regard de l'étude réalisée par Boschi-Muller *et al.* (1997). En effet, les auteurs ont montré que l'E4PDH était capable d'utiliser le G3P en tant que substrat, de façon peu efficace, dans une réaction de phosphorylation similaire à celle catalysée par la GAPDH. Dans ce cas, deux groupements phosphates sont présents dans le site actif de l'enzyme, le phosphate du substrat et la molécule de phosphate inorganique. Au vu des résultats obtenus sur la GAPDH (*cf.* Chapitre III Partie B), il est désormais clair que le site de phosphorylation, dans ces enzymes, est constitué par le site de reconnaissance anionique appelé nouveau site Pi. Un site compétent équivalent à ce site Pi existant dans le site actif de l'E4PDH, il est vraisemblable, comme l'avaient suggéré Boschi-Muller *et al.* (1997), que la faible activité phosphorylante de l'E4PDH soit due à l'existence de ce site. Il devient alors nécessaire d'imaginer que le phosphate du substrat doit céder sa place à une molécule de phosphate inorganique pour que la réaction de phosphorylation ait lieu. Dans cette hypothèse, les données actuelles ne nous permettent pas de préciser si ce groupement phosphate a ou non la possibilité d'être quelque peu stabilisé par la structure protéique.

#### **- Efficacité de l'étape de désacylation – Résidu susceptible d'activer la molécule d'eau.**

Dans la réaction catalysée par l'E4PDH avec son substrat naturel, la seconde étape de la réaction est une hydrolyse efficace de l'intermédiaire thioacylenzyme. En effet, l'absence d'effet isotopique de solvant suggère que l'étape limitante, associée à la seconde étape de la réaction, concerne la libération des produits plutôt que l'hydrolyse elle-même (Boschi-Muller *et al.*, 1997). L'efficacité de cette étape d'hydrolyse implique la présence, à proximité du site actif de l'E4PDH, d'un résidu capable de jouer le rôle de base et d'activer la molécule d'eau hydrolytique. Un candidat probable identifié au cours de cette étude pourrait être le résidu His120, situé dans l'environnement direct du site actif. La chaîne latérale de ce résidu a été observée dans deux conformations. La première conformation est rencontrée dans la grande majorité des structures et est incompatible avec un rôle d'activateur de la molécule d'eau. En revanche, une seconde conformation, liée au changement conformationnel de la boucle 96-102 est rencontrée dans le monomère C de la structure APO1. Cette conformation fait pointer la chaîne latérale de l'His120 en direction du site actif, à une distance d'environ 6 Å de l'atome S $\gamma$  de la cystéine catalytique et environ 7 Å de l'atome C1 du modèle de l'intermédiaire hémithioacétal. Le rôle potentiel du résidu His120 devra être confirmé par des études enzymatiques d'un mutant H120N. S'il s'avère que le résidu His120 joue bien le rôle de base pour activer une molécule d'eau, et donc assister l'étape de désacylation, il apparaît alors que cette action nécessite en premier lieu la libération du cofacteur réduit. En effet, lorsqu'elle est dans une conformation compatible avec le rôle d'activateur de la molécule d'eau, l'extrémité de la chaîne

latérale de l'His120 occupe une position trop proche de celle du ribose adjacent au nicotinamide du cofacteur pour coexister.

En conclusion de cette étude comparative, nous proposons le mécanisme suivant :

En premier lieu, l'E4PDH, sous forme apoenzyme, fixe les molécules de cofacteur et de substrat. L'ordre exact de ces événements reste encore à déterminer, néanmoins, il semble probable, compte tenu de l'impossibilité à isoler la forme holoenzyme de l'E4PDH, que les trois partenaires soient nécessaires pour conduire à la formation d'un complexe ternaire compétent. L'E4P se fixe dans le site actif de l'enzyme avec son groupement phosphate (C4P) stabilisé par le site de reconnaissance anionique Pi. L'attaque de la cystéine catalytique sur le carbone aldéhydique du substrat conduit à la formation transitoire d'un intermédiaire hémithioacétal, stabilisé par l'histidine essentielle (His176). Le transfert d'hydrure, favorisé grâce à l'assistance basique de l'His176, conduit alors à la formation de l'intermédiaire thioacylenzyme et à la réduction du NAD<sup>+</sup> en NADH. Le cofacteur réduit est libéré, autorisant le changement conformationnel de la Tyr99 et le repositionnement de la chaîne latérale de l'His120 en direction du site actif. Cette dernière, jouant le rôle de base, active une molécule d'eau qui peut ainsi réaliser l'attaque nucléophile sur le thioacylenzyme et conduire à la libération du produit.

# CONCLUSION ET PERSPECTIVES

L'objectif de ces travaux de thèse visait à étudier deux enzymes de la famille des Aldéhyde déshydrogénases : la GAPDH de *Bacillus stearothermophilus* et l'E4PDH d'*Escherichia coli*. Structuralement apparentées et utilisant des substrats aldéhydiques très proches (le G3P et l'E4P, respectivement), ces deux enzymes possèdent des spécificités qui leurs sont propres et procèdent par des mécanismes catalytiques différents. En effet, bien que la première étape de la réaction (acylation) soit identique, les deux enzymes se différencient au travers de la seconde étape (désacylation) qui correspond à une phosphorylation ou à une hydrolyse respectivement pour la GAPDH et l'E4PDH. De plus, la GAPDH a été très étudiée, alors que l'E4PDH n'a fait l'objet que d'un nombre d'études très limité. Par exemple, plus de 50 structures cristallographiques de la GAPDH provenant de plus de 20 organismes différents sont présentes dans la PDB alors que seule une structure de l'E4PDH d'*E. coli* a été résolue mais non publiée (Kauffmann, 2000). De même, le mécanisme catalytique de la GAPDH est très bien caractérisé sur le plan moléculaire alors que de nombreuses incertitudes demeurent sur celui de l'E4PDH. Aussi les objectifs de ces travaux de thèse étaient-ils très différents pour chacune de ces deux enzymes.

L'objectif principal de l'étude de la GAPDH visait à attribuer leur rôle respectif aux deux sites de reconnaissance anionique présents dans le site actif de l'enzyme. En effet, malgré un grand nombre de travaux, la contribution exacte des sites Pi et Ps au mécanisme réactionnel de la GAPDH restait indéterminée. Pour répondre à cette question, nous avons résolu la structure de deux intermédiaires covalents (thioacylenzyme-NADH et acylenzyme-NAD<sup>+</sup>), et celles de complexes enzyme-produit de la réaction (1,3-DPG). Ces structures, avec la structure du complexe michaelien enzyme-G3P (Didierjean *et al.*, 2003), nous ont permis de localiser la position du groupement phosphate (C3P) du substrat durant les différentes étapes de la réaction. Ainsi, cette étude met en évidence pour la première fois le mouvement de balancier (ou "flip-flop") qu'opère le substrat entre les deux sites de reconnaissance anionique au cours de la catalyse et montre que ce mouvement du groupement phosphate du substrat est nécessaire à l'étape d'échange du cofacteur qui intervient entre les étapes d'acylation et de désacylation. Cependant, la façon dont la présence de NAD<sup>+</sup> favorise l'étape de phosphorylation reste encore incomprise. Un autre point important qui émane de cette étude est la variabilité conformationnelle de l'atome O1 qui peut occuper deux positions distinctes. Si l'une de ces positions correspond à une conformation indispensable à l'acte catalytique (puisque'elle permet à l'His176 d'assurer son rôle d'acide-base), il nous est impossible à ce stade de savoir si la seconde est également représentative ou non d'un état physiologique. En

l'absence d'une preuve expérimentale de l'existence de cette conformation, des études théoriques pourraient peut être permettre d'apporter des éléments de réponse.

L'existence de conformations multiples au sein d'une même structure est inhérente à la méthode utilisée. L'idéal aurait été de pouvoir déclencher une réaction simultanée dans l'ensemble des molécules du cristal, en disposant, par exemple, d'un substrat photoactivable.

Dans l'immédiat, il semble nécessaire de valider les modèles obtenus et notamment l'estimation des taux d'occupation des différentes conformations présentes au sein de chaque modèle en poursuivant les affinements avec Buster-TNT. Une collaboration est en cours à ce sujet avec Clemens Vornrhein (Global Phasing Ltd). De même, il conviendrait de confirmer les résultats obtenus sur l'acylenzyme-NAD<sup>+</sup> par l'obtention du même type d'intermédiaire avec l'enzyme de type sauvage. L'acylation de la cystéine essentielle par le 1,3-DPG pourrait être directement suivie par microspectrophotométrie *via* la disparition de la bande de Racker (Collaboration avec Stefano Bruno et Andrea Mozzarelli, Université de Parme).

Concernant l'E4PDH, l'objectif initial de ces travaux de thèse était d'obtenir la structure de l'enzyme en présence de différents ligands : ion phosphate ou sulfate ; cofacteur et substrat (complexes binaires de type holoenzyme ou enzyme-substrat et complexe ternaire de type enzyme/substrat/cofacteur). Même si les difficultés rencontrées au cours de cette étude ne nous ont pas permis de réaliser pleinement ces objectifs, 3 structures APO1, APO2, et ECAP, (les toutes premières de cette enzyme) ont néanmoins été obtenues. Leur analyse a permis d'avancer les premiers arguments structuraux à l'origine des propriétés particulières de l'E4PDH par rapport à la GAPDH. Tout d'abord, nous avons mis en évidence l'existence d'un seul site de reconnaissance anionique. Ce site possède des caractéristiques voisines du site Pi de la GAPDH, de par sa localisation et la nature des résidus qui le composent. Ce site est un peu plus éloigné des résidus catalytiques que celui retrouvé dans la GAPDH, ce qui est pleinement compatible avec la fixation du groupement phosphate d'un substrat qui possède un carbone de plus que le G3P et est donc un peu plus long. D'autre part, cette étude a également permis de proposer des hypothèses concernant la faible affinité de l'enzyme pour son cofacteur et de proposer que l'E4PDH présente probablement des propriétés de coopérativité négative vis-à-vis du NAD. Enfin, nous avons identifié un résidu (His120) susceptible de jouer un rôle de base et de conduire ainsi à l'activation de la molécule d'eau

intervenant dans l'étape d'hydrolyse du thioacylenzyme.

Les hypothèses formulées sur la base de l'analyse structurale devront maintenant être validées par des études biochimiques couplées à la mutagenèse dirigée. Ainsi, il paraît indispensable de confirmer le rôle de l'His120 en remplaçant ce résidu par un autre n'ayant pas la capacité à jouer un rôle de base. Si notre hypothèse est vérifiée, le mutant devrait présenter une étape de désacylation fortement ralentie et, en particulier, révéler la présence d'un effet isotopique de solvant prouvant que l'étape limitante, dans ce mutant, est associée à l'hydrolyse elle-même, alors qu'elle concerne la libération des produits dans l'enzyme de type sauvage. Toujours en relation avec l'efficacité de l'étape de désacylation, il pourrait également être intéressant d'examiner les conséquences de la substitution du résidu Tyr99, dans la mesure où nous suggérons que la conformation de l'His120 adaptée à la désacylation soit liée au mouvement de la boucle 96-102. En effet, au-delà de sa participation à la fixation du cofacteur, le résidu Tyr99 développe une interaction avec la Lys209 qui stabilise la boucle 96-102, et pourrait de ce fait participer à la réorientation de l'His120 en direction du site du site actif.

Concernant la faible affinité de l'enzyme pour le NAD, il serait intéressant de tenter de "rigidifier" la conformation de la boucle 233-237 en y introduisant la proline majoritairement présente dans les GAPDH (Pro235). En effet, nous proposons que la conformation particulière de cette boucle nuise au bon positionnement de l'Asn313 et influe ainsi sur l'affinité de l'enzyme sur le cofacteur. Alternativement, la démarche inverse pourrait être tentée sur la GAPDH, en remplaçant la Pro235 par une isoleucine, résidu présent dans les E4PDH (ou remplacé par une valine).

Par ailleurs, l'analyse de la structure renfermant le 3-CAPAD montre que la molécule de 3-CAPAD n'est bien définie que dans les deux monomères reliés par l'axe binaire  $Q$ , suggérant que l'E4PDH d'*E. coli* présente également des propriétés de coopérativité négative pour le cofacteur. Cette hypothèse mériterait d'être validée par des études biochimiques. Pour la GAPDH, la détermination des constantes de dissociation intrinsèque de chaque sous-unité pour le cofacteur a été réalisée par des mesures de quenching de fluorescence du Trp310, lors de l'ajout de quantités croissantes de cofacteur (Corbier *et al.*, 1990b ; Clermont *et al.*, 1993). Ce phénomène n'existant pas avec l'E4PDH (Boschi-Muller, communication personnelle), l'affinité du NAD pour cette enzyme a été déterminée par des expériences de protection du NAD vis-à-vis de l'inactivation par l'iodoacétamide et n'a permis d'accéder qu'à une valeur de  $K_D$  moyen. Une façon d'accéder aux constantes individuelles serait peut-être de suivre la fixation du NADH par fluorescence aux



alentours de 450 nm. L'intensité de fluorescence du NADH étant dépendante de son état (libre ou lié), cette approche pourrait permettre d'accéder à la fraction de NADH lié et de procéder à une analyse mathématique de la coopérativité.

Concernant maintenant les interactions enzyme-substrat, une molécule de type 1,2-époxybutane-4-phosphate, pourrait permettre l'obtention d'un mime d'intermédiaire covalent tétraédrique, à défaut d'un complexe enzyme-substrat naturel.

Enfin, il pourrait également être intéressant de travailler sur l'E4PDH d'une autre source, ce qui présenterait l'avantage de pouvoir discerner les spécificités propres à l'E4PDH d'*E. coli* des propriétés communes aux E4PDH.



# RÉFÉRENCES BIBLIOGRAPHIQUES

- Alefounder PR and Perham RN. (1989) *Identification, molecular cloning and sequence analysis of a gene cluster encoding the class II fructose 1,6-bisphosphate aldolase, 3-phosphoglycerate kinase and a putative second glyceraldehyde 3-phosphate dehydrogenase of Escherichia coli.* Mol Microbiol. **3**(6):723-32.
- Allen G and Harris JI. (1975) *The binding of nicotinamide-adenine dinucleotide to glyceraldehyde 3-phosphate dehydrogenase from Bacillus stearothermophilus.* Biochem J. **151**(3):747-9.
- Armstrong JM and Trentham DR. (1976) *The reactions of D-glyceraldehyde 3-phosphate with thiols and the holoenzyme of D-glyceraldehyde 3-phosphate dehydrogenase and of inorganic phosphate with the acyl-holoenzyme.* Biochem J. **159**(3):513-27.
- Berni R, Mozzarelli A, Rossi GL. (1979) *Crystallographic symmetry and coenzyme binding properties of D-glyceraldehyde-3-phosphate dehydrogenase from the tail muscle of Palinurus vulgaris.* J Biol Chem. **254**(16):8004-6.
- 
- Biellmann JF, Branlant G, Foucaud BY and Jung MJ. (1974) *Preparation of 3-chloroacetylpyridine adenine dinucleotide: an alkylating analogue of NAD<sup>+</sup>.* FEBS Lett. **40**: 29-32.
- Biesecker G and Wonacott AJ. (1977) *Coenzyme binding and co-operativity in D-glyceraldehyde 3-phosphate dehydrogenase.* Biochem Soc Trans. **5**(3):647-52.
- Biesecker G, Harris JI, Thierry JC, Walker JE, Wonacott AJ. (1977) *Sequence and structure of D-glyceraldehyde 3-phosphate dehydrogenase from Bacillus stearothermophilus.* Nature. **266**:328-33.
- Boschi-Muller S, Azza S, Pollastro D, Corbier C, Branlant G. (1997) *Comparative enzymatic properties of GapB-encoded erythrose-4-phosphate dehydrogenase of Escherichia coli and phosphorylating glyceraldehyde-3-phosphate dehydrogenase.* J Biol Chem. **272**(24):15106-12.
- Boschi-Muller S and Branlant G. (1999) *The active site of phosphorylating glyceraldehyde-3-phosphate dehydrogenase is not designed to increase the nucleophilicity of a serine residue.* Arch. Biochem. Biophys. **363**(2): 259-66.
- Branlant C, Oster T and Branlant G. (1989) *Nucleotide sequence determination of the DNA region coding for Bacillus stearothermophilus glyceraldehyde-3-phosphate dehydrogenase and of the flanking DNA regions required for its expression in Escherichia coli.* Gene **75**:145-155.
- Bricogne G and Irwin J. (1996) *Macromolecular Refinement: Proceedings of the CCP4 Study Weekend, édité par E. Dodson, M. Moore, A. Ralph and S. Bailey, pp. 85-92.* Warrington: Daresbury Laboratory.

- Brinkmann H, Cerff R, Salomon M and Soll J. (1989) *Cloning and sequence analysis of cDNAs encoding the cytosolic precursors of subunits GapA and GapB of chloroplast glyceraldehyde-3-phosphate dehydrogenase from pea and spinach*. *Plant Mol Biol.* **13**(1):81-94.
- Brünger AT. (1992) *Free R value : a novel statistical quantity for assessing the accuracy of crystal structures*. *Nature (London)*, **355**: 472-475.
- Brünger AT. (1993) *Assessment of phase accuracy by cross validation: the free R value. Methods and applications*. *Acta Cryst.* **D49**: 24-36.
- Buehner M, Ford GC, Moras D, Olsen KW and Rossmann MG. (1974) *Structure determination of crystalline lobster D-glyceraldehyde-3-phosphate dehydrogenase*. *J. Mol. Biol.* **82**:563-585.
- Burmeister WP. (2000) *Structural changes in a cryo-cooled protein crystal owing to radiation damage*. *Acta Crystallogr. Sect.* **D56**: 328-41.
- Byers LD. (1978) *Enantiomeric specificity of glyceraldehyde 3-phosphate dehydrogenase*. *Arch. Biochem. Biophys.* **186**(2): 335-342.
- Carroll PA, Zhao G, Boyko SA, Winkler ME and Calderwood SB. (1997) *Identification, sequencing, and enzymatic activity of the erythrose-4-phosphate dehydrogenase gene of Vibrio cholerae*. *J Bacteriol.* **179**(1):293-6.
- Castilho MS, Pavão F, Oliva G, Ladame S, Willson M and Périé J. (2003) *Evidence for the two phosphate binding sites of an analogue of the thioacyl intermediate for the Trypanosoma cruzi glyceraldehyde-3-phosphate dehydrogenase-catalyzed reaction, from its crystal structure*. *Biochemistry.* **42**:7143-51.
- Charron C, Talfournier F, Isupov MN, Littlechild JA, Branlant G, Vitoux B and Aubry A. (2000) *The crystal structure of d-glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic archaeon Methanothermus fervidus in the presence of NADP(+) at 2.1 Å resolution*. *J Mol Biol.* **297**:481-500.
- Clermont S, Corbier C, Mely Y, Gerard D, Wonacott A and Branlant G. (1993) *Determinants of coenzyme specificity in glyceraldehyde-3-phosphate dehydrogenase: role of the acidic residue in the fingerprint region of the nucleotide binding fold*. *Biochemistry.* **32**(38):10178-84.
- Cobessi D, Tête-Favier F, Marchal S, Branlant G and Aubry A. (2000) *Structural and biochemical investigations of the catalytic mechanism of an NADP-dependent aldehyde dehydrogenase from Streptococcus mutans*. *J. Mol. Biol.* **300**(1): 141-152.
- COLLABORATIVE COMPUTATIONAL PROJECT, Number 4. (1994) *Acta Crystallogr. Sect.* **D50**: 760-763.
- Conway A and Koshland DE (1968) *Negative cooperativity in enzyme action. The binding of diphosphopyridine nucleotide to glyceraldehyde 3-phosphate dehydrogenase*. *Biochemistry.* **7**(11): 4011-23.

- Corbier C (1989) Thèse
- Corbier C, Branlant C, Wonacott A, Branlant G. (1989) *Site-directed mutagenesis of glyceraldehyde-3-phosphate dehydrogenase reveals the role of residue Ser148*. Protein Eng. **2**(7):559-62.
- Corbier C, Mougin A, Mely Y, Adolph HW, Zeppezauer M, Gerard D, Wonacott A, Branlant G. (1990a) *The nicotinamide subsite of glyceraldehyde-3-phosphate dehydrogenase studied by site-directed mutagenesis*. Biochimie. **72**(8): 545-54.
- Corbier C, Clermont S, Billard P, Skarzynski T, Branlant C, Wonacott A, Branlant G. (1990b) *Probing the coenzyme specificity of glyceraldehyde-3-phosphate dehydrogenases by site-directed mutagenesis*. Biochemistry. **29**(30): 7101-6.
- Corbier C, Michels S, Wonacott AJ and Branlant G. (1994) *Characterization of the two anion-recognition sites of glyceraldehyde-3-phosphate dehydrogenase from Bacillus stearothermophilus by site-directed mutagenesis and chemical modification*. Biochemistry. **33**: 3260-65.
- Cowan-Jacob SW, Kaufmann M, Anselmo AN, Stark W, Grütter MG. (2003) *Structure of rabbit-muscle glyceraldehyde-3-phosphate dehydrogenase*. Acta Cryst **D59**: 2218-27.
- De Vijlder JJM, Boers JM and Slater EC. (1969) *Binding and properties of NAD<sup>+</sup> in glyceraldehydephosphate dehydrogenase from lobster-tail muscle*. Biochim. Biophys. Acta. **191**: 221-228.
- Delano WL. DeLano Scientific LLC, San Carlos, CA, USA. <http://www.pymol.org>
- Diamond R. (1969) *Profile Analysis in Single Crystal Diffractometry*. Acta Crystallogr. **A25**: 43-55.
- Didierjean C, Rahuel-Clermont S, Vitoux B, Dideberg O, Branlant G, Aubry A. (1997) *A crystallographic comparison between mutated glyceraldehyde-3-phosphate dehydrogenases from Bacillus stearothermophilus complexed with either NAD<sup>+</sup> or NADP<sup>+</sup>*. J Mol Biol. **268**(4):739-59.
- Didierjean, C., Corbier, C., Fatih, M., Favier, F., Boschi-Muller, S., Branlant, G. and Aubry, A. (2003) *Crystal structure of two ternary complexes of phosphorylating glyceraldehyde-3-phosphate dehydrogenase from Bacillus stearothermophilus with NAD and D-glyceraldehyde 3-phosphate*. J. Biol. Chem. **278**: 12968-76.
- Diederichs K and Karplus A. (1997) *Improved R-Factors for diffraction data analysis in macromolecular crystallography*. Nat Struct Biol. **4**(4):269-75. Erratum in: Nat Struct Biol **4**(7):592.
- Drenth, J. (1994). Principles of Protein X-ray Crystallography. New York: Springer-Verlag.
- Duée E, Olivier-Deyris L, Fanchon E, Corbier C, Branlant G and Dideberg O. (1996) *Comparison of the structures of wild-type and a N313T mutant of Escherichia coli glyceraldehyde 3-phosphate dehydrogenases: implication for NAD binding and cooperativity*. J. Mol. Biol. **257**:814-38.

- Emsley P and Cowtan K. (2004) *Coot: model-building tools for molecular graphics*. Acta Crystallogr. Sect. **D60**: 2126-2132.
- Engh R A and Huber R. (1991) *Accurate bond and angle parameters for X-ray protein structure refinement*. Acta Cryst. **A47**:392-400.
- Evans PR. (1993). Proceedings of the CCP4 Study Weekend. Data Collection and Processing, edited by L. Sawyer, N. Isaacs and S. Bailey, pp. 114±122. Warrington: Daresbury Laboratory.
- Eyschen J, Vitoux B, Marraud M, Cung MT and Branlant G. (1999) *Engineered glycolytic glyceraldehyde-3-phosphate dehydrogenase binds the anti conformation of NAD<sup>+</sup> nicotinamide but does not experience A-specific hydride transfer*. Arch Biochem Biophys. **364**(2):219-27.
- Fermani S, Sparla F, Falini G, Martelli PL, Casadio R, Pupillo P, Ripamonti A and Trost P. (2007) *Molecular mechanism of thioredoxin regulation in photosynthetic A2B2-glyceraldehyde-3-phosphate dehydrogenase*. Proc Natl Acad Sci. **104**(26):11109-14.
- Ferri G, Comerio G, Iadarola P, Zapponi MC and Speranza ML. (1978) *Subunit structure and activity of glyceraldehyde-3-phosphate dehydrogenase from spinach chloroplasts*. Biochim Biophys Acta. **522**(1):19-31.
- Fothergill-Gilmore LA and Michels PA. (1993) *Evolution of glycolysis*. Prog Biophys Mol Biol. **59**(2):105-235.
- Furfine CS, Velick SF. (1965) *The acyl-enzyme intermediate and the kinetic mechanism of the glyceraldehyde-3-phosphate dehydrogenase reaction*. J Biol Chem. **240**:844–855.
- Garman EF and Schneider TR. (1997) *Macromolecular Cryocrystallography* J. Appl. Cryst. . **30**:211-237.
- Gaussian 03, Revision **C.02**, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A.; Gaussian, Inc., Wallingford CT, 2004.
- Harrigan PJ and Trentham DR. (1973) *Kinetic studies of the acylation of pig muscle D-glyceraldehyde 3-phosphate dehydrogenase by 1,3-diphosphoglycerate and of proton uptake and release in the overall enzyme mechanism*. Biochem J. **135**(4): 695-703.

- Harrigan PJ and Trentham DR. (1974) *Kinetic studies on oxidized nicotinamide--adenine dinucleotide-facilitated reactions of D-glyceraldehyde 3-phosphate dehydrogenase*. *Biochem. J.* **143**(2): 353-63.
- Harris JI and Waters M. (1976) *The enzymes*, 3rd Ed., edited by P. D. Boyer, ch. 13. New York: Academic Press.
- Hillman JD and Fraenkel DG. (1975) *Glyceraldehyde 3-phosphate dehydrogenase mutants of Escherichia coli*. *J Bacteriol.* **122**(3): 1175-1179.
- Iglesias AA and Losada M. (1988) *Purification and kinetic and structural properties of spinach leaf NADP-dependent nonphosphorylating glyceraldehyde-3-phosphate dehydrogenase*. *Arch Biochem Biophys.* **260**(2):830-40.
- Isupov MN, Fleming TM, Dalby AR, Crowhurst GS, Bourne PC, Littlechild JA. (1999) *Crystal structure of the glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic archaeon Sulfolobus solfataricus*. *J Mol Biol.* **291**(3):651-60.
- Jancarik J and Kim SH. (1991) *Sparse matrix sampling: a screening method for crystallization of proteins*. *J. Appl. Cryst.* **24**: 409-411.
- Kabsch W. (1993) *Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants*. *J. Appl. Cryst.* **26**: 795-800.
- Kauffmann B. (2000) Manuscrit de DEA
- Kim H, Feil IK, Verlinde CL, Petra PH and Hol WG. (1995) *Crystal structure of glycosomal glyceraldehyde-3-phosphate dehydrogenase from Leishmania mexicana: implications for structure-based drug design and a new position for the inorganic phosphate binding site*. *Biochemistry.* **34**: 14975-86.
- Kim H, Hol WG. (1998) *Crystal structure of Leishmania mexicana glycosomal glyceraldehyde-3-phosphate dehydrogenase in a new crystal form confirms the putative physiological active site structure*. *J Mol Biol.* **278**(1):5-11.
- Kirschner K. (1971) *Co-operative binding of nicotinamide-adenine dinucleotide to yeast glyceraldehyde-3-phosphate dehydrogenase. II. Stopped-flow studies at pH 8.5 and 40 degrees C*. *J Mol Biol.* **58**(1):51-68.
- Kleywegt GJ and Brünger AT. (1996) *Checking your imagination: applications of the free R value*. *Structure.* **4**(8):897-904.
- Korndörfer I, Steipe B, Huber R, Tomschy A and Jaenicke R. (1995) *The crystal structure of holo-glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic bacterium Thermotoga maritima at 2.5 Å resolution*. *J. Mol. Biol.* **246**: 511-21.
- Krissinel E and Henrick K. (2007) *Inference of macromolecular assemblies from crystalline state*. *J. Mol. Biol.* **372**: 774-797.



- Ladame S, Castilho MS, Silva CH, Denier C, Hannaert V, Périé J, Oliva G, Willson M. (2003) *Crystal structure of Trypanosoma cruzi glyceraldehyde-3-phosphate dehydrogenase complexed with an analogue of 1,3-bisphospho-d-glyceric acid*. Eur J Biochem. **270**(22):4574-86.
- Lam HM and Winkler ME. (1990) *Metabolic relationships between pyridoxine (vitamin B6) and serine biosynthesis in Escherichia coli K-12*. J Bacteriol. **172**(11):6518-28.
- Laskowski RA, MacArthur MW, Moss DS and Thornton JM. (1993) *PROCHECK: a program to check the stereochemical quality of protein structures*. J. Appl. Crystallogr. **26**: 283-291.
- Leslie AG. (1999) *Integration of macromolecular diffraction data*. Acta Cryst. **D55**: 1696-1702.
- Liu L and Huskey WP. (1992) *Progress in establishing the rate-limiting features and kinetic mechanism of the glyceraldehyde-3-phosphate dehydrogenase reaction*. Biochemistry. **31**(30):6898-903.
- Lovell SC, Davis IW, Arendall WB, de Bakker PIW, Word JM, Prisant MG, Richardson JS and Richardson DC. (2003) *Structure validation by C-alpha geometry: phi, psi, and C-beta deviation*. Proteins. **50**:437-450.
- Marchal S, Rahuel-Clermont S, Branlant G. (2000) *Role of glutamate-268 in the catalytic mechanism of nonphosphorylating glyceraldehyde-3-phosphate dehydrogenase from Streptococcus mutans*. Biochemistry. **39**(12):3327-35.
- McPherson A, Malkin AJ and Kuznetsov YG. (1995) *The Science of macromolecular crystallization*. Structure. **3**:759-68.
- Michels S, Scagliarini S, Della Seta F, Carles C, Riva M, Trost P, Branlant G. (1994) *Arguments against a close relationship between non-phosphorylating and phosphorylating glyceraldehyde-3-phosphate dehydrogenases*. FEBS Lett. **339**(1-2): 97-100.
- Michels S, Rogalska E and Branlant G. (1996) *Phosphate-binding sites in phosphorylating glyceraldehyde-3-phosphate dehydrogenase from Bacillus stearothermophilus*. Eur. J. Biochem. **235**:641-47.
- Miller R, DeTitta GT, Jones R, Langs DA, Weeks CM and Hauptman HA. (1993). *On the application of the minimal principle to solve unknown structures*. Science. **259**:1430-1433.
- Moniot S. (2004) Manuscrit de DEA
- Moniot S, Bruno S, Vornrhein C, Didierjean C, Boschi-Muller S, Vas M, Bricogne G, Branlant G, Mozzarelli A, Corbier C. (2008) *Trapping of the Thioacylglyceraldehyde-3-phosphate Dehydrogenase Intermediate from Bacillus stearothermophilus: direct evidence for a flip-flop mechanism*. J Biol Chem. **283**(31):21693-702.
- Moras D, Olsen KW, Sabesan MN, Buehner M, Ford GC and Rossmann MG. (1975) *Studies of asymmetry in the three-dimensional structure of lobster D-glyceraldehyde-3-phosphate dehydrogenase*. J. Biol. Chem. **250**: 9137-62.

- Mougou A, Corbier C, Soukri A, Wonacott A, Branlant C, Branlant G. (1988) *Use of site-directed mutagenesis to probe the role of Cys149 in the formation of charge-transfer transition in glyceraldehyde-3-phosphate dehydrogenase*. Protein Eng. **2**(1):45-8.
- Mozzarelli A, Berni R, Rossi GL, Vas M, Bartha F and Keleti T. (1982) *Protein isomerization in the NAD<sup>+</sup>-dependent activation of beta-(2-furyl)acryloyl-glyceraldehyde-3-phosphate dehydrogenase in the crystal*. J. Biol. Chem. **257**(12): 6739-44.
- Mozzarelli A, and Rossi GL. (1996) *Protein function in the crystal*. Annu. Rev. Biophys. Biomol. Struct. **25**:343-65.
- Murshudov GN, Vagin AA and Dodson EJ. (1997) *Refinement of macromolecular structures by the maximum-likelihood method*. Acta Crystallogr. Sect. **D53**: 240-255.
- Navaza J. (1994) *AMoRe: an automated package for molecular replacement*. Acta Cryst. **A50**:157-163.
- Otwinowski Z, and Minor W. (1997) *Processing X-ray diffraction data collected in oscillation mode*. Methods Enzymol. **276**: 307–326.
- Pearson AR, Mozzarelli A and Rossi GL. (2004) *Microspectrophotometry for structural enzymology*. Curr. Opin. Struct. Biol. **14**(6): 656-62.
- Racker E and Krimsky I. (1952) *The mechanism of oxidation of aldehydes by glyceraldehyde-3-phosphate dehydrogenase*. J. Biol. Chem. **198**(2): 731-43.
- Ramakrishnan C and Ramachandran GN. (1965) *Stereochemical criteria for polypeptide and protein chain conformations. II. Allowed conformations for a pair of peptide units*. Biophys. J. **5**, 909-933.
- Ravelli RB and McSweeney SM. (2000) *The 'fingerprint' that X-rays can leave on structures*. Structure Fold. Des. **8**: 315-28.
- Read RJ. (1986) *Improved Fourier coefficients for maps using phases from partial structures with errors*. Acta Cryst. **A42**: 140-149.
- Roitel O, Sergienko E, Branlant G. (1999) *Dimers generated from tetrameric phosphorylating glyceraldehyde-3-phosphate dehydrogenase from Bacillus stearothermophilus are inactive but exhibit cooperativity in NAD binding*. Biochemistry. **38**(49):16084-91.
- Roitel O, Vachette P, Azza S, Branlant G. (2003 ) *P but not R-axis interface is involved in cooperative binding of NAD on tetrameric phosphorylating glyceraldehyde-3-phosphate dehydrogenase from Bacillus stearothermophilus*. J Mol Biol. **326**(5):1513-22.
- Rossmann MG, Adams MJ, Buehner M, Ford GC, Hackert ML, Liljas A, Rao ST, Banaszak LJ, Hill E, Tsernoglou D and Webb L. (1973) *Letter: Molecular symmetry axes and subunit interfaces in certain dehydrogenases*. J Mol Biol. **76**(4):533-7.
- Rossmann MG, Moras D, Olsen KW. (1974) *Chemical and biological evolution of nucleotide-binding protein*. Nature. **250**(463):194-9.

- Rossmann MGD and Blow M. (1962) *The detection of sub-units within the crystallographic asymmetric unit*. Acta Crystallogr. **15**:24-31.
- Segal HL and Boyer PD. (1953) *The role of sulfhydryl groups in the activity of D-glyceraldehyde 3-phosphate dehydrogenase*. J. Biol. Chem. **204**:265-81.
- Seta FD, Boschi-Muller S, Vignais ML and Branlant G. (1997) *Characterization of Escherichia coli strains with gapA and gapB genes deleted*. J Bacteriol. **179**(16): 5218-21.
- Seydoux F, Bernhard S, Pfenninger O, Payne M, Malhotra OP. (1973) *Preparation and active-site specific properties of sturgeon muscle glyceraldehyde-3-phosphate dehydrogenase*. Biochemistry. **12**(21): 4290-300.
- Sirover MA (1999) *New insights into an old protein: the functional diversity of mammalian glyceraldehyde-3-phosphate dehydrogenase*. Biochim. Biophys. Acta. **1432**: 159-184.
- Sirover MA (2005) *New Nuclear Functions of the Glycolytic Protein, Glyceraldehyde-3-Phosphate Dehydrogenase, in Mammalian Cells*. J. Cell. Biochem. **95**:45–52.
- Skarzyński T, Moody PC and Wonacott AJ (1987) *Structure of holo-glyceraldehyde-3-phosphate dehydrogenase from Bacillus stearothermophilus at 1.8 Å resolution*. J. Mol. Biol. **193**: 171-187.
- Skarzyński T and Wonacott AJ. (1988) *Coenzyme-induced conformational changes in glyceraldehyde-3-phosphate dehydrogenase from Bacillus stearothermophilus*. J Mol Biol. **203**(4):1097-1118.
- Song S, Li J, Lin Z. (1998) *Structure of holo-glyceraldehyde-3-phosphate dehydrogenase from Palinurus versicolor refined at 2 Å resolution*. Acta Crystallogr Sect. **D54**: 558-69.
- Song SY, Xu YB, Lin ZJ, Tsou CL. (1999) *Structure of active site carboxymethylated D-glyceraldehyde-3-phosphate dehydrogenase from Palinurus versicolor*. J Mol Biol. **287**(4):719-25.
- Soukri A, Mougin A, Corbier C, Wonacott A, Branlant C and Branlant G. (1989) *Role of the histidine 176 residue in glyceraldehyde-3-phosphate dehydrogenase as probed by site-directed mutagenesis*. Biochemistry. **28**: 2586-92.
- Souza DH, Garratt RC, Araújo AP, Guimarães BG, Jesus WD, Michels PA, Hannaert V, Oliva G. (1998) *Trypanosoma cruzi glycosomal glyceraldehyde-3-phosphate dehydrogenase: structure, catalytic mechanism and targeted inhibitor design*. FEBS Lett. **424**(3):131-5.
- Talfournier F, Colloc'h N, Mornon JP, Branlant G. (1998) *Comparative study of the catalytic domain of phosphorylating glyceraldehyde-3-phosphate dehydrogenases from bacteria and archaea via essential cysteine probes and site-directed mutagenesis*. Eur. J. Biochem. **252**: 447-57.

- Talfournier F, Colloc'h N, Mornon JP, Branlant G. (1999) *Functional characterization of the phosphorylating D-glyceraldehyde 3-phosphate dehydrogenase from the archaeon Methanothermus fervidus by comparative molecular modelling and site-directed mutagenesis.* Eur J Biochem. **265**(1): 93-104.
- Trentham DR. (1968) *Aspects of the chemistry of D-glyceraldehyde 3-phosphate dehydrogenase.* Biochem J. **109**(4): 603-12.
- Trentham DR. (1971a) *Reactions of D-glyceraldehyde 3-phosphate dehydrogenase facilitated by oxidized nicotinamide-adenine dinucleotide.* Biochem. J. **122**: 59-69.
- Trentham DR. (1971b) *Rate-determining processes and the number of simultaneously active sites of D-glyceraldehyde 3-phosphate dehydrogenase.* Biochem. J. **122**: 71-77.
- Usón I and Sheldrick GM. (1999) *Advances in direct methods for protein crystallography.* Curr Opin Struct Biol. **9**: 643-8.
- Vagin A and Teplyakov A. (1997) *MOLREP: an automated program for molecular replacement.* J. Appl. Crystallogr. **30**: 1022-1025.
- Weeks, C. M., DeTitta, G. T., Miller, R. & Hauptman, H. A. (1993) *Applications of the minimal principle to peptide structures.* Acta Cryst. **D49**: 179-181.
- Weiss MS. (2001) *Global indicators of X-ray data quality.* J. Appl. Cryst. **34**:130-135.
- Wierenga RK, Terpstra P, Hol WG. (1986) *Prediction of the occurrence of the ADP-binding beta alpha beta-fold in proteins, using an amino acid sequence fingerprint.* J Mol Biol. **187**(1):101-7.
- Yang Y, Zhao G, Man TK, Winkler ME. (1998) *Involvement of the gapA- and epd (gapB)-encoded dehydrogenases in pyridoxal 5'-phosphate coenzyme biosynthesis in Escherichia coli K-12.* J Bacteriol. **180**(16):4294-9.
- Yun M, Park CG, Kim JY and Park HW. (2000) *Structural analysis of glyceraldehyde 3-phosphate dehydrogenase from Escherichia coli: direct evidence of substrate binding and cofactor-induced conformational changes.* Biochemistry. **39**:10702-10.
- Zhao G, Pease AJ, Bharani N, Winkler ME. (1995) *Biochemical characterization of gapB-encoded erythrose 4-phosphate dehydrogenase of Escherichia coli K-12 and its possible role in pyridoxal 5'-phosphate biosynthesis.* J Bacteriol. **177**(10):2804-12.

# ANNEXE 1 : PUBLICATION

"Trapping of the Thioacylglyceraldehyde-3-phosphate  
Dehydrogenase Intermediate from *Bacillus stearothermophilus* :  
DIRECT EVIDENCE FOR A FLIP-FLOP MECHANISM"

Moniot *et al.*, 2008

# Trapping of the Thioacylglycerinaldehyde-3-phosphate Dehydrogenase Intermediate from *Bacillus stearothermophilus*

## DIRECT EVIDENCE FOR A FLIP-FLOP MECHANISM<sup>\*[5]</sup>

Received for publication, March 24, 2008, and in revised form, May 13, 2008. Published, JBC Papers in Press, May 14, 2008, DOI 10.1074/jbc.M802286200

Sébastien Moniot<sup>‡1</sup>, Stefano Bruno<sup>§</sup>, Clemens Vornrhein<sup>¶</sup>, Claude Didierjean<sup>‡</sup>, Sandrine Boschi-Muller<sup>||</sup>, Mária Vas<sup>\*\*</sup>, Gérard Bricogne<sup>¶</sup>, Guy Branlant<sup>||</sup>, Andrea Mozzarelli<sup>§++</sup>, and Catherine Corbier<sup>‡§§2</sup>

From the <sup>‡</sup>Laboratoire de Cristallographie et de Modélisation des Matériaux Minéraux et Biologiques, UMR CNRS-Université Henri Poincaré (UHP) 7036, the <sup>||</sup>Laboratoire de Maturation des ARN et Enzymologie Moléculaire, UMR CNRS-Université Henri Poincaré 7567, and the <sup>§§</sup>Unité de Recherche sur l'Animal et les Fonctionnalités des Produits Animaux, Équipe Protéolyse et Bio-fonctionnalités des Protéines et Peptides, EA 3998 Université Henri Poincaré, the Faculté des Sciences, Nancy Université, Vandoeuvre-lès-Nancy, 54506 Cedex, France, the <sup>§</sup>Department of Biochemistry and Molecular Biology and <sup>\*\*</sup>Unit of the National Institute for Biostructures and Biosystems, University of Parma, Viale Usberti 23/A, Parma 43100, Italy, the <sup>¶</sup>Global Phasing Ltd., Sheraton House, Castle Park, Cambridge CB3 0AX, United Kingdom, and the <sup>\*\*</sup>Institute of Enzymology, Biological Research Center, Hungarian Academy of Sciences, P. O. Box 7, Budapest H-1518, Hungary

The crystal structure of the thioacylenzyme intermediate of the phosphorylating glyceraldehyde-3-phosphate dehydrogenase (GAPDH) from *Bacillus stearothermophilus* has been solved at 1.8 Å resolution. Formation of the intermediate was obtained by diffusion of the natural substrate within the crystal of the holoenzyme in the absence of inorganic phosphate. To define the soaking conditions suitable for the isolation and accumulation of the intermediate, a microspectrophotometric characterization of the reaction of GAPDH in single crystals was carried out, following NADH formation at 340 nm. When compared with the structure of the Michaelis complex (Didierjean, C., Corbier, C., Fatih, M., Favier, F., Boschi-Muller, S., Branlant, G., and Aubry, A. (2003) *J. Biol. Chem.* 278, 12968–12976) the 206–210 loop is shifted and now forms part of the so-called “new P<sub>i</sub>” site. The locations of both the O1 atom and the C3-phosphate group of the substrate are also changed. Altogether, the results provide evidence for the flipping of the C3-phosphate group occurring concomitantly or after the redox step.

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH)<sup>3</sup> is a homotetrameric enzyme catalyzing the oxidative phosphorylation of D-glyceraldehyde 3-phosphate (G3P) into 1,3-bisphos-

phoglycerate (1,3-DPG), in the presence of inorganic phosphate (P<sub>i</sub>) and nicotinamide adenosine dinucleotide (NAD<sup>+</sup>).

The reaction mechanism has been intensively investigated in particular for bacterial and eukaryotic GAPDHs (1–9) and consists of two steps as follows: (i) an oxidoreduction reaction, corresponding to the nucleophilic attack of the catalytic cysteine (Cys-149) on the aldehydic group of G3P, followed by a hydride transfer assisted by His-176 (base catalyst) from the generated thiohemiacetal to the C4 of the nicotinamide of NAD<sup>+</sup> that leads to the formation of a thioacylenzyme (7), and (ii) a phosphorylation of the resulting thioester through the nucleophilic attack of inorganic phosphate on the carbonyl group of the thioacylenzyme. The second step is preceded by the exchange of NADH with NAD<sup>+</sup>, with the latter favoring the phosphorylation step.

Two anion recognition sites accommodate the inorganic phosphate ion and the phosphate groups of G3P and 1,3-DPG. Their positions within the active site have been deduced from the location of two sulfate ions derived from the ammonium sulfate crystallization medium (9). On the basis of a model of the thiohemiacetal intermediate in the *Homarus americanus* GAPDH structure, the anion binding sites were initially attributed to the specific binding of the C3-phosphate (C3P) group of D-G3P (P<sub>s</sub> site) and of the inorganic phosphate ion (P<sub>i</sub> site). The location of the P<sub>s</sub> site in the three-dimensional structures of eukaryotic and bacterial GAPDHs is conserved and independent of the enzyme state, apo-, or holo-form, and of the presence of ligands such as sulfate ions, phosphate ions, substrate, or substrate analogs. This P<sub>s</sub> site is composed of the side chains of residues Arg-231 and Thr-179 and the 2'-hydroxyl group of the nicotinamide ribose of NAD<sup>+</sup>. On the contrary, the location of the P<sub>i</sub> site appears to vary depending on the presence and nature of the bound ligands or source organism. This location is related to the conformation adopted by the segment composed of residues 206–212. Although the most common conformation is that originally found in the holoenzyme

\* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[5] The on-line version of this article (available at <http://www.jbc.org>) contains supplemental Fig. S1.

The atomic coordinates and structure factors (code 3cmc) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (<http://www.rcsb.org/>).

<sup>1</sup> Fellow of the French Ministère de la Recherche et des Nouvelles Technologies.

<sup>2</sup> To whom correspondence should be addressed: Tel.: 33-3-83-68-42-67; Fax: 33-3-83-68-40-01; E-mail: catherine.corbier@scbiol.uhp-nancy.fr.

<sup>3</sup> The abbreviations used are: GAPDH, D-glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12); G3P, D-glyceraldehyde 3-phosphate; 1,3-DPG, 1,3-bisphosphoglycerate; C3P group, C3-phosphate group; TAE, thioacylenzyme; r.m.s.d., root mean square deviation.

## Thioacylenzyme Intermediate from *B. stearothermophilus* GAPDH

from *H. americanus* and *Bacillus stearothermophilus* (9–10), an alternative conformation was first observed in the holoenzyme from *Thermotoga maritima* and *Leishmania mexicana* (11–12) that generates a  $P_i$  site located closer to the catalytic Cys-149 residue and 3 Å away from the former position, which has been called new  $P_i$  site. Whatever the considered position, the  $P_i$  site includes the side chains of residues Ser-148 and Thr-208 and the main-chain nitrogen of Gly-209. In addition, the new  $P_i$  site also involves residue Thr-150.

Despite numerous functional (6, 7, 13) and structural studies (14–16), the individual role of  $P_i$  and  $P_s$  sites in the catalytic events remains unclear. Although it is generally assumed from geometrical considerations that inorganic phosphate must attack the intermediate from the  $P_i$  site, a possible flipping of the C3P of G3P between the two sites during catalysis was suggested by Skarzynski *et al.* (10).

Recently, using mutants of the active site cysteine, Didierjean *et al.* (16) obtained the structures of ternary complexes from *B. stearothermophilus* GAPDH with  $NAD^+$  and the physiological substrate D-G3P, the phosphate group of which was found to be bound at the  $P_s$  site. This result strongly suggests that the  $P_s$  site constitutes the binding site for the C3P group of the substrate but does not, however, exclude the possibility that the substrate could display different interactions especially with regard to its C3P group, once the covalent bond is formed between Cys-149 and D-G3P. Indeed this structure is representative of a Michaelis, non-covalent, complex and does not allow to prefigure the location of the C3P group in the intermediate stages of the reaction.

Much of the uncertainty regarding the individual contribution of the two anion binding sites during the catalytic events derives from the absence of the structure of the physiological thioacylenzyme intermediate. To obtain this structure, holo-GAPDH crystals from *B. stearothermophilus* were soaked with the substrate G3P in the absence of inorganic phosphate. Formation of NADH in the crystal was monitored by single crystal absorption microspectrophotometry (17), a method pioneered by Rossi and Bernhard (18), for the determination of protein function in the crystalline state (19) and for the design of kinetic crystallography experiments (20). A series of microspectrophotometric studies on GAPDH crystals were previously carried out using a chromophoric substrate analog of 1,3-DPG (21–23). The experimental conditions suitable for the accumulation of the catalytically-relevant thioacylenzyme intermediate in the crystals of *B. stearothermophilus* GAPDH were thus designed and allowed us to solve its structure. This constitutes the first structure of a reaction intermediate of GAPDH in the presence of its cofactor  $NAD^+$ . Analysis of the acylated structure reveals two major outcomes when compared with that of the non-covalent complex: a shift of the 206–210 loop, which delineates a new position for the  $P_i$  site comparable to that originally described for the holo-structure of *Thermotoga maritima*, and the location of the C3P group of the thioacyl intermediate, which is bound in this new  $P_i$  site.

## EXPERIMENTAL PROCEDURES

**Chemicals**—DL-Glyceraldehyde-3-phosphate (DL-G3P, Sigma) was prepared from the DL-G3P diethylacetal according to the manufacturer, and its concentration was assessed using GAPDH. Iodoacetic acid, ammonium sulfate, EDTA,  $NAD^+$ , Tris, and dithiothreitol were from Sigma and were of the best available quality and used without further purification.

**Production, Purification, and Crystallization of Wild-type *B. stearothermophilus* GAPDH**—The *Escherichia coli* strain used for wild-type GAPDH production was HB101 transformed with a pBluescript II SK containing the *gap* gene under the *lac* promoter. Purification was performed as previously described by Talfournier *et al.* (7). Purity of the enzyme was checked by electrophoresis on a 10% SDS-PAGE and by mass spectrometry.

Prior to crystallization, the enzyme was dialyzed overnight against 100 mM Tris buffer, pH 8.0, 2 mM EDTA, and 2 mM dithiothreitol, and concentrated using a 30-kDa molecular mass cut-off centrifugation system (Centricon). Protein concentration was determined spectrophotometrically at 280 nm using a molecular mass of 143,776 Da and an extinction coefficient of  $1.31 \times 10^5 \text{ M}^{-1} \text{ cm}^{-1}$ .

Crystals of the holo-form of wild-type *B. stearothermophilus* GAPDH were obtained at 293 K using the hanging-drop, vapor diffusion method in the presence of ammonium sulfate as precipitating agent. The reservoir solution contained 2.7 M ammonium sulfate, 100 mM Tris-HCl buffer, pH 6.9. Drops were prepared by mixing in different ratios a solution containing 15 mg·ml<sup>-1</sup> enzyme and 2 mM  $NAD^+$  and the reservoir solution. Crystals appear within 1 week from drops with mixed volumes of protein and reservoir solutions in ratios of 1/2, 1/1, or 2/1.

**Single Crystal-polarized Absorption Microspectrophotometric Measurements**—Single crystals of GAPDH were resuspended at least six times in a stabilizing solution containing 2.7 M ammonium sulfate, 0.1 M Tris-HCl buffered at pH 8.2. Crystals were loaded in a quartz flow cell mounted on the stage of a Zeiss MPM03 microspectrophotometer, equipped with a 10× ultrafluar objective and a thermostatic apparatus (24). Polarized absorption spectra were recorded between 300 and 500 nm, with the electric vector of the linearly polarized light parallel to crystal edges.

To follow the reaction of GAPDH crystals with D-G3P, crystals were suspended in the stabilizing solution containing DL-G3P at concentrations up to 60 mM. Polarized absorption spectra were recorded as a function of time to monitor reaction equilibration. To obtain the reference spectrum for the fully reduced NADH-enzyme complex, single crystals were treated with 1 mM sodium arsenate in the presence of G3P. The arsenate anion irreversibly reacts with the thioacylenzyme, shifting the equilibrium toward the full reduction of  $NAD^+$ . At each G3P concentration, the fraction of reduced NADH was calculated by fitting the corresponding absorption spectrum to a linear combination of the spectra of the native form and the arsenate-treated form. As a control experiment, enzyme crystals were suspended in a solution containing 10 mM iodoacetic acid, 2.7 M ammonium sulfate, 0.1 M Tris-HCl buffer, pH 8.2. This experiment was aimed to demonstrate the disappearance of the Racker band, diagnostic of an active enzyme, due to the

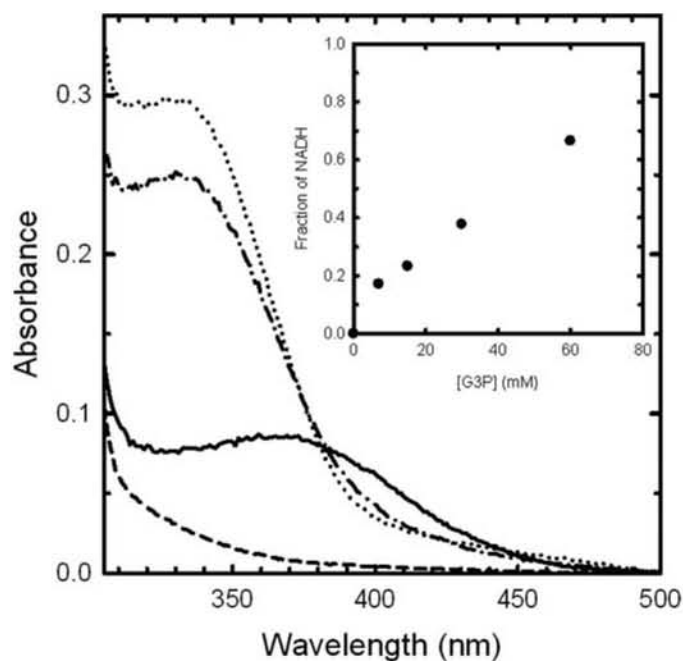
alkylation of the reactive Cys-149. All experiments were carried out at 15 °C.

**X-ray Diffraction Experiments**—Crystals of the holoenzyme were soaked for 1.5–10 min in a crystallization medium containing 60 mM DL-G3P. Reaction was stopped by flash freezing crystals after a quick soak in a cryoprotective solution containing 20% (v/v) glycerol in addition to the crystallization medium. Several datasets were collected at 100 K using a synchrotron beam either at European Synchrotron Radiation Facility (Grenoble, France) or Deutsches Elektronen-Synchrotron (Hamburg, Germany). 360 frames each of 0.5° oscillation step were systematically collected to assure collection of complete and redundant datasets, which were processed using the XDS package (25) and scaled using SCALA (26). The structure reported here (see below) derived from a crystal, soaked for 3 min, which diffraction data were collected at the X13 beamline (Deutsches Elektronen-Synchrotron) using a MARCCD detector (165 mm) and a wavelength of 0.806 Å. Statistics of the corresponding 1.77-Å resolution dataset are summarized in Table 1. Holo-crystals belong to monoclinic space group  $P2_1$  with typical unit cell dimensions of  $a = 82$  Å,  $b = 123$  Å,  $c = 82$  Å, and  $\beta = 111^\circ$ , containing one physiological homotetramer per asymmetric unit (as previously described by Skarzynski *et al.* (10)).

**Phasing and Refinement**—The structure was solved by molecular replacement using the MOLREP program (27) of the CCP4 suite (28) with the wild-type holo-structure (Protein Data Bank code 1gd1 (10)) as a starting model and data up to 3.6-Å resolution. Cycles of refinement (Refmac5) alternated with manual rebuilding (Coot) were carried out to improve the model (29, 30). Water molecules were added using the built-in find-water function of Coot and individually checked for significant signal and consistent contact with H-bond donor/acceptor. The thioacylenzyme model obtained from soaked ammonium sulfate-grown crystal was refined up to final convergence, using all the reflections ( $\sigma$  cut-off = 0) with  $R$  and  $R_{\text{free}}$  of 18.2% and 23.9%, respectively. The final cycles of refinement were performed using the BUSTER-TNT program (31), which allowed refinement of occupancy factors of the catalytic cysteine, reaction intermediate, 206–212 loop, and sulfate anions within the four monomers. This gave a final  $R$  and  $R_{\text{free}}$  of 16.1% and 19.8%, respectively.

**Structure Analysis and Final Structure Parameters**—The geometry of the model was checked with PROCHECK (32). Statistics concerning the geometry of the final model are given in Table 1. All non-glycine residues are located in favorable regions of the Ramachandran plot (33) except Asp-186 and Val-237, two residues already reported to be outside the allowed regions of the Ramachandran plot in the crystal structures of GAPDH isolated from *B. stearotherophilus* and other sources (34). The coordinates and structure factors of the thioacylenzyme structure have been deposited to the Protein Data Bank at Research Collaboratory for Structural Bioinformatics (ID code 3cmc). Figs. 2–4 were drawn with PyMOL.<sup>4</sup>

<sup>4</sup> W. L. DeLano (2002) PyMOL, DeLano Scientific, San Carlos, CA.



**FIGURE 1. Polarized absorption spectra of single crystals of *B. stearotherophilus* holo-GAPDH in the absence and presence of G3P and iodoacetic acid.** GAPDH crystals, grown in ammonium sulfate, were soaked in a stabilizing solution containing 2.7 M ammonium sulfate, 0.1 M Tris-HCl, buffered at pH 8.2 (solid line). The spectra were recorded with the light polarized perpendicular to the  $b$  axis of the crystal. The crystal was then resuspended three times with the same solution containing 10 mM iodoacetic acid (dashed line). Within the time required for the washing steps, a complete disappearance of the Racker band was observed. A different crystal was soaked with the stabilizing solution containing increasing concentrations of G3P (10, 15, 30, and 60 mM). The absorption spectrum obtained at 60 mM G3P is reported (dash-dot line). Further reduction of  $\text{NAD}^+$  was obtained by soaking enzyme crystals with the stabilizing solution containing 60 mM G3P and 1 mM sodium arsenate (dotted line). The fractional reduction of  $\text{NAD}^+$  by G3P was calculated at different G3P concentrations by fitting the observed spectra to a linear combination of the spectrum of the native form and the arsenate-treated form. *Inset:* dependence of the fractional reduction of  $\text{NAD}^+$  as a function of G3P concentration (10, 15, 30, and 60 mM). Spectra recorded on different crystals were normalized using the spectrum of the holoenzyme as reference.

## RESULTS

**Polarized Absorption Spectra of *B. stearotherophilus* Holo-GAPDH Crystals in the Absence and Presence of G3P**—Polarized absorption spectra of holo-GAPDH crystals from *B. stearotherophilus* grown in ammonium sulfate exhibited a broad band centered at around 360 nm (Fig. 1). This spectral feature, known as the Racker band, is associated with a charge transfer between the catalytic cysteine residue and the  $\text{NAD}^+$  nicotinamide ring (35). Carboxymethylation of GAPDH crystals at the S $\gamma$  of Cys-149 with iodoacetic acid resulted in the disappearance of the band (Fig. 1), as in solution, indicating that GAPDH in the crystal is in a catalytically active conformation. A similar observation was obtained on GAPDH crystals from *Palinurus vulgaris* (21).

Upon addition of the physiological substrate G3P to holoenzyme- $\text{NAD}^+$  crystals, an intense band centered at around 340 nm, associated with the formation of  $\text{NADH}$ , appeared with the concomitant disappearance of the Racker band (Fig. 1). This result indicates that the redox reaction has occurred in the crystalline state, with the hydride transfer from C1 of G3P to C4 of nicotinamide ring and concom-



## Thioacylenzyme Intermediate from *B. stearothermophilus* GAPDH

itant formation of the thioacylenzyme-NADH complex, as observed in solution (2–3, 36). The reduction of all four NAD<sup>+</sup> molecules of the tetramer could be achieved by adding sodium arsenate to the soaking solution containing G3P (Fig. 1). The arsenate anion irreversibly reacts with the acylenzyme, thus shifting the equilibrium toward the reduction of NAD<sup>+</sup>. The fraction of acylated sites and NADH molecules depends on the concentration of G3P in the soaking solution (Fig. 1, *inset*). A titration of GAPDH crystals with increasing concentrations of G3P revealed a relatively low apparent affinity for GAPDH with respect to solution (Fig. 1, *inset*), likely due to the inhibitory effect of the high sulfate concentration in the suspending medium, competing for the anion binding sites. Indeed, a similar behavior was observed for *P. vulgaris* GAPDH crystals, grown in ammonium sulfate (data not shown), whereas on the contrary, crystals of *B. stearothermophilus*, grown from polyethylene glycol solutions, reacted with G3P with complete reduction of NAD<sup>+</sup> at ~1 mM G3P (data not shown). However, the reaction of G3P with these polyethylene glycol-grown

GAPDH crystals was systematically accompanied by a significant decrease of the diffraction power, preventing their use for structural determination.

**Overall Structure and Model Quality**—The final model is composed of one tetramer in the asymmetric unit, each monomer containing residues 0–333 and one molecule of cofactor. The asymmetric unit also contains 4 bound G3P molecules, 18 sulfate ions (including two in each active site), 8 molecules of glycerol, and 1 molecule of ethylene glycol derived from the cryoprotective treatment (see “Experimental Procedures”) and 1743 water molecules. Except for some solvent-exposed side chains, residues of the model are well defined in the electron density maps. Only a few residual peaks, mostly in the solvent-exposed region, were not interpreted in the difference maps. The model has tight stereochemical restraints with r.m.s.d. on bond lengths and on bond angles of 0.011 Å and 1.27°, respectively (Table 1).

The four monomers are named O, P, Q and R, following the nomenclature initially described for *H. americanus* GAPDH (37). Using the O subunit as a reference, the root-mean-square deviations (r.m.s.d.) of backbone atoms after superimposition of the P, Q, and R subunits are 0.15, 0.14, and 0.12 Å, respectively, showing that the conformations of all four subunits in the final refined tetramer are very similar. Superimposition of the backbone atoms of our model onto the holo-structure described by Skarzynski *et al.* (10) (same space group and unit cell dimensions; pdb code 1gd1) gives an r.m.s.d. of 0.45 Å for the whole tetramer and a mean value of 0.27 Å for the different pairwise superimpositions of the monomers. No major difference was observed between the two structures except locally for the 206–212 loop, which exhibits deviations three times higher than the r.m.s.d. This conformational difference will be further described below.

As expected from the stoichiometry of the reaction of GAPDH crystals with 60 mM DL-G3P (fraction of NADH per monomer of ~0.70), determined by polarized absorption microspectrophotometry (Fig. 1, *inset*), the active site of each subunit was found to be present in two distinct, differently populated chemical states, which have been assumed to correspond to the native holo-structure with a free cysteine and a sulfate anion bound to the P<sub>i</sub> site (mean refined fractional occupancy of ~0.4, conformation A) and to the thioacylenzyme-NADH intermediate (mean refined fractional occupancy of ~0.6, conformation B) (Table 2). The refined occupancies and B values of the ligands are similar in the four monomers and support an equivalent distribution of the thioacylenzyme intermediate in all subunits. The differences between the two states are also

**TABLE 1**  
Data collection and refinement statistics

Values in parentheses refer to the outermost resolution shell.

Thioacylenzyme structure (soaked ammonium sulfate crystal)	
<b>Data collection</b>	
Space group	P2 <sub>1</sub>
Unit cell (Å)	<i>a</i> = 81.7; <i>b</i> = 122.6; <i>c</i> = 81.8; β = 111.1
Z	2
Nominal resolution (Å)	1.77
Outermost resolution shell (Å)	1.88–1.77
Temperature (K)	100
Unique reflections	142,733 (20,259)
Completeness (%)	97.7 (92.9)
Redundancy	3.7 (3.2)
R <sub>merge</sub> <sup>a</sup> (%)	2.8 (18.9)
Mean I/σ(I)	32.1 (6.5)
<b>Refinement</b>	
R-factor <sup>b</sup> (%)	16.1
R <sub>free</sub> <sup>c</sup> (%)	19.8
r.m.s.d. from ideal geometry	
Bond lengths (Å)	0.011
Bond angles (°)	1.27
Dihedral angles (°)	17.0
Improper angles (°)	26.0
Mean B-factor (Å <sup>2</sup> )	
Overall model	25.7
Protein atoms	23.6
NAD	17.1
Reaction intermediate	27.4
Sulfate atoms	40.4 (45.8–34.7)
(P <sub>i</sub> and P <sub>s</sub> site sulfates)	
Water molecules	37.7

<sup>a</sup> R-factor for symmetry-related intensities.

<sup>b</sup> Crystallographic R-factor.

<sup>c</sup> Crystallographic R-factor for a randomly selected 5% of reflections not included in refinement.

**TABLE 2**  
Occupancy factor refinement and thermal agitation values

	Subunit	O	P	Q	R
		Thioacyl intermediate	Refined occupancy Mean B value (Å <sup>2</sup> )	0.64 27.3	0.59 27.9
Sulfate (P <sub>i</sub> site)	Real space correlation coefficient <sup>a</sup>	0.95	0.91	0.94	0.94
	Occupancy B value (Å <sup>2</sup> )	0.33 46.1	0.47 45.6	0.41 45.1	0.27 46.3
206–212 loop (confA/confB)	Refined occupancy	0.47/0.53	0.66/0.34	0.65/0.34	0.44/0.56
	Mean B value (Å <sup>2</sup> )	24.8/25.6	24.5/25.2	23.8/24.5	24.9/25.5
	Real space correlation coefficient <sup>a</sup>	0.93/0.87	0.95/0.90	0.95/0.89	0.94/0.88

<sup>a</sup> Correlation existing between model density and 2mF<sub>o</sub> – DF<sub>c</sub> density.

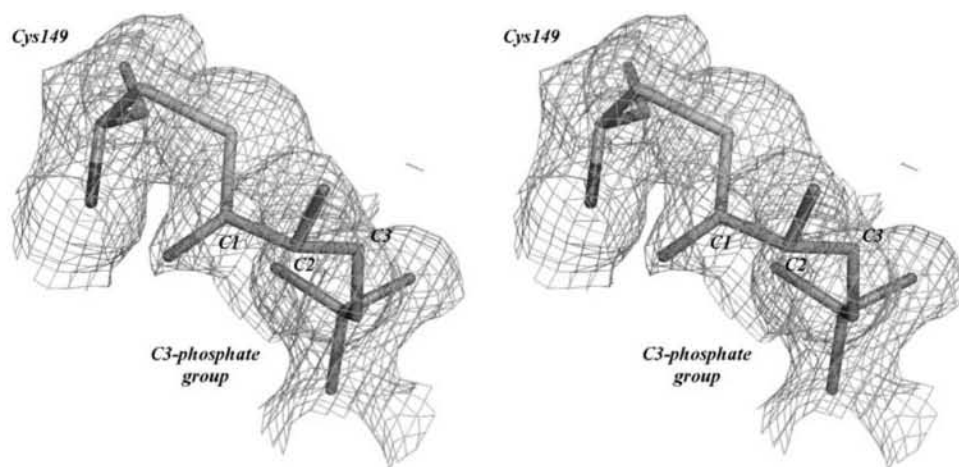


FIGURE 2. **Stereoscopic view of the thioacylzyme intermediate (subunit O).** The catalytic cysteine (Cys-149) and the bound substrate are presented in stick mode in  $2mF_o - DF_c$  omit maps contoured at  $1\sigma$ . These maps were obtained by removing from the model all atoms of the protein, NAD, sulfate, and water molecules within 3.6 Å of the thioacylzyme intermediate residues. This omit model was refined again with Buster, eliminating the excluded regions also from the bulk solvent part of the model. Continuity of the density covering all atoms of the substrate is satisfying especially regarding the relatively low occupancy of these atoms (0.64). Note the trans conformation retained by the two oxygen atoms bound to C1 and C2, respectively.

reflected in the 206–212 loop that was built in the electron density maps in two alternate conformations, A and B, each associated with one of the two states presented above. Refinement of the occupancies of both 206–212 loop alternate conformations shows the existence of two subunit types. O and R subunits exhibit almost equal occupancies for the two conformations, whereas the native conformation is prevailing in P and Q (Table 2). Although this loop is not directly involved in crystal packing, such difference could arise from the distinct crystalline environment of each subunit extensively described by Skarzynski *et al.* (10). Despite these slight occupancy variations, conformations of the 206–212 loop and of bound ligands are similar in the four monomers. Therefore, all following descriptions and comparisons, except when specified, will be based on the structure of the O monomer.

**Conformation of the Thioacyl-GAPDH-NADH Intermediate**—The thioacylzyme intermediate is present in each subunit with a refined occupancy factor varying from 0.59 to 0.67. It is well defined with good connectivity in monomers O and R and a real space correlation coefficient of 0.94 (Table 2 and Fig. 2). In monomers P and Q, density is poorer with a lack of signal at the C2 atom level (Table 2 and supplemental materials). Although no NCS restraints were applied on the bound substrate, the conformation of the intermediate and the interactions shared with the protein structure are almost identical in the four monomers.

The 3-phosphoglyceroyl moiety is bound via a planar thioester bond formed between the  $sp^2$  C1 atom and the  $S\gamma$  atom of the catalytic cysteine (Cys-149) (Fig. 2). The plane of the thioester bond is almost parallel to the one defined by the nicotinamide ring, and the carbonyl oxygen (O1) points in a direction opposite to the catalytic His-176 to accept a H-bond from the Cys-149 main chain and from a water molecule (Fig. 3A). Whereas DL-G3P was used for soaking experiments, the C-2 carbon is found in an R configuration, which accounts for the enantiomer specificity of the GAPDH enzyme toward D-G3P

(38). The C2 hydroxyl group points away from the carbonyl group at C1 to adopt a trans conformation. The O2 oxygen atom is assumed to form H-bonds with both the NO7 atom of the cofactor and with the Ne atom of His-176. The C3P group is located in the new  $P_i$  site, formed by the 206–212 loop under its B conformation (see below). Its oxygen atoms are hydrogen-bonded to the side chains of Ser-148, Thr-150, and Thr-208 and to the main chain of Thr-150 and Gly-209, all of these residues being invariant. Additional interactions are provided by the non-conserved Arg-195 and by the invariant Arg-231 via two water molecules (see Table 3 for interaction distances).

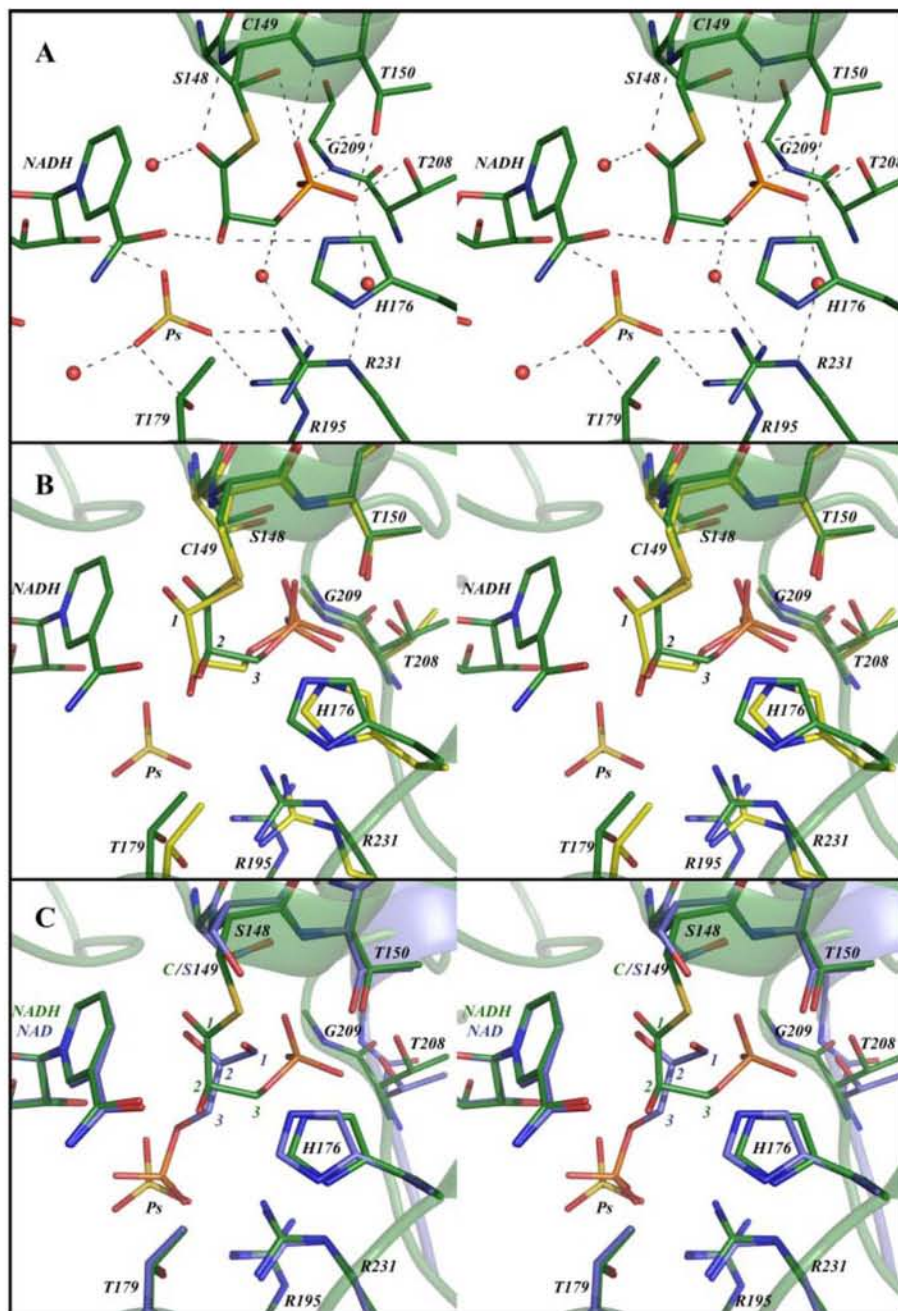
Given the refined fractional occupancy of this intermediate (Table 2),

one has to assume that Cys-149 also exists in its free state. Based on the shape of the density map and on refinements, this residue was built under a unique conformation that is representative of both states. The side chain of the catalytic cysteine superimposes perfectly with that of the known holo-structure from *B. stearotherophilus*. In its free form, this side chain is well positioned to form an ion pair with the Ne atom of His-176 distant of 3.6 Å.

In each monomer, the cofactor molecule is well defined in density maps with full occupancy. Because the reaction with G3P in the crystal was not complete, a mixture of oxidized and reduced forms is expected for the cofactor. However, the two redox states cannot be distinguished from the 1.77-Å electron density map, and there is no evidence to support any change of conformation for the cofactor molecule after hydride transfer.

A sulfate ion is bound with full occupancy in the  $P_s$  site through five hydrogen bonds to Arg-231, Arg-195, and Thr-179 side chains, the 2'-OH atom of the ribose adjacent to the nicotinamide, and a water molecule (Fig. 3A). A second sulfate ion is located in the  $P_i$  site with partial occupancy (see above). It interacts with the side chains of Thr-208 and Ser-148 and with the main-chain N atom of Gly-209 (conformation A of the 206–212 loop). This anion exhibits a high  $B$  value ( $45.8 \text{ \AA}^2$ ) in comparison with the model mean value ( $25.7 \text{ \AA}^2$ ) and is poorly defined in electron density maps. However, a water molecule cannot account for the observed electron density, and the location of this anion matches the one bound in the  $P_i$  site of the holoenzyme structure (10).

**Alternate Conformations of the 206–212 Loop**—Delineating one side of the active site, the 206–212 loop contributes to the formation of one of the two anion binding sites. As mentioned above, this loop has been unequivocally fitted into the density maps under two conformations (Fig. 4A and Table 2). The first conformation, called conformation A, is strictly equivalent to the conformation described in the holo-structure (10, 16). In this conformation, the loop participates in the binding of a sul-



**FIGURE 3. Active site and comparison with other representative complexes.** The reaction intermediates, the substrate and cofactor molecules, and the residues involved in catalysis and substrate binding are presented in *stick mode*. Labels 1, 2, and 3 refer to the C1, C2, and C3 atoms of the substrate, respectively. *A*, stereoscopic view of the active site of the thioacyl intermediate structure. Hydrogen bonds are presented as *dashed lines*. Important water molecules are shown as *spheres*. *B*, stereoscopic view of the superimposition between this model (*green*) and the hemithioacetal structure of *E. coli* apo-GAPDH (*yellow*) (1DC4) (14). *C*, stereoscopic view of the superimposition between this model (*green*) and the Michaelis ternary complex from *B. stearothermophilus* holo-GAPDH (C149S mutant, in *blue*) (1NQO) (16).

fate anion in the “classic”  $P_i$  site. The second conformation of the loop (conformation B), with complementary occupancy, displays a maximum shift of 3.4 Å for the Gly-209  $C\alpha$  atom toward the catalytic cysteine. In this alternate position, the 206–212 loop is located too close to the sulfate molecule to allow both of them to coexist in the same position, confirming that the presence of the sulfate anion is associated with the above described loop conformation A (Fig. 4A). The loop in conformation B is very similar to the one initially described for

*T. maritima* GAPDH (see Fig. 4B) (11) and contributes to the formation of the anion binding site called new  $P_i$ . Both conformations have nearly identical thermal motion factors (24.5 and 25.2 Å<sup>2</sup>, for conformations A and B, respectively), which fall in the same range of the overall protein (23.6 Å<sup>2</sup>), suggesting that both states of the 206–212 loop are stable conformations.

**Comparison with Other Structures**—The position of the new  $P_i$  binding site, here observed for the first time in *B. stearothermophilus* GAPDH, is equivalent to that described for structures of GAPDH from other sources (11–12, 14). Superimposition of the main-chains atoms of the O subunit of our model with those of the hemithioacetal structure of *E. coli* apo-GAPDH (pdb code 1dc4, A chain (14)) locates the two phosphorus atoms only 0.33 Å away from each other (Fig. 3B). As well, the central atoms of the sulfate and phosphate anions bound in the new  $P_i$  site of *T. maritima* (pdb code 1hdg, O chain (11)) and *L. mexicana* (pdb code 1gyp, A chain (12)), respectively, after similar superimposition procedures, are situated only 0.17 and 0.42 Å away from the phosphorus atom of the thioacylzyme intermediate, respectively. The position and geometry of the substrate in this intermediate are very close to those described by Yun *et al.* (14) for the hemithioacetal structure (pdb code 1dc4) of *E. coli* apo-GAPDH. Only a slight shift can be observed between the two bound G3P molecules (Fig. 3B) that can be related, first, to the different hybridizations of the two C1 atoms, and second, to the steric influence of the cofactor molecule, which is missing in the hemithioacetal apostructure. As previously

described by Yun *et al.* (14) for the hemithioacetal intermediate, the oxygen bound to C1 points away from the hydroxyl group of C2 to adopt a trans conformation. Thus, the O1 atom of G3P turns away from the catalytic His-176 in a counterintuitive manner with respect to its demonstrated catalytic role (see “Discussion”).

When the thioacyl-GAPDH structure is compared with the non-covalent Michaelis complex obtained by Didierjean *et al.* (16) for *B. stearothermophilus* GAPDH (pdb code 1nqo), the

TABLE 3

## Substrate-protein interactions

Numbers in parentheses refer to the water molecule numbers as defined in the atomic coordinate file.

	O	P	Q	R
O1-G3P-N-Cys-149	2.7	2.6	2.7	2.7
O1-G3P-H <sub>2</sub> O	3.0 (W558)	2.7 (W1547)	3.0 (W1056)	2.7 (W1213)
O2-G3P-NO7-NAD	2.7	2.7	2.7	2.8
O2-G3P-Nε-His-176	3.1	3.1	3.2	3.1
O1P-G3P-H <sub>2</sub> O	3.0 (W395)	3.0 (W631)	2.9 (W1743) 3.0 (W367)	3.0 (W111)
O2P-G3P-N-Gly-209	2.3	2.4	2.4	2.5
O2P-G3P-H <sub>2</sub> O	/	/	3.2 (W367)	3.1 (W111) 3.1 (W1701)
O3P-G3P-Oγ <sub>1</sub> -Thr-150	2.6	2.5	2.5	2.5
O3P-G3P-Oγ <sub>1</sub> -Thr-208	2.7	2.8	2.9	2.7
O3P-G3P-H <sub>2</sub> O	2.8 (W181)	2.7 (W253)	2.6 (W189)	2.8 (W216)
O4P-G3P-Oγ-Ser-148	3.0	3.0	3.2	2.9
O4P-G3P-N-Thr-150	2.9	3.1	3.0	3.0
O4P-G3P-Oγ <sub>1</sub> -Thr-150	2.7	2.8	2.6	2.7

main difference concerns the binding of the C3P group, which is located in the P<sub>s</sub> site in the non-covalent complex (Fig. 3C), ~6 Å away from the new P<sub>i</sub> site. This shift of the C3P group is also accompanied by a complete rearrangement of the carbon chain, which results in significant modifications of its interaction pattern with the protein. Thus, the orientations of the C1–O1 and C2–O2 bonds differ drastically in the two structures. Although the O1 atom points toward the catalytic histidine and the C2–O2 bond is almost perpendicular to the nicotinamide ring in the Michaelis complex, the O1 atom points away from the histidine and the C2–O2 bond is almost parallel to the nicotinamide ring in the thioacylenzyme.

## DISCUSSION

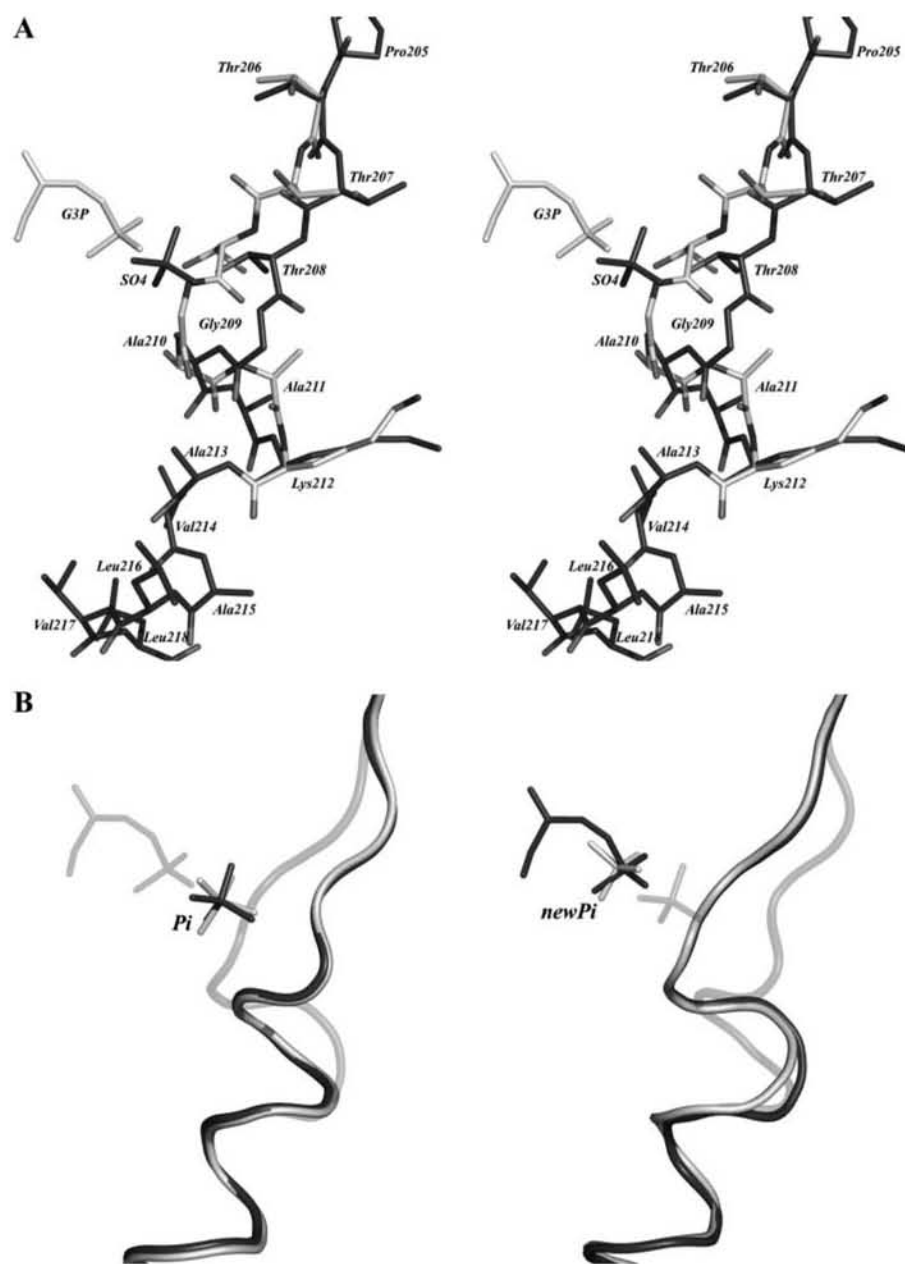
This report presents for the first time the crystal structure of a thioacylenzyme intermediate of GAPDH. The strategy used to obtain this intermediate takes advantage of the slow rate of the deacylation step in the absence of phosphate ions. Under this condition, the rate of hydrolysis of the intermediate is 10<sup>-2</sup> s<sup>-1</sup>, i.e. 6 × 10<sup>4</sup> times slower than acylation (6), which allows the intermediate to accumulate in the crystal. Despite the partial occupancy of the intermediate, the quality of the diffraction data and of the resulting structure allowed us to build two distinct states of the enzyme, which are each present in the four monomers of the enzyme. The most populated state corresponds to the structure of the thioacylenzyme intermediate covalently bound to the catalytic cysteine and the other one corresponds to a state in which the catalytic cysteine is free and a sulfate ion is bound in the classic P<sub>i</sub> site. Each state is associated with a particular conformation of the 206–212 loop in which invariant Thr-208 and Gly-209 interact with either the C3P of the thioacyl intermediate or a sulfate anion.

**Structure of the Thioacylenzyme Intermediate and Implications for Catalysis**—The phosphate group of the thioacyl intermediate is bound in the new P<sub>i</sub> site. This result is consistent with the location of the substrate phosphate group found in the structure of the hemithioacetal (*sp*3) intermediate obtained with the apo-GAPDH from *E. coli*: the two structures superimpose well (Fig. 3B) except at C1 due to different hybridization (*sp*2 instead of *sp*3). On the other hand, structures considered to be representative of a productive enzyme-NAD-D-G3P non-covalent complex (16) showed that the P<sub>s</sub> site constitutes the binding site for the C3P group of the substrate in this Michaelis

complex. Taken altogether, these results suggest that, even in the presence of a fully formed P<sub>s</sub> site (presence of the ribose of the cofactor to interact with the P<sub>s</sub> site, a condition that was not fulfilled for the hemithioacetal structure from *E. coli* GAPDH in which the enzyme was under its apo-form), the C3P group of the substrate has shifted from P<sub>s</sub> site to the new P<sub>i</sub> site during the oxidoreduction step. The exact stage at which this shift occurs remains unknown. However, although one cannot rule out the possibility that the formation of the covalent bond between Cys-149 and D-G3P promotes the repositioning of the C3P in the new P<sub>i</sub> site, it seems more likely that the first step of the reaction occurs while the C3P group is bound at the P<sub>s</sub> site and that the relocation of the phosphate group occurs once hydride transfer is achieved. Indeed, a hydride transfer, while the C3P group is still bound at P<sub>s</sub>, would present the advantage of holding both the substrate and cofactor in tight interaction and might promote this step of the reaction.

Besides the C3P group location, an unexpected feature concerns the orientation of the oxygen atom bound at C1, which points opposite to His-176, almost parallel to the nicotinamide ring (Fig. 3A), and adopts an energetically favorable trans conformation when related to the O2 atom position. It is well established that an efficient hydride transfer requires either a base-catalyst or an oxyanion hole to stabilize the negative charge developed on the tetrahedral transition state. Whereas non-phosphorylating GAPDHs possess the geometric features of an oxyanion hole reminiscent to those from serine proteases (39), these criteria are clearly not fulfilled in phosphorylating GAPDHs. His-176 was shown instead to play an unequivocal role of base-catalyst during the redox step (5, 7). This means that hydride transfer must have occurred with the C1-OH group well oriented with respect to His-176. In the structure presented here, which is representative of a stage after hydride transfer, the oxygen atom at C1 adopts another position that results in the disruption of the alignment between the Nε atom of His-176, the O1 and C1 atoms of the substrate, the hydride ion, and the C4 atom of the pyridinium ring.

It appears thus that the overall redox step involves important conformational adjustments of the substrate and requires the relocation of both O1 atom that shifts away from His-176 and of the C3P group that flips from the P<sub>s</sub> site toward the new P<sub>i</sub> site. Whether these two events are coupled or not is still unclear.



**FIGURE 4. Alternate conformations of the 206–212 loop and re-location of the  $P_i$  anion binding site.** A, stereoscopic view of the O subunit region spanning residues 205–218. The 206–212 loop is represented in two alternate conformations: in *dark gray*, the classic conformation delineates the  $P_i$  anion binding site where a sulfate is bound (conformation A). In *light gray*, the alternate conformation exhibits a shift toward the C3P of the intermediate (conformation B). This motion allows Thr-208 and Gly-209 to participate to the formation of the new  $P_i$  site and thus to the binding of the C3P group of the intermediate. B, superimposition of the 206–212 loop with representative GAPDH structures. The superimposition of the 206–212 loop under its classic conformation (*dark gray*) with the holo-structure from *B. stearotherophilus* (in *white*, pdb code 1gd1 (10)) is reported on the *left*, and the superimposition of the 206–212 loop under its alternate conformation with the structure of *T. maritima* GAPDH (in *white*, pdb code 1hdg (11)) is reported on the *right*. All structures are represented in *schematic mode* with their corresponding bound anion in *stick mode*. To facilitate comparison, both conformations of our model were represented in each superimposition, *shading* the one that is not considered.

However, the fact that both relocations seem to occur at the same stage, that is once hydride transfer is achieved, strongly suggests these two events to be coupled. One should note that, except for the 206–212 loop, the active site conformation remains unchanged between the Michaelis complex and the thioacylenzyme intermediate structures and that only the substrate conformation is modified along the redox step. The path-

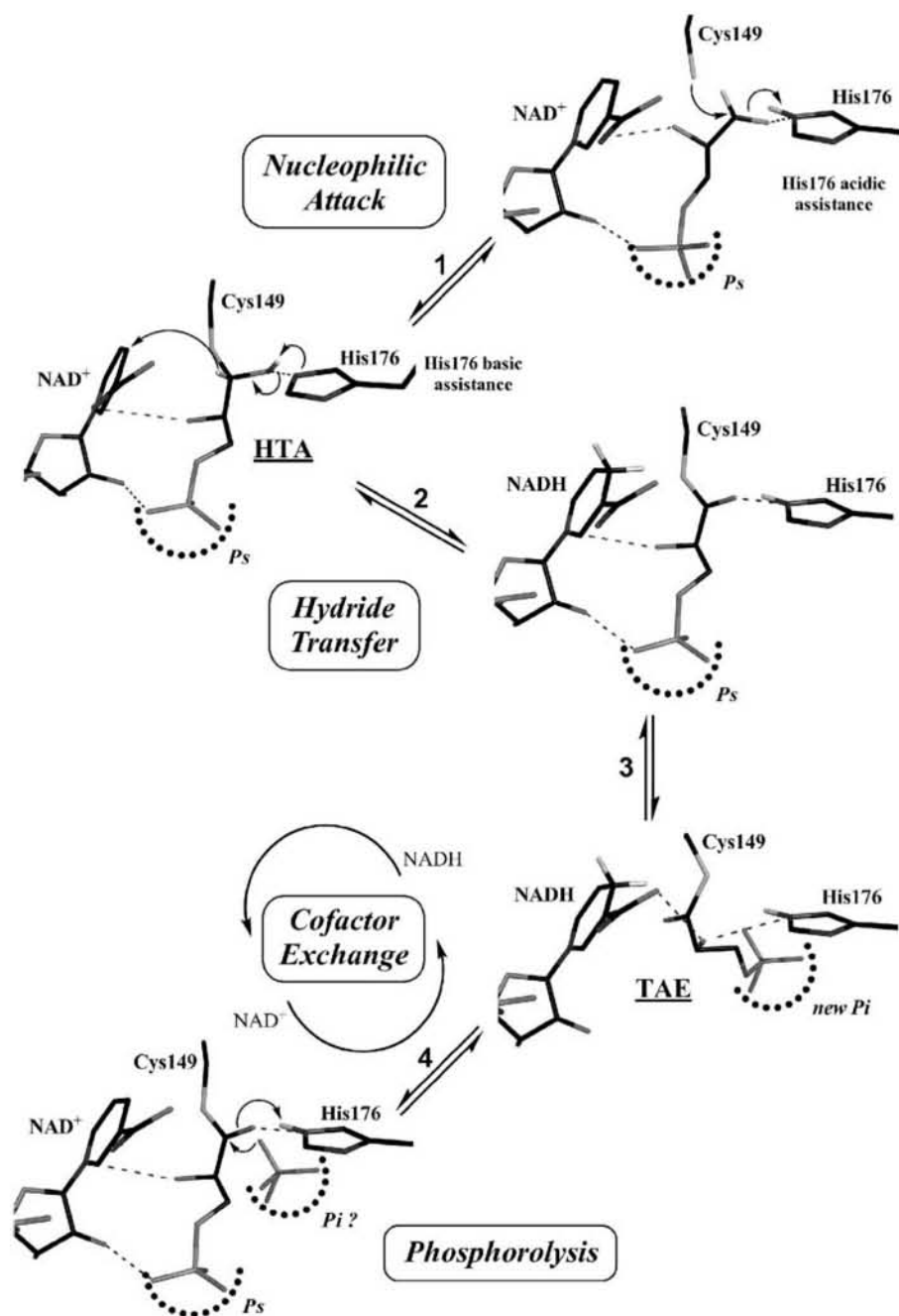
way leading from one conformation to the other implies mainly a rotation of the trans plane O1–C1–C2–O2 of  $\sim 180^\circ$ . A rotation around the C2–C3 bond is also needed for the adequate positioning of the C3P group in the new  $P_i$  site.

The next step consists of the release of the cofactor NADH, which was shown to be rate-limiting in the overall enzymatic process, and the entry of a new molecule of  $NAD^+$ , which is known to enhance the rate of phosphorolysis (2).<sup>5</sup> The events promoting NADH release are still unclear. However, the loss of the interaction between O2' of the ribose and the C3P group of the intermediate following the flipping of the C3P group from the  $P_s$  site to the new  $P_i$  site could be the first signal that triggers the dissociation of NADH. Note that, although a mixture of  $NAD^+$  and NADH is present in the structure (due to an incomplete reaction), the two states cannot be distinguished from the electron density maps and have been built as one conformation. Therefore, except for the above mentioned point, no evidence exists in the structure that might account for a destabilization of the cofactor-enzyme complex.

The entry of a new  $NAD^+$  molecule likely promotes conformational changes on the substrate molecule required for phosphorolysis. Indeed, taking into account the facts that (i) His-176 has to play a role as an acid/base catalyst (10) in the phosphorolysis step as it does in the acylation step and (ii) the inorganic phosphate has to bind to either the classic  $P_i$  (10) or the new  $P_i$  site (12), the consequence of the re-entry of  $NAD^+$  should be a flip back of both the O1 atom and the C3P groups to their initial position to interact with His-176 (Ne) and to bind to the  $P_s$  site, respectively. Again, these rearrangements can be coupled or occur in two distinct steps.

**Proposed Scenario**—G3P initially binds to the active site with its C3P located in the  $P_s$  site. Reaction begins through the nucleophilic attack of the Cys-149, whose thiol function is first

<sup>5</sup> F. Talfournier, unpublished results.



**FIGURE 5. Structural scenario proposed for the catalytic mechanism.** Only hydrogen atoms directly involved in the catalytic mechanism are represented. G3P initially binds to the active site of GAPDH with its C3P group located in the  $P_s$  site. Reaction begins through the nucleophilic attack of the Cys-149 thiolate function on the aldehydic carbon C1 of the substrate (*step 1*), which leads to a tetrahedral intermediate called hemithioacetal (*HTA*). This step is followed by hydride transfer (*step 2*), assisted by His-176, from the C1 atom of G3P to the C4 atom of  $NAD^+$ . At this stage, C3P is thought to be located in the  $P_s$  site, and the O1 atom of G3P is expected to interact with His-176 (Ne) (see "Discussion"). To allow NADH release, the substrate must undergo a conformational change during which the orientation of the O1 atom changes while the C3P group flips toward the new  $P_i$  site (*step 3*). It results in the loss of the interaction between the substrate and the cofactor required for NADH release (*TAE*: thioacylzyme intermediate corresponding to the structure depicted here).  $NAD^+$  enters the active site and likely promotes the relocation of O1 and C3P to their initial position. The  $P_i$  site (either in the classic or new position, indicated as " $P_i?$ ") is free to bind an inorganic phosphate for the nucleophilic attack on the thioacylzyme. This last step is also assisted by His-176 and leads to the formation of 1,3-bisphosphoglycerate. Subsequent release of the product leaves the enzyme in its holo-state, ready to accommodate a new G3P molecule.

activated through the formation of an ion pair with His-176, on the aldehydic carbon C1 of the substrate (Fig. 5, *step 1*). Hydride transfer occurs then from the C1 atom of the thiohemiacetal

zyme intermediate is still speculative. The structure of a Michaelis-like complex with 1,3-DPG and  $NAD^+$ , and of a thioacylzyme- $NAD^+$  complex, will provide further information

intermediate to the C4 atom of the nicotinamide of  $NAD^+$  through general base catalysis by His-176 (Fig. 5, *step 2*). At this stage, the C3P group of the thioacyl intermediate is likely still bound in the  $P_s$  site in close interaction with the cofactor (Fig. 5, *HTA stage*). Then, the thioacylzyme intermediate undergoes conformational adjustments (Fig. 5, *step 3*) during which the orientation of the O1 atom changes in such a way that it now points away from His-176 while C3P flips toward the "new"  $P_i$  site. These events lead to the thioacylzyme intermediate (Fig. 5, *TAE stage*) whose structure is presented here. The next step consists of the exchange of cofactor ( $NADH$  release, entry of a new molecule of  $NAD^+$ ) (Fig. 5, *step 4*). Note that this step is facilitated by the loss of interaction between the substrate and the cofactor molecule due to the relocation of the C3P group in the new  $P_i$  site. The entry of a new  $NAD^+$  molecule likely promotes conformational changes on the substrate molecule that are required for phosphorolysis, that is the flip back of both O1 atom and C3P group to their initial position. Inorganic phosphate finally binds to the active site and attacks the thioacylzyme intermediate from the  $P_i$  site leading to product (1,3-DPG) release. At the end, the enzyme returns to its holo-form, ready to accommodate a new G3P molecule.

In this scenario, even if the thioacylzyme structure presented here highlights the structural changes associated with the redox step, further work has to be carried out to understand how this putatively competent complex leads to 1,3-DPG formation. In that context, direct evidence is still lacking as to whether and/or how the redox state of the cofactor promotes the rearrangements required for the phosphorolysis step. In addition, the exact site from which inorganic phosphate attacks the thioacyl-

## Thioacylenzyme Intermediate from *B. stearotherophilus* GAPDH

regarding the mechanism of this “old” but still imperfectly known enzyme.

*Acknowledgments*—We thank Prof. G. L. Rossi for many stimulating discussions on GAPDH and microspectrophotometry. We gratefully acknowledge access to synchrotron radiation at the EMBL Outstation, Deutsches Elektronen-Synchrotron, Hamburg, and at the European Synchrotron Radiation Facility, Grenoble, France.

### REFERENCES

1. Segal, H. L., and Boyer, P. D. (1953) *J. Biol. Chem.* **204**, 265–281
2. Trentham, D. R. (1971) *Biochem. J.* **122**, 59–69
3. Trentham, D. R. (1971) *Biochem. J.* **122**, 71–77
4. Harris, J. I., and Waters, M. (1976) in *The Enzymes*, 3rd Ed. (Boyer, P. D., ed) ch. 13, Academic Press, New York
5. Soukri, A., Mougou, A., Corbier, C., Wonacott, A., Branlant, C., and Branlant, G. (1989) *Biochemistry* **28**, 2586–2592
6. Michels, S., Rogalska, E., and Branlant, G. (1996) *Eur. J. Biochem.* **235**, 641–647
7. Talfournier, F., Colloc'h, N., Mornon, J. P., and Branlant, G. (1998) *Eur. J. Biochem.* **252**, 447–457
8. Boschi-Muller, S., and Branlant, G. (1999) *Arch. Biochem. Biophys.* **363**, 259–266
9. Moras, D., Olsen, K. W., Sabesan, M. N., Buehner, M., Ford, G. C., and Rossmann, M. G. (1975) *J. Biol. Chem.* **250**, 9137–9162
10. Skarzyński, T., Moody, P. C., and Wonacott, A. J. (1987) *J. Mol. Biol.* **193**, 171–187
11. Korndörfer, I., Steipe, B., Huber, R., Tomschy, A., and Jaenicke, R. (1995) *J. Mol. Biol.* **246**, 511–521
12. Kim, H., Feil, I. K., Verlinde, C. L., Petra, P. H., and Hol, W. G. (1995) *Biochemistry* **34**, 14975–14986
13. Corbier, C., Michels, S., Wonacott, A. J., and Branlant, G. (1994) *Biochemistry* **33**, 3260–3265
14. Yun, M., Park, C. G., Kim, J. Y., and Park, H. W. (2000) *Biochemistry* **39**, 10702–10710
15. Castilho, M. S., Pavão, F., Oliva, G., Ladame, S., Willson, M., and Périé, J. (2003) *Biochemistry* **42**, 7143–7151
16. Didierjean, C., Corbier, C., Fatih, M., Favier, F., Boschi-Muller, S., Branlant, G., and Aubry, A. (2003) *J. Biol. Chem.* **278**, 12968–12976
17. Pearson, A. R., Mozzarelli, A., and Rossi, G. L. (2004) *Curr. Opin. Struct. Biol.* **14**, 656–662
18. Rossi, G. L., and Bernhard, S. A. (1971) *J. Mol. Biol.* **55**, 215–230
19. Mozzarelli, A., and Rossi, G. L. (1996) *Annu. Rev. Biophys. Biomol. Struct.* **25**, 343–365
20. Bourgeois, D., and Royant, A. (2005) *Curr. Opin. Struct. Biol.* **15**, 538–547
21. Berni, R., Mozzarelli, A., Pellacani, L., and Rossi, G. L. (1977) *J. Mol. Biol.* **110**, 405–415
22. Vas, M., Berni, R., Mozzarelli, A., Tegoni, M., and Rossi, G. L. (1979) *J. Biol. Chem.* **254**, 8480–8486
23. Mozzarelli, A., Berni, R., Rossi, G. L., Vas, M., Bartha, F., and Keleti, T. (1982) *J. Biol. Chem.* **257**, 6739–6744
24. Mozzarelli, A., Peracchi, A., Rossi, G. L., Ahmed, S. A., and Miles, E. W. (1989) *J. Biol. Chem.* **264**, 15774–15780
25. Kabsch, W. (1993) *J. Appl. Crystallogr.* **26**, 795–800
26. Evans, P. R. (1993) in *Proceedings of the CCP4 Study Weekend: Data Collection and Processing* (Sawyer, L., Isaacs, N., and Bailey, S., eds) pp. 114–122, Daresbury Laboratory, Warrington, UK
27. Vagin, A., and Teplyakov, A. (1997) *J. Appl. Crystallogr.* **30**, 1022–1025
28. Collaborative Computational Project, Number 4. (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50**, 760–763
29. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **53**, 240–255
30. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60**, 2126–2132
31. Bricogne, G., and Irwin, J. J. (1996) in *Macromolecular Refinement. Proceedings of the CCP4 Study Weekends*, (Dodson, E., Moore, M., Ralph, A., and Bailey, S., eds) pp. 85–92 Daresbury Laboratory Warrington, UK
32. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26**, 283–291
33. Ramakrishnan, C., and Ramachandran, G. N. (1965) *Biophys. J.* **5**, 909–933
34. Duée, E., Olivier-Deyris, L., Fanchon, E., Corbier, C., Branlant, G., and Dideberg, O. (1996) *J. Mol. Biol.* **257**, 814–838
35. Racker, E., and Krimsky, I. (1952) *J. Biol. Chem.* **198**, 731–743
36. Harrigan, P. J., and Trentham, D. R. (1974) *Biochem. J.* **143**, 353–363
37. Buehner, M., Ford, G. C., Moras, D., Olsen, K. W., and Rossmann, M. G. (1974) *J. Mol. Biol.* **82**, 563–585
38. Byers, L. D. (1978) *Arch. Biochem. Biophys.* **186**, 335–342
39. Cobessi, D., Tête-Favier, F., Marchal, S., Branlant, G., and Aubry, A. (2000) *J. Mol. Biol.* **300**, 141–152





# ANNEXE 2 : CONTRIBUTION À L'ÉTUDE DE LA PROTÉINE DING DE *PSEUDOMONAS FLUORESCENS*

## *Introduction*

Les protéines DING se caractérisent par leur séquence N-terminale DINGGG caractéristique et ont été identifiées dans une large variété d'organismes eucaryotes (Berna *et al.*, 2002 et 2007). En dépit de nombreux efforts, aucune séquence génomique complète n'est disponible pour ces protéines et une seule séquence protéique complète est connue. Cette dernière correspond à une protéine DING humaine, la "Human Phosphate-Binding Protein" (HPBP), purifiée à partir de

fractions de haute densité des lipoprotéines du plasma (Renault *et al.*, 2006 ). Sa séquence a été établie à partir de la structure cristallographique de la protéine et a été confirmée par un séquençage peptidique en spectrométrie de masse (Fokine *et al.*, 2003 ; Morales *et al.*, 2006 ; Diemer *et al.*, 2008).

La structure ternaire de HPBP a permis de confirmer la capacité de cette enzyme à fixer une molécule de phosphate inorganique, à la manière des protéines bactériennes SBP ("solute-binding proteins") qui interviennent au sein du système de transport ABC du phosphate. Les structures des SBP et de HPBP adoptent toutes un repliement connu sous le nom de "Venus fly-trap" qui est constitué de deux domaines globulaires articulés autour d'une charnière. L'anion phosphate est localisé au niveau de cette charnière dans un site de fixation formé par 8 résidus conservés qui établissent un réseau complexe de liaisons hydrogène avec la molécule de phosphate inorganique (Morales *et al.*, 2006 ; Luecke and Quioco, 1990 ; Chang and Roth, 2001).

Des séquences partielles de gènes et de protéines DING sont également disponibles pour plusieurs plantes (Berna *et al.*, 2003 et 2007), parmi lesquelles celle isolée à partir du millepertuis (*Hypericum perforatum*) sur la base d'une activité anti-HIV sur des lymphocytes. Bien que tronquée en C-terminale vraisemblablement à la suite d'une hydrolyse lors de l'extraction (Darbinian-Sarkissian *et al.*, 2006), cette protéine présente une analogie évidente avec les séquences de protéines DING existantes. Si les résidus impliqués dans la fixation du phosphate y sont conservés, aucune preuve expérimentale n'est disponible sur les propriétés de fixation du phosphate par cette protéine.

Les protéines SBP sont retrouvées de manière ubiquitaire dans le génome des bactéries, cependant, certains génomes de l'espèce *Pseudomonas* présentent, en plus, des gènes homologues à ceux codant pour les protéines DING eucaryotiques. Ces gènes codent pour deux familles de protéines fixant le phosphate et qui pourraient être impliquées dans son piégeage (Berna *et al.*, 2007). L'un de ces gènes, issu de *Pseudomonas fluorescens*, a été cloné et surexprimé par l'équipe du Professeur K. Scott de l'université d'Auckland en Nouvelle Zélande (Scott and Wu, 2005). La protéine obtenue, la *PfluDING*, fixe une molécule de phosphate et présente, au même titre que d'autres protéines DING humaines (Hain *et al.*, 1996 ; Bush *et al.*, 1998 ; Adams *et al.*, 2003 ; Belenky *et al.*, 2003), une activité mitogène envers les fibroblastes humains en culture. La séquence de cette protéine *PfluDING* présente une forte homologie de séquence (plus de 60 %) avec la séquence partielle de la protéine DING du millepertuis.

En l'absence de gène eucaryote complet codant pour ces protéines DING, la *PfluDING*

constitue un modèle de choix pour l'étude des relations structure-fonction de cette famille de protéines. C'est dans le cadre d'une collaboration entre le LCM3B (Dr. E. Chabrière) et l'équipe du Professeur K. Scott que j'ai été amené à travailler sur cette protéine. L'objectif initial de cette étude était de déterminer la structure cristallographique de la *PfluDING* pour étudier les relations structure-fonction de cette enzyme.

Le haut pouvoir de diffraction des cristaux de *PfluDING* ( limite de résolution inférieure à 1 Å) mis en évidence au début de cette étude nous a alors conduits à étudier précisément le réseau de liaisons hydrogène à l'origine de la fixation du phosphate. En effet, cette protéine DING, comme les PBP ("phosphate binding protein") bactériennes, présente une grande spécificité vis-à-vis de la reconnaissance du phosphate inorganique ( $K_D = 1,7 \mu\text{M}$ ). Cette spécificité repose sur l'existence, au niveau du site de fixation des PBP, d'un résidu d'acide aminé aspartate, accepteur exclusif de liaisons hydrogène (Luecke and Quioco, 1990). Ainsi, l'aspartate, permettrait de fixer et de reconnaître spécifiquement l'atome d'oxygène protoné du phosphate (le phosphate possédant toujours au moins un proton à des pH physiologiques) et constituerait grâce à sa charge négative un antidéterminant de la fixation d'anion sulfate, complètement déprotoné.

Le site de fixation du phosphate dans la *PfluDING* est très conservé et renferme le résidu aspartate (Asp62). Compte tenu du haut pouvoir de diffraction de ses cristaux, cette protéine constitue un modèle potentiellement très informatif pour étudier la spécificité de ce type d'enzyme pour l'anion phosphate. De plus, nous avons mis en évidence la possibilité d'incuber les cristaux à différents pH sans altération de leur pouvoir de diffraction. Cette opportunité nous a alors conduits à comparer la fixation des formes monobasique ( $\text{H}_2\text{PO}_4^-$  à pH 4,5) et dibasique ( $\text{HPO}_4^{2-}$  à pH 8,5) du phosphate par la *PfluDING*.

## ***Matériel et méthodes***

### ***Production de la protéine***

La protéine *pfluDING* recombinante a été produite et purifiée par l'équipe du Professeur K. Scott. La protéine a été produite dans des cellules d'*E. coli* BL21 (DE3) transformées par un plasmide pET22b(+) contenant le gène de la protéine fusionné à une étiquette hexahistidine en C-terminale puis purifiée par chromatographie d'affinité. Les procédures expérimentales exactes ont été publiées par Scott and Wu (2005).

## ***Cristallisation***

Préalablement à la cristallisation, la *PfluDING* a été dialysée pendant 24 h contre un tampon contenant 20 mM de Tris à pH 8,5 et 5 mM de  $\beta$ -mercaptoéthanol puis concentrée par centrifugation en utilisant un système Centricon à seuil de coupure de 3,5 kDa. La concentration de la protéine, déterminée par spectrophotométrie en utilisant un coefficient d'extinction molaire de  $42650 \text{ M}^{-1} \cdot \text{cm}^{-1}$ , a ensuite été ajustée à  $10 \text{ mg} \cdot \text{mL}^{-1}$  pour les essais de cristallisation.

Le criblage initial des conditions de cristallisation de la *PfluDING* a été réalisé à  $20^\circ\text{C}$  sur les matrices d'échantillonnages Crystal Screen I et II (Hampton Research) et Wizard I et II (Emerald Biostructures) en utilisant la méthode de la diffusion de vapeur en goutte assise. La condition n°17 du Wizard I [30% (p/v) PEG 8000, 100 mM acétate pH 4,5 et 200 mM  $\text{Li}_2\text{SO}_4$ ] ayant conduit à l'obtention de polycristaux, elle a ensuite été optimisée en utilisant la méthode de la diffusion de vapeur en goutte suspendue. Des polycristaux croissent en quelques jours en présence de fortes concentrations de PEG (28-32%). Des monocristaux exploitables d'une taille moyenne de  $0,1 \times 0,1 \times 0,05$  mm apparaissent en une semaine dans des gouttes composées de  $2 \mu\text{l}$  de solution protéique et  $2 \mu\text{L}$  de la solution du puits (24-28% (p/v) PEG 8000, 100 mM acétate pH 4,5 et 200 mM  $\text{Li}_2\text{SO}_4$ ). La *PfluDING* garde cependant une forte propension à former des polycristaux et des techniques de micro-ensemencements ont été mises en œuvre pour améliorer la qualité des cristaux. Des cristaux d'une taille supérieure à  $0,2 \times 0,2 \times 0,1$  mm sont à ce jour obtenus de manière reproductible.

Une fois les cristaux obtenus, la solution composant la goutte de cristallisation peut être progressivement échangée contre une solution similaire, dans laquelle le tampon acétate a été remplacé par un tampon Tris à pH 8,5. Les cristaux incubés dans cette nouvelle condition sont stables pendant 2 à 3 jours sans dégradation visuelle ni altération de leur pouvoir de diffraction.

## ***Collecte des données de diffraction***

Préalablement à leur congélation dans l'azote liquide, les cristaux sont rapidement plongés dans une solution cryoprotectrice dérivée de la condition de cristallisation (à 28 % de PEG 8000) dans laquelle 10% (p/v) de glycérol ont été ajoutés. Le premier jeu de données de diffraction, composé de 200 images représentant chacune  $1^\circ$  d'oscillation, a été collecté sur la ligne de lumière BM14 de l'ESRF en utilisant une longueur d'onde de  $0,9535 \text{ \AA}$  et un détecteur MAR CCD (Mosaic 225). Les données présentant une résolution limite de  $1,43 \text{ \AA}$  de résolution ont été traitées en utilisant le programme XDS (Kabsch, 1993) (Tableau A-1).

**Tableau A-1 : Statistiques de collecte et d'affinement des structures cristallographiques de la PfluDING.** Le modèle DING1 a été affiné en utilisant le programme Refmac5 et les modèles DING-4.5 et DING-8.5 avec le programme Shelxl.

Données de diffraction	DING 1 (pdb 2q9t)	DING-4.5	DING-8.5
Groupe d'espace	$P 2_1$	$P 2_1$	$P 2_1$
Maille (Å)	$a = 36,7$ $b = 123,7$ $c = 40,8$ $\beta = 116,7^\circ$	$a = 36,9$ $b = 124,0$ $c = 40,9$ $\beta = 116,3^\circ$	$a = 36,9$ $b = 124,0$ $c = 40,9$ $\beta = 116,3^\circ$
Limite de résolution (Å)	61,9 - 1,43	62,1 - 2,14	50,0 - 0,88
Couronne de plus haute résolution (Å)	1,50 - 1,43	1,08 - 0,98	0,94 - 0,88
Température (K)	100	100	100
Réflexions uniques	58905(7166)	185558 (44846)	253101 (42675)
Complétude (%)	98,1 (89,2)	98,1 (94,0)	97,4 (91,8)
Redondance	4,0 (2,9)	5,4 (3,1)	6,1 (3,1)
R-merge <sup>a</sup> (%)	2,6 (10,2)	4,3 (7,5)	5,3 (43,8)
I/ $\sigma$ (I)	38,9 (10,5)	24,4 (10,4)	16,6 (2,6)
<b>Affinement</b>			
R <sup>b</sup> (%)	13,1	8,5	12,9
R <sub>free</sub> <sup>c</sup> (%)	16,4	10,8	14,6
<b>r.m.s.d.<sup>d</sup> par rapport à la géométrie standard</b>			
Longueur de liaisons (Å)	0,007	0,006	0,006
Angles de valence (°)	1,12	1,19	1,18
<b>Facteur B moyens (Å<sup>2</sup>)</b>			
Modèle	10,15	9,6	9,3
Atomes de la protéine	6,6	7,7	7,9
Anion phosphate	3,1	3,1	3,2
Molécules d'eau	22,6	17,1	14,5

<sup>a</sup> Facteur R pour les intensités reliées par symétrie.

<sup>b</sup> Facteur R cristallographique. Facteur R cristallographique calculé sur 5 % des réflexions choisies aléatoirement et non incluses dans l'affinement.

<sup>d</sup> Écart quadratique moyen.

Parmi les jeux de données de diffraction collectés ensuite, 2 jeux de données présentant des limites de résolution inférieures à 1 Å ont pu être collectés sur des cristaux natifs (pH 4,5) ou des cristaux incubés à pH 8,5 pendant au moins 20 h. Ces jeux de données ont été collectés sur les lignes de lumière ID14eh4 ( $\lambda = 0,94$  Å ; Quantum ADSC Q315) et ID29 ( $\lambda = 0,75$  Å ; Quantum ADSC Q315) de l'ESRF en réalisant dans chaque cas 3 collectes successives sur un même cristal pour obtenir tour à tour les données de haute, de moyenne et de basse résolution. Les données de diffraction ont été traitées en utilisant le programme XDS (Kabsch, 1993) (Tableau A-1).

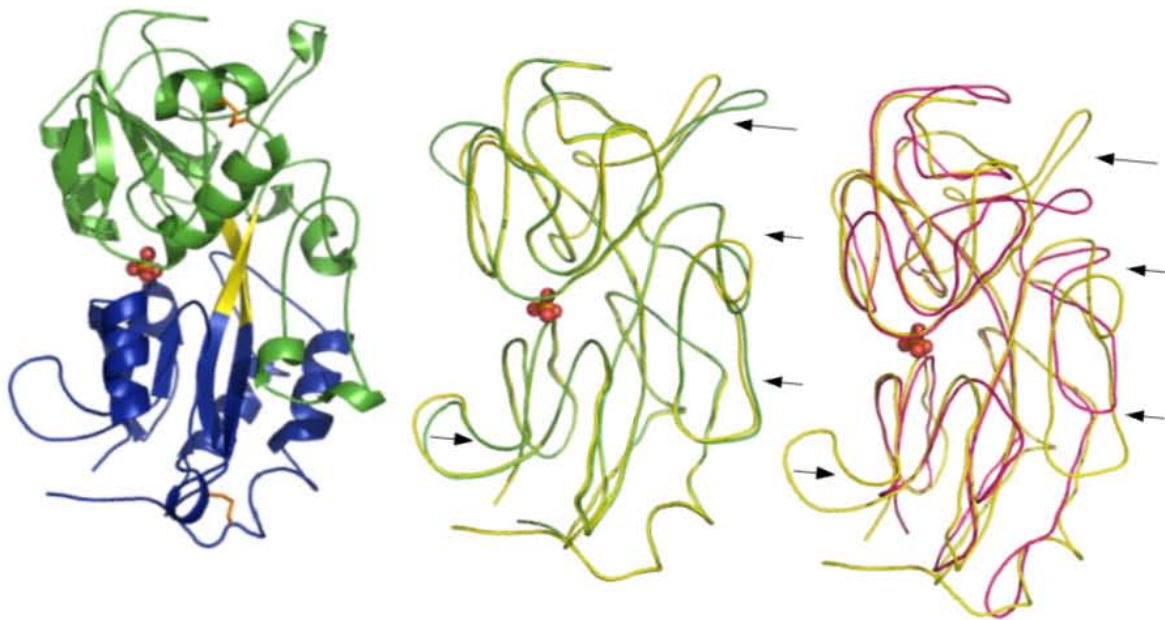
### ***Résolution de la structure et affinements cristallographiques***

Les phases initiales ont été obtenues par remplacement moléculaire en utilisant le logiciel MOLREP (Vagin and Teplyakov, 1997) et les coordonnées atomiques de la structure de HPBP (code pdb 2cap, 71 % d'identité de séquence avec la *PfluDING*) en tant que modèle. Pour les trois structures, l'affinement a été réalisé avec les programmes Refmac5 (Murshudov *et al.*, 1997) et Coot (Emsley and Cowtan, 2004) jusqu'à convergence des facteurs R et  $R_{\text{free}}$ . Le modèle DING1, à 1,43 Å de résolution a convergé à des valeurs de 13,1 et 16,4 % respectivement pour les R et  $R_{\text{free}}$  et a été ensuite analysé. Ce modèle a été déposé dans la Protein Databank sous le code 2q9t et publié (Ahn *et al.*, 2007). Pour les deux autres jeux de données, l'affinement a été poursuivi avec le programme SHELXL (Sheldrick & Schneider, 1997) jusqu'à des valeurs de R et  $R_{\text{free}}$  de respectivement 8,5 et 10,8 % dans le cas du modèle DING-4.5 (cristal natif obtenu à un pH de 4,5) et de 12,9 et 14,6 % dans le cas du modèle DING-8.5 (cristal incubé à pH 8,5). Les travaux d'affinement des modèles à ultra-haute résolution sont encore en cours, menés par Dorothée Liebschner au LCM3B en utilisant le programme Mopro (Guillot *et al.*, 2001).

## ***Résultats***

### ***Structure ternaire de la PfluDING***

La structure cristallographique de la *PfluDING* révèle un repliement allongé composé de deux domaines globulaires adjacents. Chacun des domaines est constitué d'un feuillet  $\beta$  central flanqué par des hélices  $\alpha$  et contient un pont disulfure (Cys114-Cys159 et Cys300-Cys363) (Figure A-1). Connectés au travers d'un long feuillet antiparallèle à deux brins se comportant comme une charnière, les deux domaines forment, à leur interface, une cavité profonde dans laquelle est lié un



**Figure A-1 : Superposition des structures secondaires.** (*gauche*) La structure de la protéine DING de *Pseudomonas fluorescens* est représentée en mode ruban et flèche. Les deux domaines sont colorés en bleu et vert et la charnière en jaune. Les 2 ponts disulfures sont représentés en orange. La molécule de phosphate inorganique est représentée en mode sphère et bâton (en rouge). (*milieu*) Comparaison de la structure de la PfluDING (en jaune) avec celle de la DING humaine (HPBP) (2cap) (en vert). (*droite*) Comparaison de la structure de la PfluDING (en jaune) avec la structure de la PstS d'*E. coli* (1ixh) (en rose). L'anion phosphate fixé dans la structure de la PfluDING est représenté en mode sphère et bâton (en rouge). Les boucles protubérantes existant dans les structures de la PfluDING et de HPBP sont indiquées par des flèches.

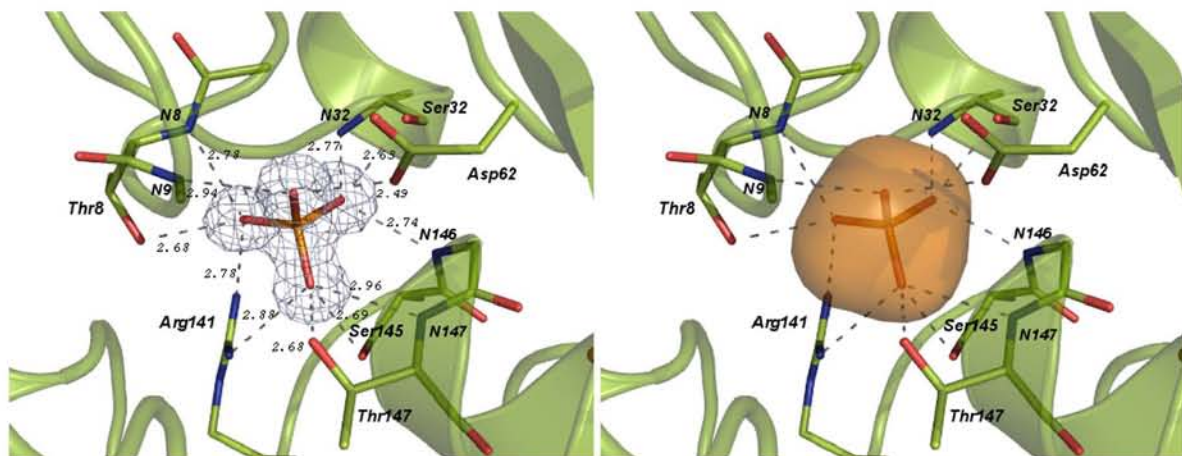
anion phosphate (Figure A-1). Ce repliement, connu sous le nom de "Venus fly-trap" ressemble étroitement au repliement adopté par la DING humaine (HPBP) et par les protéines bactériennes PstS. Ainsi, la superposition des carbones de C $\alpha$  du modèle DING1 avec les structures de HPBP (2cap; Morales *et al.*, 2006) et de la PstS d'*E. coli* (1ixh, Wang *et al.*, 1997) conduit à des rmsd de 0,65 et 1,88 Å, respectivement. Comme la protéine HPBP, la structure de la PfluDING présente 4 boucles protubérantes, qui n'existent pas dans la structure de l'enzyme d'*E. coli* (Figure A-1). Bien que ces boucles soient légèrement plus courtes dans le cas de la PfluDING, elles se superposent presque parfaitement avec les boucles de HPBP (Figure A-1). Ces observations, au même titre que la forte identité de séquence existant entre les deux protéines, confirment que la protéine bactérienne PfluDING est plus proche des protéines DING eucaryotiques que des PstS bactériennes (Scott and Wu, 2005).

### **Le site de fixation du phosphate**

Le site de reconnaissance anionique de la PfluDING, localisé à l'interface entre les deux domaines, est complètement enfoui et contient un anion phosphate qui a été co-purifié avec l'enzyme. La molécule de phosphate inorganique est stabilisée par 12 liaisons hydrogène formées par 8 résidus d'acides aminés répartis autour du site de fixation (Thr8, Leu9, Ser32, Asp62, Arg141, Ser145, Gly146 et Thr147) (Figure A-2). La longueur des liaisons hydrogène (entre atomes lourds) varie de 2,49 à 3,16 Å. La liaison la plus courte implique un atome d'oxygène du phosphate et un atome d'oxygène de la chaîne latérale de l'Asp62. Des liaisons hydrogènes particulièrement courtes (jusqu'à 2,43 Å) ont également été rapportées dans les structures des PstS d'*E. coli* (1ixh), de *M. tuberculosis* (1pc3) et dans HPBP (2cap). Étant le seul accepteur exclusif de liaison hydrogène, le résidu Asp62 joue vraisemblablement un rôle clé dans la spécificité de reconnaissance en acceptant les espèces protonées du phosphate (Luecke and Quiosho, 1990). Les résidus impliqués dans la fixation du phosphate sont strictement identiques à ceux identifiés dans la structure de HPBP (Morales *et al.*, 2006).

La structure de la PfluDING révèle que le phosphate occupe presque entièrement la cavité du site de fixation (Figure A-2). Pour illustration, la molécule d'eau la plus proche de l'anion phosphate est localisée à une distance de 6 Å. Compte tenu de la possibilité pour cette protéine de fixer des ligands plus encombrants que le phosphate tels que les nucléotides triphosphate (ATP, TTP, CTP, GTP) (K. Scott, communication personnelle), la fixation de ces molécules implique donc nécessairement des changements conformationnels de la structure, qui s'apparentent très vraisemblablement au mouvement d'ouverture du repliement en "Venus fly-trap" qui a déjà été





**Figure A-2 : Vues du site de fixation du phosphate de la PfluDING.** Les résidus impliqués dans la fixation du phosphate sont représentés sous forme de bâtons et annotés. **(gauche)** Les 12 liaisons hydrogène formées entre l'anion phosphate et la protéine sont représentées par des pointillés et les distances mesurées entre les atomes lourds sont reportées en Å. La carte omit mFo-DFc, calculée sans le phosphate, est contournée à 5 sigma. **(droite)** La cavité du site actif : la surface de van der Waals de la protéine a été tracée en utilisant une sphère de 1,6 Å de rayon.

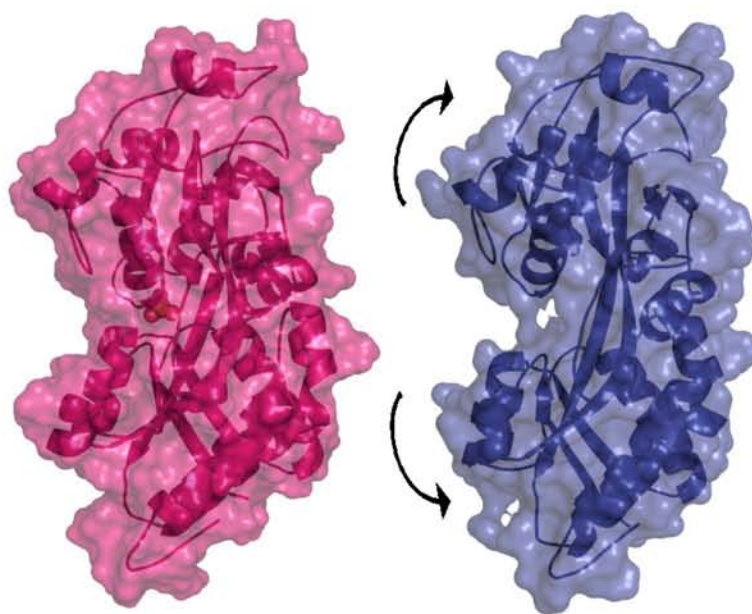
décrit notamment pour la protéine PstS d'*E. coli* (Yao *et al.*, 1996) (Figure A-3).

### ***Étude des structures à ultra-haute résolution***

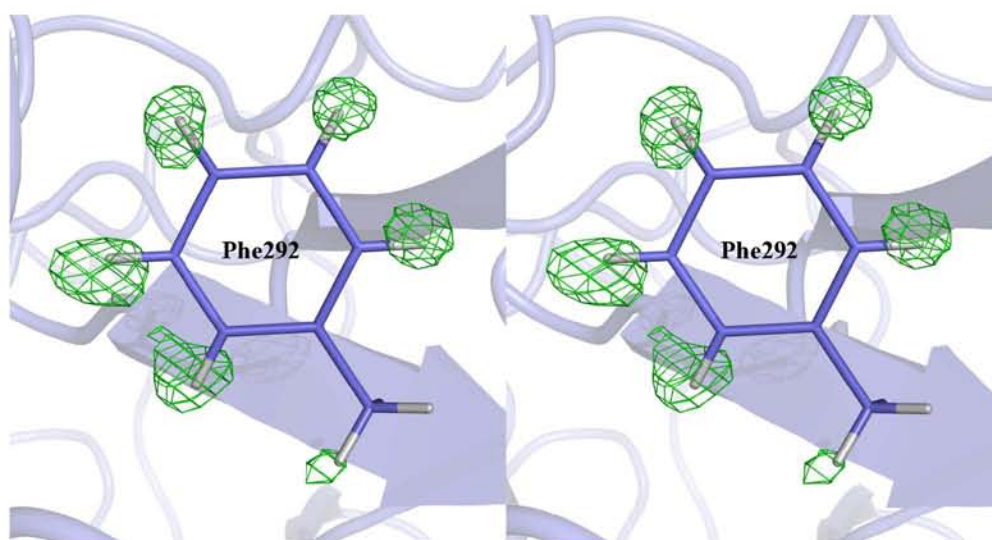
Il est possible, avec des données de diffraction d'une résolution inférieure à l'angström, de visualiser la position de certains atomes d'hydrogène de la structure d'une protéine (Figure A-4). Cependant, une telle résolution n'est que rarement atteinte (moins de 1 % des structures déposées dans la PDB). Les cristaux de la *PfluDING* possèdent un haut pouvoir de diffraction, dont les limites sont encore à déterminer. La résolution maximale atteinte à ce jour est de 0,88 Å, mais l'amélioration de la qualité des cristaux devrait encore permettre d'augmenter cette limite. En effet, généralement corrélé au pouvoir de diffraction, le facteur d'agitation thermique de ces jeux de données a été estimé par le "Wilson plot" à environ 4 Å<sup>2</sup> et montre que les cristaux de la *PfluDING* sont particulièrement ordonnés. Ce phénomène peut être en partie expliqué par le faible taux de solvant contenu dans les cristaux, qui est inférieur à 40% (Moniot *et al.*, 2007). En l'attente de jeux de données de meilleure résolution et qualité, des analyses préliminaires des modèles à 0,98 et 0,88 Å de résolution ont été conduites. Ces deux modèles ont été obtenus respectivement à partir d'un cristal natif (obtenu à pH 4,5) et d'un cristal incubé plus de 20 h à un pH de 8,5. Nous avons, en effet, montré qu'il était possible, une fois les cristaux obtenus, de modifier le pH de leur environnement (de 4,5 à 8,5) sans effet notable sur leur pouvoir de diffraction. Cette possibilité, nous autorise aujourd'hui à envisager de pouvoir comparer l'état de protonation de la protéine en fonction du pH et surtout de comparer le mode fixation des formes monobasique (H<sub>2</sub>PO<sub>4</sub><sup>-</sup> à pH 4,5) et dibasique (HPO<sub>4</sub><sup>2-</sup> à pH 8,5) du phosphate dans le site de fixation de la *PfluDING*.

Les deux modèles, d'abord affinés avec Refmac5, ont ensuite été affinés de façon anisotrope par le programme SHELXL jusqu'à la convergence des facteurs R et R<sub>free</sub> (Tableau A-1). Les deux modèles présentent une qualité géométrique satisfaisante et ne renferment aucun résidu à l'extérieur des zones autorisées du diagramme de Ramachandran. Dans les deux cas, une molécule de phosphate inorganique est présente dans le site de reconnaissance anionique, stabilisée par des interactions identiques à celles précédemment décrites. Le calcul de cartes omit des atomes d'hydrogène du site de fixation a ensuite été réalisé sur chacune des deux structures dans le but d'établir précisément le réseau de liaisons hydrogène impliqué dans la fixation du phosphate inorganique à pH de 4,5 et 8,5 (c'est-à-dire d'identifier les donneurs et les accepteurs de liaisons H) puis de les comparer (Figure A-5).

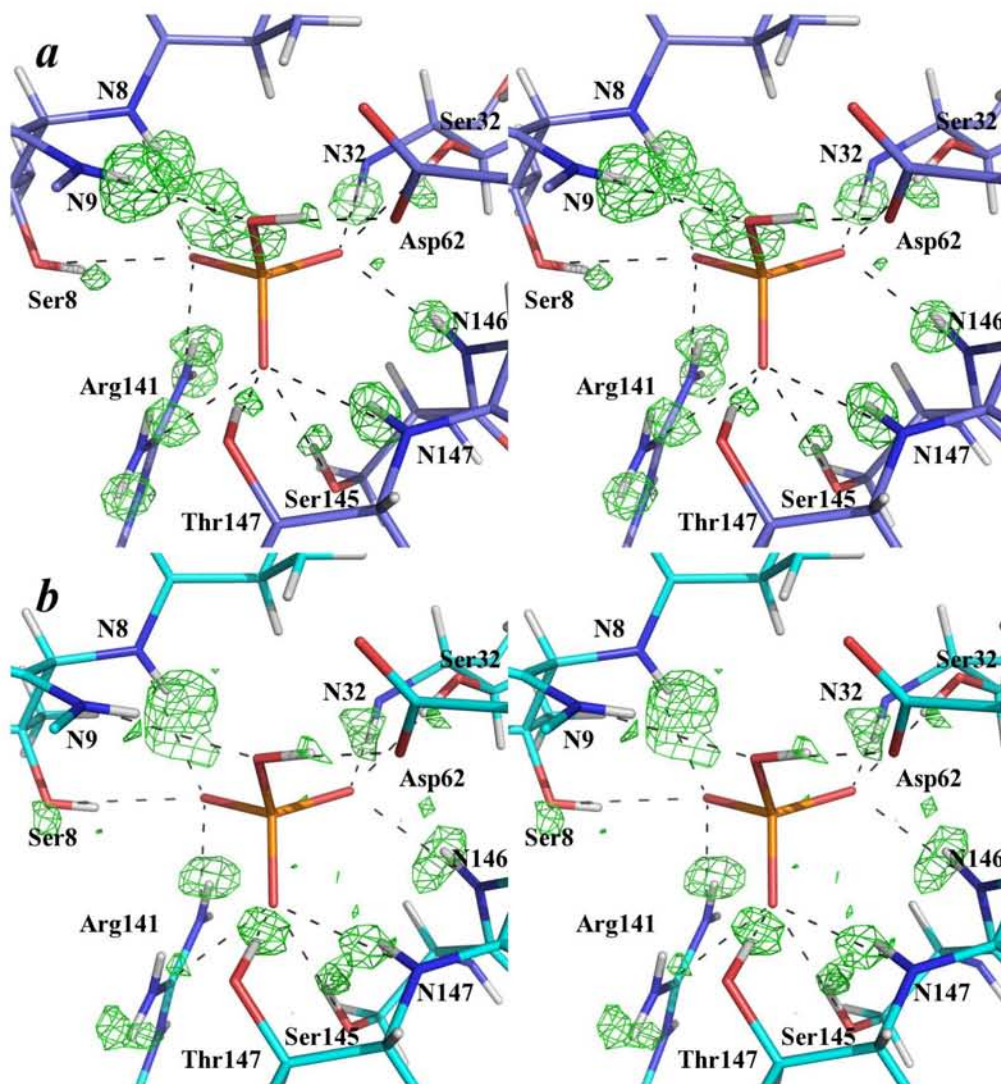
Les cartes de densité électronique omit Fo-Fc, contourées à 2,8 sigma, permettent de



**Figure A-3 : Le mouvement du repliement en "Venus fly-trap".** Le mouvement du repliement en "Venus fly-trap" autour de la charnière centrale est présenté au travers de la comparaison des formes "fermée" (enzyme de type sauvage, en mauve) et "ouverte" (mutant T141D, en bleu) de la PstS d'E. coli (code pdb 1oib et 1ixh). Les structures sont présentées en mode ruban et flèche et par leur surface moléculaire. La molécule de phosphate inorganique liée à la forme "fermée" est représentée en mode sphère et bâton en rouge.



**Figure A-4 : Carte de densité électronique omit différence des atomes d'hydrogène du résidu Phe292 du modèle DING-4.5.** La carte de densité électronique de type Fo-Fc et contourée à 3 sigma a été calculée avec Shelxl en ayant omis les atomes d'hydrogène du résidu Phe292.



**Figure A5 : Cartes de densité électronique omit des sites actifs des modèles DING-4.5 (a) et DING-8.5 (b).**  
 Les cartes de type Fo-Fc et contourées à 2,8 sigma ont été calculées avec Shelxl après avoir omis de l'affinement les atomes d'hydrogène du site de fixation et du phosphate. Les résidus impliqués dans la fixation du phosphate inorganique sont représentés sous forme de bâtons et annotés avec le code 3 lettres des acides aminés. Les atomes d'azote appartenant à la chaîne principale sont annotés Nx, avec x le numéro du résidu.

visualiser l'ensemble des atomes d'hydrogène du site actif des formes DING-4.5 et DING-8.5. Bien qu'ils présentent généralement un signal plus faible que les atomes d'hydrogène en position fixe (NH de chaîne principale, par exemple), les atomes d'hydrogène des fonctions hydroxyle des sérines et thréonines du site de fixation sont également observés de manière non ambiguë dans les cartes omit (Figure A-5). Dans les deux formes, DING-4.5 et DING-8.5, un signal de densité positive localisé à proximité de l'atome d'oxygène O1 de la molécule de phosphate inorganique a été attribué à la présence d'un atome d'hydrogène qui forme une liaison hydrogène avec le résidu Asp62 (Figure A-5). La carte omit différence du site de fixation de la forme DING-4.5 montre un signal de densité positive à proximité de cet atome O1 qui n'a pu être expliqué. Néanmoins, aucun signe de la présence d'un second atome d'hydrogène n'est observé à proximité d'un autre atome d'oxygène du phosphate qui semble donc, dans les deux cas, correspondre à la forme dibasique,  $\text{HPO}_4^{2-}$ , monoprotinée du phosphate inorganique. Si cette observation semble en accord avec la forme adoptée en solution par le phosphate à un pH de 8,5 et donc pour la forme DING-8.5, on aurait pu s'attendre, pour la forme DING-4.5 obtenue à un pH de 4,5, à observer la forme monobasique du phosphate ( $\text{H}_2\text{PO}_4^-$ ) portant deux atomes d'hydrogène. Ainsi il semblerait que la *PfluDING* fixe préférentiellement la forme dibasique du phosphate peut-être en modifiant localement le pKa pour favoriser l'existence de cette forme monoprotinée. On ne peut toutefois pas exclure que le phosphate, co-purifié avec la protéine, comme les résidus du site actif, complètement enfouis et inaccessibles au solvant, puissent avoir conservé, lors de la cristallisation à pH 4,5, un état de protonation correspondant au pH du milieu de purification (pH 8,5 ; Scott and Wu, 2005). Il est important de noter que le phosphate inorganique n'est vraisemblablement pas le ligand naturel de la *PfluDING* puisque nombre de molécules contenant du phosphate, tels que les nucleotides di- et triphosphate, le glucose-6-phosphate ou encore des molécules comme la carboxycytidine ont un effet inhibiteur notable sur la fixation du phosphate inorganique (Scott and Wu, 2005 ; K. Scott communication personnelle).

## ***Conclusion et perspectives***

Ces travaux ont permis tout d'abord d'obtenir la première structure cristallographique d'une protéine DING bactérienne qui, avec à la structure de la protéine DING humaine, HPBP, complète les connaissances structurales disponibles pour cette famille de protéines. Bien qu'elles soient structurellement très proches des protéines solubles du système de transport ABC du phosphate chez les bactéries (PstS, PBP), la structure des protéines DING présentent quelques différences notables

comme l'existence de 4 boucles protubérantes. Les études biochimiques ont par ailleurs mis en évidence que la fixation de phosphate par la protéine affectait son activité mitogénique. Ainsi une forme "ouverte" de la protéine (mutant T147N dont la fixation du phosphate est fortement altérée) ou encore tronquée en N- et C-terminal (correspondant à un seul des deux domaines) présente une activité accrue par rapport à la protéine native (Ahn *et al.*, 2007). Ces observations semblent indiquer que l'activité mitogénique est révélée lors de l'ouverture du repliement en "Venus fly-trap" qui pourrait ainsi rendre accessible un site de signalisation de l'activité mitogénique. La disponibilité de ces deux formes de l'enzyme pour des études cristallographiques devrait permettre d'apporter des éléments de réponse quant à l'identification des déterminants structuraux de l'activité mitogénique de la *PfluDING*.

Au delà, de l'étude des relations structure-fonction, cette protéine constitue également un modèle de choix de l'étude du mode de fixation du phosphate par ces protéines. Les analyses préliminaires des structures à ultra-haute résolution semblent indiquer que seule la forme dibasique du phosphate est reconnue par le site de fixation de la *PfluDING*. Ces résultats devront néanmoins être confirmés en s'assurant notamment que la protéine ait pu échanger l'anion phosphate avec lequel elle est co-purifiée, lors d'une dialyse en tampon phosphate à pH 4,5, par exemple. L'obtention de jeu de données de diffraction de meilleure qualité pour les deux formes sera également déterminante. En effet, bien qu'elles soient prometteuses, les cartes de densité électronique différence présentent dans les deux cas des incertitudes qui fragilisent les interprétations. Ainsi, l'existence d'un pic de densité positive important (près de 4 sigma) et non attribué à proximité du phosphate dans le cas du modèle DING-4.5 ou encore le niveau de bruit assez élevé dans les cartes de différence du modèle DING-8.5 sont autant de limites à l'exploitation complète de ces modèles.

Enfin, au vu des résolutions d'ores et déjà atteintes et des facteurs d'agitation thermique particulièrement faibles de ces structures, il semble possible d'améliorer encore les limites de résolution de ces cristaux. Des données de diffraction avec une résolution inférieure à 0,8 Å, particulièrement rares, permettent d'établir précisément la densité de charge et donnent accès aux propriétés de charge atomique ou de potentiel électrostatique (localisation des doublets non liants, micro/macro dipôle). De telles données sont d'un intérêt particulier pour le LCM3B qui développe le logiciel d'affinement cristallographique MoPro spécialement orienté vers l'affinement des structures cristallographiques de résolution subatomique en utilisant un modèle atomique multipolaire (Guillot *et al.*, 2001).

## Références Bibliographiques

- Adams, L., Davey, S. and Scott, K. (2002) *Biochim. Biophys. Acta* 1586, 254–264.
- Ahn, S., Moniot, S., Elias, M., Chabriere, E., Kim, D., Scott, K. (2007) *FEBS Lett.* 581(18):3455-60.
- Belenky, M., Prasain, J., Kim, H. and Barnes, S. (2003) *J. Nutr.* 133, 2497S–2501S.
- Berna, A., Bernier, F., Chabriere, E., Perera, T. and Scott, K. (2008) *Int J Biochem Cell Biol.* 40(2):170-5
- Berna, A., Bernier, F., Scott, K. and Stuhlmuller, B. (2002) *FEBS Lett.* 524, 6–10.
- Bush, D., Fritz, H., Knight, C., Mount, J. and Scott, K. (1998) *Biol. Chem.* 379, 225–229.
- Chang, G. and Roth, C.B. (2001) *Science* 293, 1793–1800.
- Darbinian-Sarkissian, N., Darbinyan, A., Otte, J., Radhakrishnan, S., Sawaya, B.E., Arzumanyan, A., Chipitsyna, G., Popov, *et al.* (2003) *Gene Therapy* 13, 288–295.
- Diemer, H., Elias, M., Renault, F., Contreras-Martel, C., Schaeffer, C., Van Dorsselaer, A. and Chabriere, E. (2008) *Proteins.* 71(4):1708-20.
- Fokine, A., Morales, R., Contreras-Martel, C., Carpentier, P., Renault, F., Rochu, D. and Chabrière, E. (2003) *Acta Crystallogr. Sect. D-Biol. Crystallogr.* 59, 2083–2087.
- Guillot, B., Viry, L., Guillot, R., Lecomte, C. & Jelsch, C. *J. Applied Crystallography* (2001). 34, 214-223.
- Hain, N.A.K., Stuhlmuller, B., Hahn, G.R., Kalden, J.R., Deutzmann, R. and Burmester, G.R. (1996) *J. Immunol.* 157, 1773–1780.
- Luecke, H. and Quioco, F.A. (1990) *Nature* 347, 402–406.
- Moniot, S., Elias, M., Kim, D., Scott, K. and Chabriere, E. (2007) *Acta Cryst. F* 63, 1–3.
- Morales, R., Berna, A., Carpentier, P., Contreras-Martel, C., Renault, F., Nicodeme, M., Chesne-Seck, M.-L., Bernier, F., Dupuy, J., Schaeffer, *et al.* (2006) *Structure* 14, 601–609.
- Renault, F., Chabriere, E., Andrieu, J.P., Dublet, B., Masson, P. and Rochu, D. (2006) *J. Chromatogr. B* 836, 15–21.
- Sheldrick, G., M., and Schneider, T., R. (1997) *Methods in enzymology* vol. 277, 319-343.
- Scott, K. and Wu, L. (2005) *Biochim. Biophys. Acta* 1744, 234–244.
- Wang, Z., Choudhary, A., Ledvina, P.S. and Quioco, F.A. (1994) *J. Biol. Chem.* 269, 25091–25094.
- Yao, N., Ledvina, P.S., Choudhary, A. and Quioco, F.A. (1996) *Biochemistry* 32, 2079–2085.





Sebastien Moniot,<sup>a</sup> Mikael Elias,<sup>a</sup>  
Donghyo Kim,<sup>b</sup> Ken Scott<sup>b</sup> and  
Eric Chabriere<sup>a\*</sup>

<sup>a</sup>Laboratoire de Cristallographie et Modélisation des Matériaux Minéraux et Biologiques, CNRS–Université Henri Poincaré, 54506 Vandoeuvre-lès-Nancy, France, and <sup>b</sup>School of Biological Sciences, University of Auckland, Auckland, New Zealand

Correspondence e-mail:  
eric.chabriere@lcm3b.uhp-nancy.fr

Received 15 March 2007

Accepted 7 June 2007

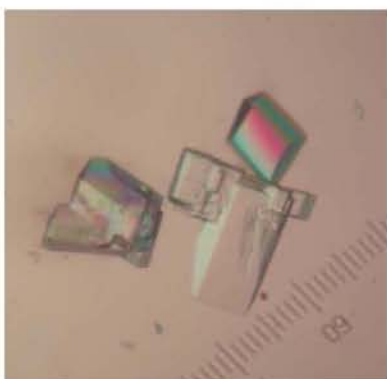
## Crystallization, diffraction data collection and preliminary crystallographic analysis of DING protein from *Pseudomonas fluorescens*

PfluDING is a phosphate-binding protein expressed in *Pseudomonas fluorescens*. This protein is clearly distinct from the bacterial ABC transporter soluble phosphate-binding protein PstS and is more homologous to eukaryotic DING proteins. Interestingly, bacterial DING proteins have only been detected in certain *Pseudomonas* species. Although DING proteins seem to be ubiquitous in eukaryotes, they are systematically absent from eukaryotic genomic databases and thus are still quite mysterious and poorly characterized. PfluDING displays mitogenic activity towards human cells and binds various ligands such as inorganic phosphate, pyrophosphate, nucleotide triphosphates and cotinine. Here, the crystallization of PfluDING is reported in a monoclinic space group ( $P2_1$ ), with typical unit-cell parameters  $a = 36.7$ ,  $b = 123.7$ ,  $c = 40.8$  Å,  $\alpha = 90$ ,  $\beta = 116.7$ ,  $\gamma = 90^\circ$ . Preliminary crystallographic analysis reveals good diffraction quality for these crystals and a 1.43 Å resolution data set has been collected.

### 1. Introduction

PstSs are 'soluble' phosphate-binding proteins of molecular weight ranging from 34 to 38 kDa that are part of the bacterial ABC cassette phosphate-uptake system. Crystallographic structures reveal that PstSs adopt a characteristic 'Venus flytrap' fold composed of two hinged domains around a conserved phosphate-binding site (Luecke & Quioco, 1990; Vyas *et al.*, 2003).

Recently, eukaryotic proteins that show significant sequence homology between their N-terminal region and PstSs have been reported (Berna *et al.*, 2002). These proteins are named DING proteins owing to their conserved DINGGG N-terminal amino-acid sequence. Interestingly, DING proteins have been independently isolated from a wide range of eukaryotic sources: they have been identified in animals (Riah *et al.*, 2000; Weebadda *et al.*, 2001; Kumar *et al.*, 2004), plants (Berna *et al.*, 2002) and fungi (Chen *et al.*, 2007). For example, a human synovial DING protein was described by Hain *et al.* (1996) and a very similar protein has been characterized from human fibroblast and tumour cells (Adams *et al.*, 2002). A 38 kDa DING protein was isolated from breast carcinoma cells by genistein-agarose affinity chromatography (Belenky *et al.*, 2003). Human phosphate-binding protein (HPBP), the first DING protein to also be identified as a phosphate-binding protein, was purified from a human plasma lipoprotein fraction (Morales *et al.*, 2006). The first plant DING was identified in tobacco by virtue of its strong binding to an *Arabidopsis thaliana* germin-like protein (Berna *et al.*, 2002). Recently, a C-terminally shortened DING protein has been cloned and expressed from *Hypericum perforatum* (Darbinian-Sarkissian *et al.*, 2006). A DING protein has also been identified in a fungus (*Ganoderma lucidum*; Du *et al.*, 2007). Although DING proteins seem to be ubiquitous in eukaryotes, genes coding for DING proteins are systematically absent from genomic databases (Berna *et al.*, 2007). Of course, genes coding for this protein family are expected to exist and, although no complete nucleotide sequences are known, a few partial DNA sequences have been cloned or identified in unannotated parts of genomes (Berna *et al.*, 2007). The only complete amino-acid sequence that is available is that of HPBP, which was completely sequenced using coupled crystallographic and mass-spectrometry analyses (H. Diemer, M. Elias, C. Schaeffer, F. Renault,



© 2007 International Union of Crystallography  
All rights reserved

C. Contreras-Martel, A. Van Dorsselaer & E. Chabriere, in preparation). The crystallographic structure of HPBP complexed with a phosphate ion confirmed that the DING-protein family is related to PstS proteins and adopts the characteristic 'Venus flytrap' motif (Morales *et al.*, 2006).

Interestingly, some *Pseudomonas* species, but to our knowledge no other bacteria, have been shown to express DING proteins that are more homologous to eukaryotic DINGs (Berna *et al.*, 2007). Since these eukaryote-related DING proteins are expressed in particular secretory operons, they are clearly distinguishable from PstS. Furthermore, the low sequence identity (16.1%) between two members of these respective families (*P. fluorescens* DING and *Escherichia coli* PstS) confirms this fact. The eukaryote-related DING protein from *P. fluorescens* (PfluDING) was cloned and expressed. This 372-amino-acid protein was shown to bind phosphate and various ligands such as pyrophosphate, nucleotide triphosphates and cotinine with high affinity and to possess significant mitogenic activity towards human fibroblasts (Scott & Wu, 2005).

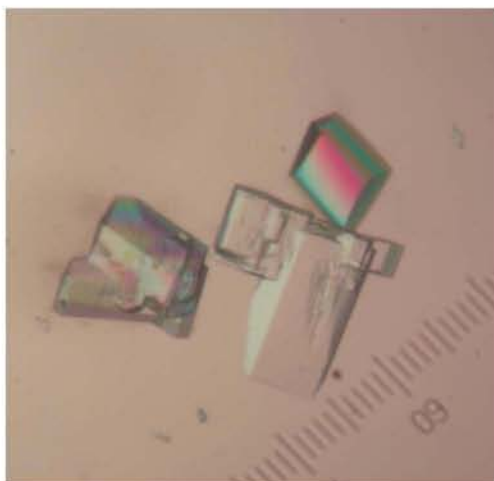
PfluDING is closely related to eukaryotic DINGs (71.3% sequence identity with HPBP) and displays very similar properties. Since the PfluDING gene is available, this protein is a good model for structure–function relationship studies on the DING-protein family.

In this paper, we report the crystallization of PfluDING in a monoclinic crystal form and preliminary crystallographic characterization.

## 2. Materials and methods

### 2.1. Protein expression and purification

Recombinant DING protein from *P. fluorescens* (PfluDING SBW25) was expressed in *E. coli* BL21 (DE3) as a fusion protein with a C-terminal hexahistidine tag and purified from bacterial lysates by affinity chromatography, as described previously (Scott & Wu, 2005). Based on SDS–PAGE experiments, the purity of the recombinant protein was estimated to be greater than 95%. Protein was dialyzed overnight against 20 mM Tris buffer pH 8.5 containing 5 mM  $\beta$ -mercaptoethanol and concentrated using a Centricon centrifugation system with a 3.5 kDa molecular weight cutoff membrane. Protein concentration was determined using UV spectrophotometry ( $\epsilon_{280} = 42\,650\,M^{-1}\,cm^{-1}$ ) and set to 10 mg ml<sup>-1</sup> for crystallization experiments.



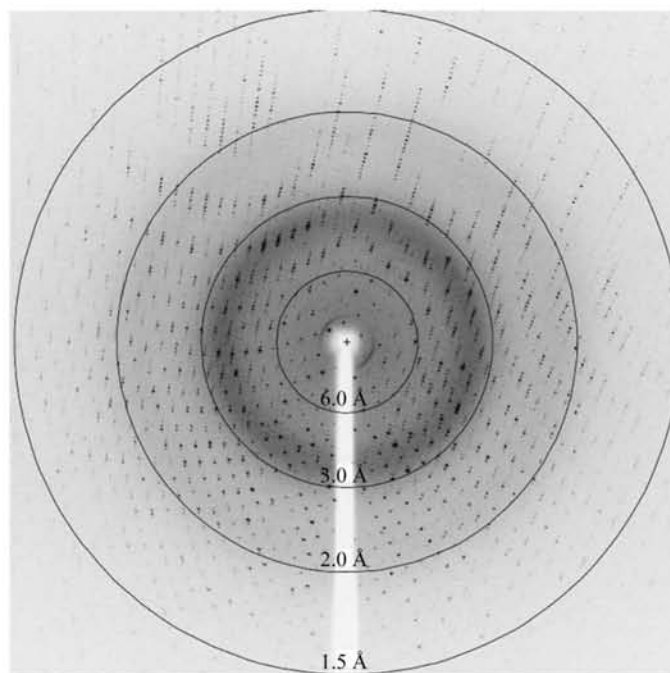
**Figure 1**  
Typical PfluDING crystals. One division represents 20  $\mu$ m.

### 2.2. Protein crystallization

Initial screening was performed at 293 K using sitting-drop vapour diffusion and the commercially available sparse-matrix screens (Jancarik & Kim, 1991) Crystals Screens I and II (Hampton Research Inc.) and Wizard Screens I and II (Emerald Biostructures). A micropipetting crystallization system (Mosquito, LabTech, UK) was used to test three protein:precipitant volume ratios (1:2, 2:2 and 2:1; one unit volume is 150 nl) for each crystallization condition. Condition No. 17 of Wizard Screen I [30% (w/v) PEG 8000, 100 mM acetate buffer pH 4.5 and 200 mM Li<sub>2</sub>SO<sub>4</sub>] turned out to be a promising crystallization condition and was then optimized using the hanging-drop vapour-diffusion method. Polynucleated crystals usually appear within a few days at high PEG concentrations (28–32%). After a week, PfluDING monocrystals had grown in 2  $\mu$ l droplets of protein solution mixed with the same amount of reservoir solution [24–28% (w/v) PEG 8000, 100 mM acetate buffer pH 4.5 and 200 mM Li<sub>2</sub>SO<sub>4</sub>]. Monocrystals had typical dimensions of 0.10  $\times$  0.10  $\times$  0.05  $\mu$ m (Fig. 1). Despite much effort, this protein has a tendency to form polynucleated crystals and seeding strategies are currently being tested to improve crystal quality.

### 2.3. Data collection and processing

Prior to data collection, crystals were soaked for a few seconds in a cryoprotective solution containing 10% (w/v) glycerol, 28% PEG 8000, 100 mM acetate buffer pH 4.5 and 200 mM Li<sub>2</sub>SO<sub>4</sub>. The crystal was then mounted on a loop (LithoLoop, Molecular Dimensions Ltd) and flash-cooled in liquid nitrogen. Data collection was performed at the BM14 UK MAD beamline (ESRF, Grenoble, France) using a wavelength of 0.9535 Å and a MAR CCD (Mosaic 225) detector. A complete data set was obtained by collecting 200 frames with an oscillation step of 1° and 20 s exposure time. Data were processed and scaled using the XDS package (Kabsch, 1993). Crystallographic analyses were performed using the CCP4 suite (Collaborative



**Figure 2**  
Typical diffraction pattern of a crystal of PfluDING. The edge of the frame is at 1.43 Å.

**Table 1**

Crystal data and intensity statistics.

Values in parentheses are for the last resolution shell.

Space group	$P2_1$
Unit-cell parameters ( $\text{\AA}$ , $^\circ$ )	$a = 36.7$ , $b = 123.7$ , $c = 40.8$ , $\alpha = 90.0$ , $\beta = 116.7$ , $\gamma = 90.0$
Temperature (K)	100
Wavelength ( $\text{\AA}$ )	0.9535
Resolution range ( $\text{\AA}$ )	60–1.43 (1.50–1.43)
No. of observed reflections	235730 (21073)
No. of unique reflections	58905 (7166)
Redundancy	4.0 (2.9)
Completeness (%)	98.1 (89.2)
$R_{\text{sym}}$ (%)	2.6 (10.2)
Mean $I/\sigma(I)$	38.9 (10.52)

Computational Project, Number 4, 1994). Crystal data are given in Table 1.

### 3. Results

Despite its relatively moderate size, the PfluDING crystal exhibits remarkable diffraction quality, as indicated by the high resolution of the reported data set (1.43  $\text{\AA}$ ; Fig. 2). Moreover, the very good data statistics at the high-resolution limit [ $R_{\text{sym}}$  of 10.2% and  $I/\sigma(I)$  of 10.52 for the last shell] means that the diffraction potential of the crystal has not yet been fully exploited. The average thermal factor ( $B$  factor), a good indicator of diffraction quality, was estimated to be 8.9  $\text{\AA}^2$  from the Wilson plot. Such a small value, in addition to previous observations, allows us to expect potential improvement in future data collections. Given the unit-cell parameters and the presence of 372 residues in the PfluDING molecule, the Matthews program indicated that the asymmetric unit contains only one PfluDING molecule. Thus, the calculated Matthews coefficient is 1.98  $\text{\AA}^3 \text{Da}^{-1}$ , which corresponds to 37.9% solvent content (Matthews, 1968). A consistent molecular-replacement solution has

been found using the *MOLREP* program with HPBP (PDB code 2cap) as a model. The initial molecular-replacement solution (using data to 3.75  $\text{\AA}$  resolution) resulted in an  $R_{\text{cryst}}$  of 0.38 and a correlation value of 0.58. Refinement and interpretation of the structure of PfluDING at 1.43  $\text{\AA}$  resolution are currently in progress.

### References

- Adams, L., Davey, S. & Scott, K. (2002). *Biochim. Biophys. Acta*, **1586**, 254–264.
- Belenky, M., Prasain, J., Kim, H. & Barnes, S. (2003). *J. Nutr.* **133**, 2497S–2501S.
- Berna, A., Bernier, F., Chabriere, E., Perera, T. & Scott, K. (2007). *Int. J. Biochem. Cell Biol.* doi: 10.1016/j.biocel.2007.02.004.
- Berna, A., Bernier, F., Scott, K. & Stuhlmüller, B. (2002). *FEBS Lett.* **524**, 6–10.
- Chen, Z., Franco, C. F., Baptista, R. P., Cabral, J. M., Coelho, A. V., Rodrigues, C. J. Jr & Melo, E. P. (2007). *Appl. Microbiol. Biotechnol.* **73**, 1306–1313.
- Collaborative Computational Project, Number 4 (1994). *Acta Cryst. D50*, 760–763.
- Darbinian-Sarkissian, N., Darbinyan, A., Otte, J., Radhakrishnan, S., Sawaya, B. E., Arzumanyan, A., Chipitsyna, G., Popov, Y., Rappaport, J., Amini, S. & Khalili, K. (2006). *Gene Ther.* **13**, 288–295.
- Du, M., Zhao, L., Li, C., Zhao, G. & Hu, X. (2007). *Eur. Food Res. Technol.* **224**, 659–665.
- Hain, N. A. K., Stuhlmüller, B., Hahn, G. R., Kalden, J. R., Deutzmann, R. & Burmester, G. R. (1996). *J. Immunol.* **157**, 1773–1780.
- Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
- Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
- Kumar, V., Yu, S., Farrell, G., Toback, F. G. & Lieske, J. C. (2004). *Am. J. Physiol. Renal Physiol.* **287**, 373–383.
- Luecke, H. & Quijcho, F. A. (1990). *Nature (London)*, **347**, 402–406.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Morales, R. *et al.* (2006). *Structure*, **14**, 601–609.
- Riah, O., Dousset, J. C., Bofill-Cardona, E. & Courriere, P. (2000). *Cell. Mol. Neurobiol.* **20**, 653–664.
- Scott, K. & Wu, L. (2005). *Biochim. Biophys. Acta*, **1744**, 234–244.
- Vyas, N. K., Vyas, N. M. & Quijcho, F. A. (2003). *Structure*, **11**, 765–774.
- Weebadda, W. K., Hoover, G. J., Hunter, D. B. & Hayes, M. A. (2001). *Comput. Biochem. Physiol., B*, **130**, 299–312.



# Structure–function relationships in a bacterial DING protein

Soyeon Ahn<sup>a,1</sup>, Sebastien Moniot<sup>b</sup>, Mikael Elias<sup>b</sup>, Eric Chabriere<sup>b</sup>, Donghyo Kim<sup>a,2</sup>, Ken Scott<sup>a,\*</sup>

<sup>a</sup> School of Biological Sciences, The University of Auckland, Auckland, New Zealand

<sup>b</sup> Laboratoire de Cristallographie et Modélisation des Matériaux Minéraux et Biologiques, CNRS-Université Henri Poincaré, 54506 Vandoeuvre-lès-Nancy, France

Received 9 May 2007; revised 20 June 2007; accepted 20 June 2007

Available online 27 June 2007

Edited by Richard Cogdell

**Abstract** A recombinant DING protein from *Pseudomonas fluorescens* has been previously shown to have a phosphate-binding site, and to be mitogenic for human cells. Here we report the three-dimensional structure of the protein, confirming a close similarity to the “Venus flytrap” structure seen in other human and bacterial phosphate-binding proteins. Site-directed mutagenesis confirms the role of a key residue involved in phosphate binding, and that the mitogenic activity is not dependent on this property. Deletion of one of the two hinged domains that constitute the Venus flytrap also eliminates phosphate binding whilst enhancing mitogenic activity.

© 2007 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

**Keywords:** DING protein; Phosphate binding; Human phosphate-binding protein; pstS protein

## 1. Introduction

DING proteins were identified in a wide range of eukaryotic organisms, on the basis of conserved DINGGG N-terminal amino acid sequences [1,2]. Despite much effort, no complete gene sequences are known for these proteins, and only a single complete amino acid sequence. The latter was determined from the X-ray crystallographic structure determination of a human phosphate-binding protein (HPBP), purified from the high-density lipoprotein fraction of plasma, and confirmed by mass spectrometric peptide sequencing [3–6].

The tertiary structure of HPBP confirms the ability of this protein to bind a single phosphate ion, in the same manner as the bacterial pstS proteins, which sequester phosphate for cellular uptake by the ABC phosphate transporter. The pstS and the HPBP structures fit a model known as the “Venus flytrap”, in which two globular domains hinge together to form the phosphate-binding site, with eight conserved residues H-bonding to phosphate [5,7,8].

Partial gene and protein sequences are also known for several plant DING proteins [1,2], one of which was isolated

and cloned from St. John’s Wort (*Hypericum perforatum*) on the basis of anti-HIV activity in lymphocytes. This p27<sup>SJ</sup> protein is clearly homologous to existing DING sequences [2], but is C-terminally truncated, possibly as a result of proteolysis by coagulation factor Xa action during the preparation of plant extracts [9]. In the p27<sup>SJ</sup> protein, all of the phosphate-binding residues are conserved, but it is not known if the protein binds a phosphate ion.

PstS proteins are ubiquitous in bacteria, but some *Pseudomonas* genomes contain genes which are much more homologous to the eukaryotic DING proteins. These genes form two families of phosphate-binding proteins, which may be involved in phosphate scavenging [2]. One such gene, from *P. fluorescens*, has been cloned and expressed [10]. The resulting protein (PfluDING) has a single phosphate-binding site. It also possesses mitogenic activity towards human fibroblasts, which is consistent with similar activity ascribed to human DING proteins [11–14]. In stimulating DNA synthesis, the bacterial DING protein does not stimulate cellular phosphate uptake [10]. The PfluDING sequence most closely resembles that of p27<sup>SJ</sup> (Fig. 1).

Because there are no complete DING gene sequences from eukaryotes, PfluDING offers the most convenient model for pursuing structure–function relationships in DING proteins. We here describe experiments to establish the tertiary structure of the PfluDING protein, and to investigate further the phosphate-binding role in relation to mitogenic activity.

## 2. Methods

A recombinant PfluDING protein was expressed in *E. coli* BL21 (DE3) as a fusion protein with a C-terminal hexahistidine tag, coded in the pET22b(+) plasmid vector, and purified from bacterial lysates by affinity chromatography, as previously described [10].

Crystallization of PfluDING was performed using the hanging drop vapor diffusion method and by mixing the reservoir solution composed of 25% (w/v) PEG 8000, 100 mM acetate buffer pH 4.5 and 200 mM Li<sub>2</sub>SO<sub>4</sub> with a PfluDING solution at 10 mg/mL. A single dataset was collected at the BM14 UK MAD beamline (ESRF, Grenoble, France) at a wavelength of 0.953 Å. Crystallization and data collection are further described in Moniot et al., 2007 [15]. Statistics of the dataset are summarized in Table 1.

The initial phase was obtained using the molecular replacement method with the MOLREP program [16]. Since the PfluDING amino acid sequence displays 71% identity with the HPBP sequence, HPBP structure was used as a starting model (PDB code 2CAP). The PfluDING model was built through refinement and manual building cycles using REFMAC5 [17] and Coot [18]. The final model shows *R* and *R*<sub>free</sub> factor values of 13.1% and 16.4%, respectively (Table 1). Coordinates and associate structure factor files have been deposited within the Protein Data Bank (ID code 2Q9T). Figs. 2–5 were drawn with PyMOL (<http://www.pymol.org>).

\*Corresponding author. Fax: +64 9 3737 668.

E-mail address: k.scott@auckland.ac.nz (K. Scott).

<sup>1</sup>Present Address: Murdoch Childrens Research Institute, The University of Melbourne, Melbourne, Australia.

<sup>2</sup>Present Address: The Liggins Institute, The University of Auckland, Auckland, New Zealand.

```

PfluDING  DINGGGATLPQALYQTSVGLTAGFAPYIGVSGNGKAAFLNNDYTKFQAGVTNKNVHWAGSDSKLSA
HPBP      DINGGGATLPQKLYLTPDVLTAGFAPYIGVSGKGIKIAFLENKYQFG-TDTTKNVHWAGSDSKLTA
HypDING   MADINGGGATLPQALYQTSVGLTAGFAPYIGVSGNGKAAFLNNDYTKFQAGVTNKNVHWAGSDSKLSA

PfluDING  TELSTYASAKQPTWGKLIQVPSVGTVAIPFNKSGSAAVDLSVQELCGVFSGRINTWDGISGSGRTG
HPBP      TELATYAADKEPGWGKLIQVPSVATVAIPFRKAGANAVDLSVKELCGVFSGRITADWSGITGAGRSG
HypDING   TELSTYASAKQPTWGKLIQVPSVGTAVAIPFNKSGTAAVDLSVSELCGVFSGRITDWSGISGSGRTG

PfluDING  PIVVYRSESSTTELFTRFLNAKNAETGNFAVTTTFGTFSFSGGL--PAGAVAATGSQGVMTALAA
HPBP      PIQVVYRAESSSTTELFTRFLNAKCTTQPGTFVTTVFANSYSLGLSPLAGAVAAIGSVGVMAADND
HypDING   AITVVYRSESSTTELFTRFLNAK-AETGTFNISTTFGTSTYTGGL--PAGAVSAAGSQGVMTALAG

PfluDING  G---DGRITYMSPDFAAPTLAGLDDATKVARVGNVA--TNTQGVSPAAANVSAIIGAVVPVPAADR
HPBP      VTTAQGRITYISPDFAAPTLAGLDDATKVARVTKGSSSSGGGAEKSPAAANVSAISVPLPAADR
HypDING   A---DGGTTYMSPDFAAPTLAGLDDATKVARVGNVA--TNTAGVSPAAANVSAIINAVVPVPASTEK

PfluDING  SNPDWVVPVFGPDNTAGVQPYPTSGYPILGFTNLIFSQCADATQTQVRDFFTKHYGASNNNDAAI
HPBP      GDNVWVTPVFGAVTGGGVVAYPDSGYPIILGFTDLIFSECYANATQTQVRNFFTKHYGTANDNAAI
HypDING   P-----

PfluDING  TANAFVPLPTAWKATVRASFLTASNALSIGNNTNVCNGIGRPLLE
HPBP      QANAFVPLPSNWKA AVRASYLTASNALSIGDSAVCGGKGRPE
HypDING   -----

```

Fig. 1. Alignment of the amino acid sequences of three DING proteins. PfluDING: the recombinant PfluDING protein sequence. HPBP: human plasma phosphate-binding protein sequence. HypDING: *Hypericum perforatum* DING protein sequence, p27<sup>SJ</sup>. Matches in blue, with the conserved residues involved in phosphate binding bold and underlined. Sequences are aligned with reference to the structural superposition of the PfluDING and HPBP ternary structures.

Table 1  
Dataset and refinement statistics

<i>Data collection</i>	
Wavelength (Å)	0.9535
Resolution (last bin) (Å)	62–1.43 (1.50–1.43)
Space group	P2 <sub>1</sub>
Unit cell (Å)	$a = 36.7$ , $b = 123.7$ , $c = 40.8$ , $\alpha = 90.0$ , $\beta = 116.7$ , $\gamma = 90.0$
Observed reflections (last bin)	235 730 (21 073)
Unique reflection (last bin)	58 905 (7166)
Completeness (last bin) (%)	98.1 (89.2)
R <sub>sym</sub> (last bin) (%)	2.6 (10.2)
Mean $I/\sigma(I)$ (last bin)	38.9 (10.5)
<i>Refinement statistics</i>	
Resolution range (Å)	61.90–1.43
$R_{work}/R_{free}$	13.1/16.4
Number of protein atoms	3743
Number of water molecules	766
Average $B$ factor (Å <sup>2</sup> )	10.15
Rms bond length (Å)	0.007
Rms bond angle (°)	1.119

$R_{work} = \frac{\sum (F_o - F_c)}{\sum F_o}$ , where  $F_o$  denotes the observed structure factor amplitude and  $F_c$  denotes the structure factor amplitude calculated from the model,  $R_{free}$  is as for  $R_{work}$ , but it is calculated with 5% of randomly chosen reflections omitted from the refinement.

Prior to mutagenesis, possible structural changes in putative variants were explored, using the Psipred software [19]. A mutated protein with a T147N substitution was created by overlap extension PCR [20] with PfluDING cDNA as the starting template, using the following primers (forward SAD1: 5'GC GAA AGC AGT GGT AAC ACT GAG CTG TTC, reverse SAD2: 5'GAA CAG CTC AGT GTT ACC ACT GCT TTC GC) in conjunction with the forward and reverse primers originally used to amplify the PfluDING coding sequence (MD5 and MD6), and using the same conditions for the PCR reaction [10]. Following a sequence check of the mutated cDNA, restriction, ligation and cloning into the pEt22b+ plasmid were carried out as before [10].

To create the truncated version of PfluDING, the primers NT1 (5'CAA CCA TGG TGG GGA CTT CGG TTG CC) and CT2 (5'CGT CTC GAG TGC GCC GAT AGC GGC AGA CAC) were created, and used for PCR with the normal PfluDING cDNA as tem-

plate. NT1 causes truncation immediately before the codon CTG, corresponding to V90 in PfluDING, and CT2 caused truncation immediately after the codon GCA, for A252. The resulting cDNA was of about 500 bps, corresponding to the predicted size of 506 bps, and had the correct nucleotide sequence. The primers contained restriction sites for NcoI and XhoI, respectively, so the cDNA was restricted, ligated and cloned into the pET22b(+) plasmid, as for other PfluDING preparations. Both mutated PfluDING derivatives were expressed and purified as for the normal protein.

Western blotting was carried out with a rabbit antiserum to a conjugated human DING protein N-terminal peptide, as previously described [13]. Double immunodiffusion was carried out with a rabbit antiserum prepared against purified recombinant PfluDING, using standard immunization protocols.

Filter-binding assays for <sup>32</sup>P-phosphate binding to DING proteins, and mitogenic assays, based on incorporation of <sup>3</sup>H-thymidine, were also carried out as before [10]. For the phosphate-binding experiments, 3.6 μg of each protein was used, in triplicate assays. Differing concentrations of each protein (0.6–3.0 μM) were added to cell cultures for mitogenic assays.

### 3. Results and discussion

The three-dimensional model of PfluDING reveals an elongated fold composed of two adjacent globular domains. Each domain is constituted by a central β-sheet core flanked by α-helices and contains a disulfide bridge (C114–C159 and C300–C363). Interconnected by an antiparallel two-stranded β-sheet acting as a hinge, the two domains form a deep cleft wherein is bound a phosphate molecule. This fold, known as a Venus flytrap, is very similar to those of the sixth family of solute binding proteins (SBP) [21,22]. Structural superposition (Fig. 2) shows good correspondence between our model and the structures of *E. coli* phosphate-binding protein (pdb code 1IXH) and HPBP (pdb code 2CAP), with root mean square deviations (rmsd) of 1.881 Å on 276 Cα atoms and 0.649 Å on 366 Cα atoms, respectively. Similar to the eukaryotic HPBP [5], the PfluDING structure exhibits four protruding loops when compared with the *E. coli* binding protein. Though these loops are a little shorter, they superimpose almost perfectly

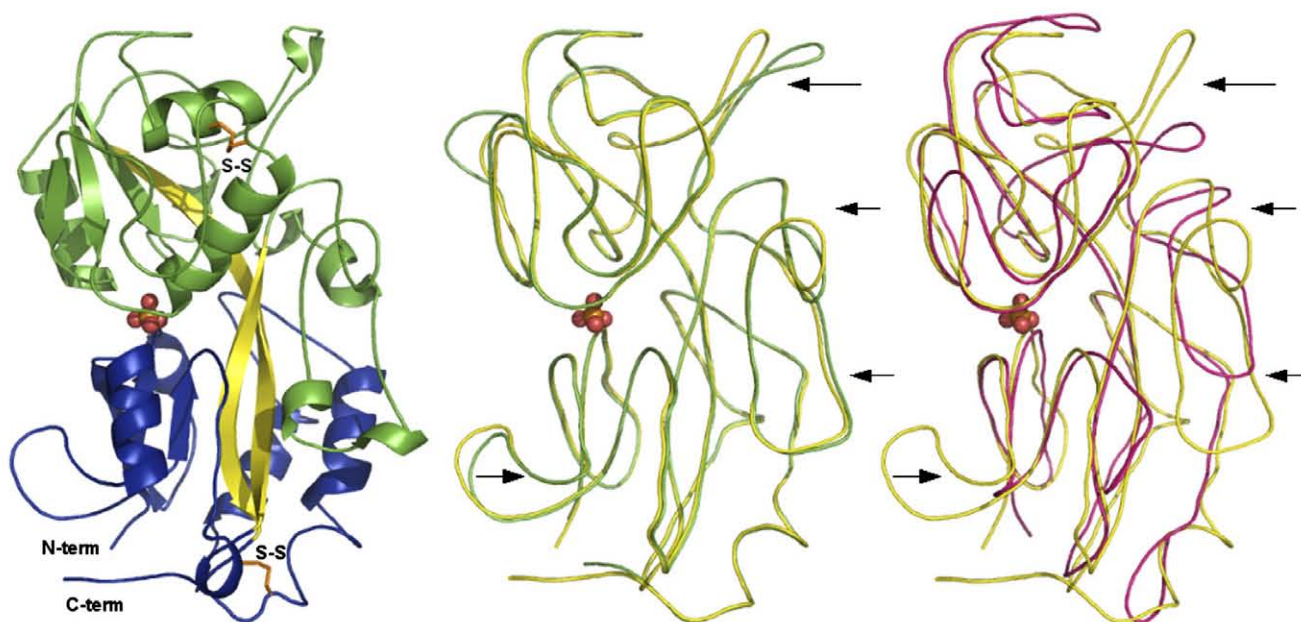


Fig. 2. Ternary structure superposition. (left) The overall structure of the DING protein from *Pseudomonas fluorescens* is shown in cartoon mode. The two domains of the protein are colored in blue and green and the hinge is presented in yellow. Both disulfide bridges are presented in orange. The phosphate molecule is represented in ball and stick mode (in red). (middle) Structural comparison between the *P. fluorescens* DING protein (in yellow) and the human phosphate-binding protein (HPBP) (pdb code 2CAP) (in green). (right) Structural comparison between the *P. fluorescens* DING protein (in yellow) and the phosphate-binding protein from *Escherichia coli* (pdb code 1IXH) (in purple). The phosphate from the PfluDING structure is represented in ball and stick mode (in red). Protruding loops of PfluDING and HPBP are indicated by arrows.

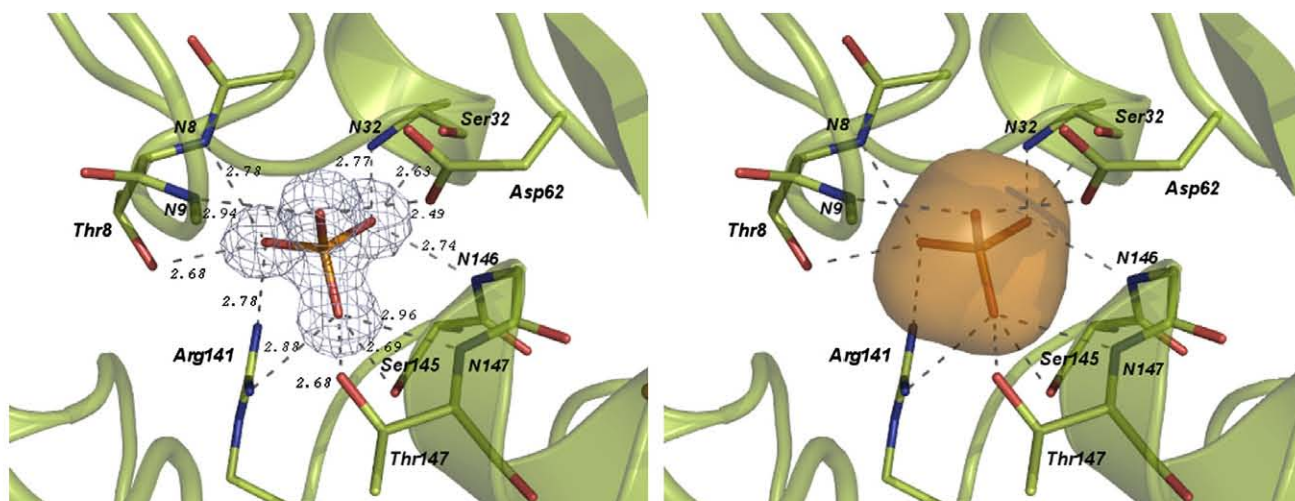


Fig. 3. Views of the phosphate-binding site. Residues involved in the binding of the phosphate molecule are labeled. The 12 H-bonds between the protein and the phosphate molecule are represented in dashed lines. Side chains are shown in sticks and main chain is represented in cartoon mode. (A) The  $2F_o - F_c$  map calculated omitting the phosphate molecule was contoured at  $5\sigma$ . Hydrogen bond distances are indicated in Å. (B) Binding site cavity. Surface is compute from the protein Van der Waals surface using a 1.6 Å probe sphere radius.

with those of HPBP (Fig. 2). Structural comparison thus reveals that PfluDING superimposes more closely on HPBP, and possesses similar protruding loops. This strongly supports the hypothesis deduced from the high sequence identity between both proteins (71%) that bacterial PfluDING is closely related to the eukaryotic DING protein family [2].

The binding site of PfluDING, located between the two domains, is totally buried and sequesters a phosphate molecule that copurified with the protein. The phosphate is tightly bound by 12 hydrogen bonds formed with 8 residues distrib-

uted on either side of the cleft (Fig. 3). Residues implicated in phosphate binding are T8, L9, S32, D62, R141, S145, G146 and T147. Hydrogen bonds distances range from 2.49 to 3.16 Å. The shortest hydrogen bond involves a phosphorus oxygen atom and a carboxylic oxygen from the D62 side chain. Such short bonds have also been reported in *E. coli* pstS (1IXH), *M. tuberculosis* pstS-1 (1PC3), and HPBP (2CAP) structures with distances as small as 2.43 Å. Being potentially the only hydrogen bond acceptor in the binding site cavity of PfluDING, D62 should play a key role in phosphate specificity

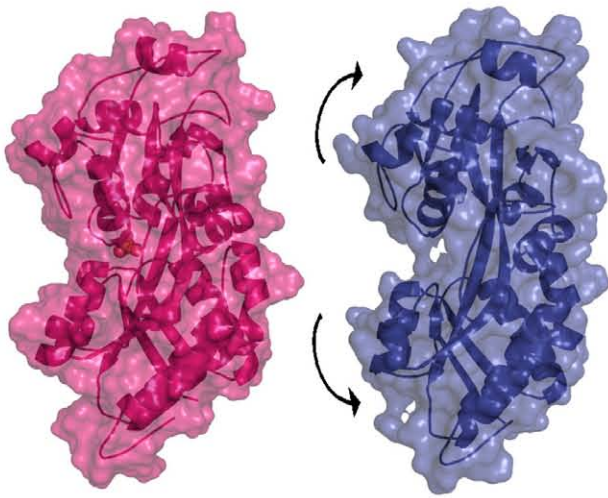


Fig. 4. The Venus “flytrap” motion. The bending motion of the Venus “flytrap” fold is presented through comparison of “closed” (wild-type) (in purple) and “open” (T141D mutant) (in blue) forms of *E. coli* PBP (pdb code 1OIB and 1IXH). Both proteins are presented in a cartoon mode with their corresponding molecular surface. The inorganic phosphate bound in the closed form is represented as balls and sticks (red).

by accepting protonated phosphate species [7]. The binding sites residues are completely identical to those identified in the HPBP structure.

Previous studies have shown that PfluDING bound  $^{32}\text{P}$ -phosphate, corresponding to approximately one mole of phosphate per mole of protein, with a  $K_D$  of  $1.7\ \mu\text{M}$  [10]. The  $K_D$  values for the *E. coli* and *P. aeruginosa* pstS proteins are approximately  $1\ \mu\text{M}$  and  $0.34\ \mu\text{M}$ , respectively [7,23]. It is

not immediately obvious from the structural comparisons why the PfluDING and pstS proteins should have different affinities for phosphate.

The “Venus flytrap” fold was shown to undergo large conformational changes upon ligand binding. These structures were shown to adopt a “closed” conformation whilst ligand is bound, and an “open” conformation without ligand, through a bending motion of the two domains around the hinge [22]. Yao et al. obtained a T141D mutant of *E. coli* pstS (equivalent to a T147 substitution in PfluDING) with a reduced phosphate affinity [24,25]. From this mutant it was possible to obtain both “closed” and “open” forms of this protein (Fig. 4).

With the aim of obtaining similar properties, the T147N PfluDING mutant was prepared, successfully expressed and purified. The normal PfluDING protein ( $3.6\ \mu\text{g}$ ) bound  $3.3 \times 10^4$  ( $\pm 0.2 \times 10^4$ ) dpm of the radioisotopic phosphate, whereas in the same quantity of the PfluDING T147N mutant protein, phosphate binding was reduced by 85%, to  $0.5 \times 10^4$  ( $\pm 0.05 \times 10^4$ ) dpm.

Another putative mutant version of PfluDING was created, with the substitutions R141L, S145I and T147I. Although expressed in our *E. coli* system, the resulting protein was almost entirely insoluble (not shown).

Truncation of PfluDING was considered because of the discovery of the biologically-active DING protein from *H. perforatum*, which comprises 263 residues with 89% sequence identity with the N-terminal sequence of the PfluDING protein (see Fig. 1). Though the nature of the activity (inhibition of gene expression and replication in HIV-infected cells [7]) is very different from those observed with most other DING proteins, we thought it possible that such proteins might retain other activities if similarly truncated. Accordingly, we designed

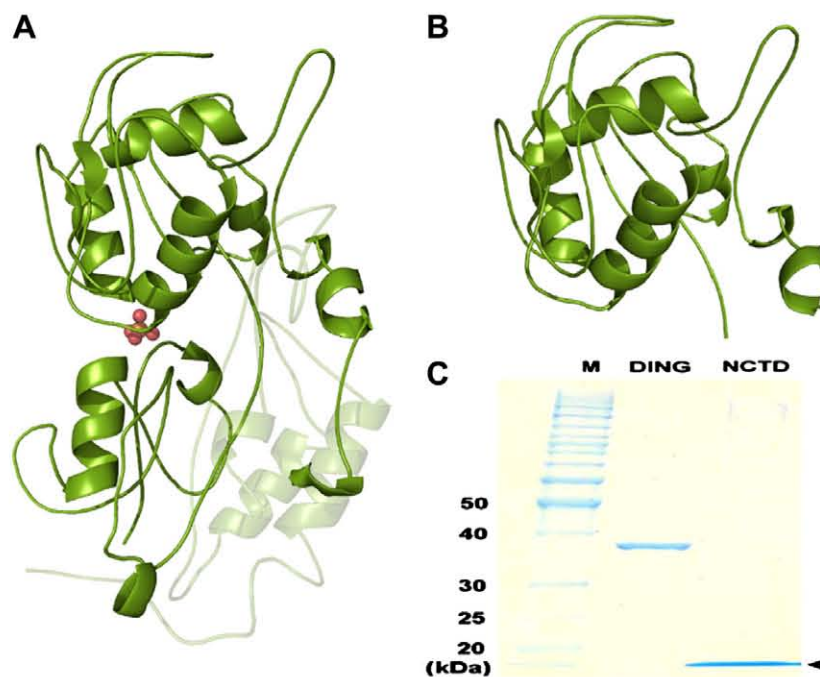


Fig. 5. N- and C-terminal truncation. (A) C-terminal truncated form (Asp1–Arg261) deduced from the complete structure of the PfluDING (truncated part is shown as transparent). (B) N- and C-terminal truncated form (Val190–Ala252) deduced from the complete structure. (C) SDS-PAGE of standard protein markers (lane M), wild-type PfluDING (lane DING) and the N- and C-terminally truncated mutant (lane NCTD).



a primer to amplify PfluDING cDNA, in conjunction with MD5, between the N-terminal codon and the codon for R261. Although the corresponding cloned, truncated protein could be readily expressed, only traces of the expected 27 kDa protein were seen, with much larger quantities of 25, 23 and 20 kDa proteins, corresponding to products of N-terminal proteolysis (not shown). It is clear from the structure (Fig. 5A) that the deletion of the C-terminal part of the “lower lobe” of the “Venus flytrap” could expose the remainder of the structure to proteolysis, either during bacterial protein processing, or during bacterial lysis and protein purification. We therefore resolved to remove the proteolytically-sensitive region by genetic truncation. Consideration of the structure of PfluDING suggested that a concomitant extension of the C-terminal truncation would effectively create a 17 kDa protein comprising just the “upper lobe” of the complete PfluDING (Fig. 5B). Thus, primers were designed to amplify the PfluDING coding sequence from the V90 codon to that for A252.

The details of the production of this NCTD protein are summarized in Fig. 5C. It has a molecular weight of 17 kDa, as expected from the genetic truncation. It retains a partial immunological cross-reaction with the native PfluDING, but cannot be detected in Western blots with an antiserum specific for the DING N-terminus (not shown). When used in phosphate-binding experiments, the extent of binding was reduced by over 92%, compared to native PfluDING ( $0.3 \times 10^4$  ( $\pm 0.2 \times 10^4$ ) dpm for the truncated mutant, compared to  $3.3 \times 10^4$  ( $\pm 0.2 \times 10^4$ ) dpm for the native protein). This is unsurprising, given that four of the eight phosphate-coordinating residues (T8, L9, S32 and D62) have been eliminated, together with one side of the phosphate-binding cleft.

The mitogenic activity of the two mutants was compared with the native protein (Table 2). We have previously shown that PfluDING is mitogenic for human fibroblasts, albeit much less efficiently than is a DING protein secreted by these cells [10,12,13]. Here we show that, in the absence of bound phosphate, the protein not only retains mitogenic activity, but the activity is enhanced. This observation may indicate that the DING protein has to be in an “open” conformation for mitogenic activity, and that the “closing of the flytrap”, which accompanies phosphate binding, may reduce this activity. A similar phenomenon occurs when one domain of the protein is absent, indicating that the “upper” domain, V90-A252, contains a site for mitogenic signalling. Further work is clearly needed to elucidate this effect, but the availability of the NCTD mutant should make it easier to identify the site

responsible for mitogenic activity. X-ray structures of both the T147N mutant, possibly in an “open” conformation, and the NCTD mutant should also be of great interest in helping to understand the molecular mechanism of modulation of mitogenic activity.

We previously showed that organophosphates would also bind to the phosphate-binding site in PfluDING [10]. In particular, ADP appeared to be an effective competitive inhibitor for phosphate binding. Sphingosine 1-phosphate was tested as a putative inhibitor, as this reagent has been shown to function as both an intracellular or extracellular component of signalling systems [26]. However, at concentrations of 0.1 and 1 mM, sphingosine 1-phosphate had essentially no effect on phosphate binding to the PfluDING protein. Other bioactive organophosphates, nucleotides and phospholipids are obvious candidates for future investigation in this regard.

The PfluDING structure (Figs. 2 and 3) reveals that the bound phosphate ion almost fills the buried cavity. To illustrate, the nearest crystallographic solvent molecules are situated 6 Å away from the phosphate. Accommodation of any larger phosphate metabolites would require conformational changes, presumably in accordance with the “Venus flytrap” motion. Organophosphate binding should thus induce a partial opening of the cleft, but it is not yet clear what effect this will have upon the mitogenic activity of the protein.

**Acknowledgements:** We thank Manasa Ramakrishna for assistance with the running of the SP<sup>3</sup> structure prediction programme, to model truncated variants of PfluDING. Financial support was provided by the University of Auckland Staff Research Fund.

## References

- [1] Berna, A., Bernier, F., Scott, K. and Stuhlmuller, B. (2002) Ring up the curtain on ding proteins. *FEBS Lett.* 524, 6–10.
- [2] Berna, A., Bernier, F., Chabriere, E., Perera, T. and Scott, K. (in press) DING proteins; novel members of a prokaryotic phosphate-binding protein superfamily which extends into the Eukaryotic Kingdom. *Intl. J. Biochem. Cell Biol.* Feb 2007. <<http://dx.doi.org/10.1016/j.biocel.2007.02.004>>.
- [3] Fokine, A., Morales, R., Contreras-Martel, C., Carpentier, P., Renault, F., Rochu, D. and Chabriere, E. (2003) Direct phasing at low resolution of a protein copurified with human paraoxonase (PON1). *Acta Crystallogr. Sect. D-Biol. Crystallogr.* 59, 2083–2087.
- [4] Renault, F., Chabriere, E., Andrieu, J.P., Dublet, B., Masson, P. and Rochu, D. (2006) Tandem purification of two HDL-associated partner proteins in human plasma, paraoxonase (PON1) and phosphate binding protein (HPBP) using hydroxyapatite chromatography. *J. Chromatogr. B* 836, 15–21.
- [5] Morales, R., Berna, A., Carpentier, P., Contreras-Martel, C., Renault, F., Nicodeme, M., Chesne-Seck, M.-L., Bernier, F., Dupuy, J., Schaeffer, C., Diemer, H., Van-Dorsselaer, A., Fontecilla-Camps, J.C., Masson, P., Rochu, P. and Chabriere, E. (2006) Serendipitous discovery and X-ray structure of a human phosphate binding apolipoprotein. *Structure* 14, 601–609.
- [6] Diemer, H., Elias, M., Renault, F., Contreras-Martel, C., Schaeffer, C., Van Dorsselaer, A. and Chabriere, E. (in preparation) Tandem use of X-ray crystallography and mass spectrometry to obtain ab initio the complete and exact amino acids sequence of HPBP, a human 38 kDa apolipoprotein.
- [7] Luecke, H. and Quioco, F.A. (1990) High specificity of a phosphate transport protein determined by hydrogen bonds. *Nature* 347, 402–406.
- [8] Chang, G. and Roth, C.B. (2001) Structure of MsbA from *E. coli*: a homolog of the multidrug resistance ATP binding cassette (ABC) transporters. *Science* 293, 1793–1800.
- [9] Darbinian-Sarkissian, N., Darbinyan, A., Otte, J., Radhakrishnan, S., Sawaya, B.E., Arzumanyan, A., Chipitsyna, G., Popov,

Table 2  
Relative stimulation of thymidine incorporation in human fibroblasts by native and mutated PfluDING

Reagent	Incorporation of 3H thymidine (dpm $\times 10^{-3}$ )	Relative incorporation (%)
3% FCS	1.2 ( $\pm 0.4$ )	100
PfluDING		
0.63 $\mu$ M	1.0 ( $\pm 0.3$ )	83
1.3 $\mu$ M	2.3 ( $\pm 0.2$ )	192
T147N		
0.63 $\mu$ M	5.5 ( $\pm 0.4$ )	458
1.3 $\mu$ M	6.4 ( $\pm 0.4$ )	533
Trunc.		
1.5 $\mu$ M	7.3 ( $\pm 0.6$ )	608
3.0 $\mu$ M	8.2 ( $\pm 0.2$ )	683

- Y., Rappaport, J., Amini, S. and Khalili, K. (2003) p27<sup>SJ</sup>, a novel protein in St. John's Wort, that suppresses expression of HIV-1 genome. *Gene Therapy* 13, 288–295.
- [10] Scott, K. and Wu, L. (2005) Functional properties of a recombinant bacterial DING protein: comparison with a homologous human protein. *Biochim. Biophys. Acta* 1744, 234–244.
- [11] Hain, N.A.K., Stuhlmüller, B., Hahn, G.R., Kalden, J.R., Deutzmann, R. and Burmester, G.R. (1996) Biochemical characterisation and microsequencing of a 205 kDa synovial protein stimulatory for T cells and reactive with rheumatoid factor containing sera. *J. Immunol.* 157, 1773–1780.
- [12] Bush, D., Fritz, H., Knight, C., Mount, J. and Scott, K. (1998) A hirudin-sensitive, growth-related proteinase from human fibroblasts. *Biol. Chem.* 379, 225–229.
- [13] Adams, L., Davey, S. and Scott, K. (2002) The DING protein: an autocrine growth-stimulatory protein related to the human synovial stimulatory protein. *Biochim. Biophys. Acta* 1586, 254–264.
- [14] Belenky, M., Prasain, J., Kim, H. and Barnes, S. (2003) DING, a genistein target in human breast cancer: a protein without a gene. *J. Nutr.* 133, 2497S–2501S.
- [15] Moniot, S., Elias, M., Kim, D., Scott, K. and Chabriere, E. (in press) Crystallization, diffraction data collection and preliminary crystallographic analysis of DING protein from *Pseudomonas fluorescens*. *Acta Cryst. F* 63, 1–3.
- [16] Vagin, A. and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement. *J. Appl. Cryst.* 30, 1022–1025.
- [17] Murshudov, G.N., Vagin, A.A. and Dodson, E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Cryst. D* 53, 240–255.
- [18] Emsley, P. and Cowtan, K. (2004) *Coot*: model-building tools for molecular graphics. *Acta Cryst. D* 60, 2126–2132.
- [19] McGuffin, L.J., Bryson, K. and Jones, D.T. (2000) The PSIPRED protein structure prediction server. *Bioinformatics* 16, 404–405.
- [20] Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K. and Pease, L.R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* 77, 51–59.
- [21] Tam, R. and Saier, M.H. (1993) Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. *Microbiol. Rev.* 57, 320–346.
- [22] Felder, C.B., Graul, R.C., Lee, A.Y., Merkle, H.P. and Sadee, W. (1999) The Venus Flytrap of periplasmic binding proteins: an ancient protein module present in multiple drug receptors. *AAPS Pharm. Sci.* 1, 2.
- [23] Poole, K. and Hancock, R.E.W. (1984) Phosphate transport in *Pseudomonas aeruginosa*. *Eur. J. Biochem.* 144, 607–612.
- [24] Yao, N., Ledvina, P.S., Choudhary, A. and Quioco, F.A. (1996) Modulation of a salt link does not affect binding of phosphate to its specific active transport receptor. *Biochemistry* 32, 2079–2085.
- [25] Wang, Z., Choudhary, A., Ledvina, P.S. and Quioco, F.A. (1994) Fine tuning the specificity of the periplasmic phosphate transport receptor. Site-directed mutagenesis, ligand binding, and crystallographic studies. *J. Biol. Chem.* 269, 25091–25094.
- [26] Watterson, K., Sankala, S., Milstien, S. and Spiegel, S. (2003) Pleiotropic actions of sphingosine 1-phosphate. *Prog. Lipid Res.* 42, 344–357.







## Résumé

La glycéraldéhyde-3-phosphate déshydrogénase (GAPDH) et l'érythrose-4-phosphate déshydrogénase (E4PDH) sont deux enzymes structurellement apparentées, appartenant à la famille des aldéhyde déshydrogénases. Ces deux enzymes catalysent respectivement la phosphorylation oxydative et l'oxydation irréversible de leur substrat aldéhydique, le glycéraldéhyde-3-phosphate (G3P) et l'érythrose-4-phosphate (E4P), en présence de cofacteur  $\text{NAD}^+$ . Dans les deux cas, le mécanisme réactionnel procède en deux étapes, nommées acylation et désacylation. Commune aux deux enzymes, la première étape est une oxydoréduction qui conduit, *via* l'attaque nucléophile d'une cystéine essentielle suivie d'un transfert d'hydrure ( $\text{H}^+ + 2\text{e}^-$ ), à l'oxydation de la molécule de cofacteur et à la formation d'un intermédiaire covalent entre l'enzyme et le substrat appelé thioacylenzyme. Les deux enzymes diffèrent ensuite par la nature de la seconde étape de la réaction. Dans le cas de la GAPDH, l'intermédiaire thioacylenzyme subit une phosphorolyse qui conduit à la libération du produit de la réaction, le 1,3-diphosphoglycérate (1,3-DPG), alors que dans le cas de l'E4PDH, c'est une hydrolyse qui libère le 4-phosphoérythronate (4PE).

Les travaux réalisés au cours de la thèse ont eu pour but d'élucider les relations structure-fonction de ces deux enzymes *via* la résolution de structures cristallographiques de ces enzymes en complexe avec leurs ligands physiologiques ou avec des analogues.

Si la GAPDH est une enzyme bien caractérisée sur le plan biochimique et structural, la contribution des deux sites de reconnaissance anionique lors de la catalyse reste incomprise et nécessitait la détermination des structures de l'intermédiaire thioacylenzyme et du complexe enzyme-produit. Dans ce but, nous avons développé une approche qui consiste à utiliser le fait qu'en l'absence de phosphate, l'étape de désacylation est fortement ralentie, car la réaction d'hydrolyse est très peu efficace. Dans ces conditions, l'intermédiaire thioacylenzyme formé après diffusion du substrat dans les cristaux d'holoenzyme s'accumule avant d'être piégé par congélation. Le complexe enzyme-produit a également pu être caractérisé grâce à la diffusion du 1,3-DPG dans des cristaux de mutants de la cystéine essentielle. Ce manuscrit présente la mise au point des conditions de cristallisation, d'accumulation de l'intermédiaire thioacylenzyme (dont la formation a été suivie *in situ* dans le cristal par microspectrophotométrie), les collectes de données de diffraction ainsi que l'analyse des structures et les implications pour le mécanisme catalytique. Les résultats obtenus mettent principalement en évidence un mouvement de "va-et-vient" du substrat au cours de la catalyse entre les deux sites de reconnaissance anionique. Ce mouvement est lié à l'étape d'échange du cofacteur intervenant à l'issue de l'acylation et préalablement à la désacylation.

Bien que structurellement proche des GAPDH, l'E4PDH est une enzyme pour laquelle peu de données sont disponibles. De nombreuses questions restent donc posées concernant notamment les déterminants de l'affinité pour le cofacteur, la nature du (ou des) site(s) de reconnaissance anionique, ou encore le mécanisme d'activation de la molécule d'eau nécessaire à une étape d'hydrolyse efficace. Sont présentées dans ce manuscrit trois structures cristallographiques de l'E4PDH d'*E. coli*, sous forme apoenzyme, en présence de phosphate ou en complexe avec une analogue de cofacteur. L'analyse de ces structures a notamment permis de montrer l'existence d'un site de reconnaissance anionique dans le site actif de l'enzyme, de proposer des hypothèses quant à la faible affinité de l'enzyme pour le cofacteur, et d'identifier un candidat possible pour l'activation de la molécule d'eau hydrolytique.

Mots-clés : Aldéhyde déshydrogénase, complexe michaelien, cristallographie des rayons X, intermédiaire réactionnel, microspectrophotométrie, thioacylenzyme.