



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

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.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>



FACULTE DES SCIENCES

U.F.R. S.T.M.P.

Groupe de Formation Doctorale Matériaux



Thèse

Présentée pour l'obtention du titre de

Docteur de l'Université Henri Poincaré, Nancy I

en Sciences des Matériaux

par Claude DIDIERJEAN

**Reconnaissance spécifique des cofacteurs NAD/NADP
par les oxydoréductases : Etude cristallographique du site de fixation
du coenzyme dans la Glycéraldéhyde-3-phosphate déshydrogénase
de *Bacillus stearothermophilus***

Soutenue le 25 avril 1997 devant la commission d'examen :

Membres du Jury :

Rapporteurs :	M. Jean Paul BRIAND	Directeur de Recherche, IBMC, Strasbourg
	M. Otto DIDEBERG	Directeur de Recherche, IBS (CNRS-CEA), Grenoble
Examineurs :	M. André AUBRY	Directeur de thèse, Directeur de Recherche, UHP, Nancy 1
	M. Guy BRANLANT	Professeur, UHP, Nancy 1
	M. Gerd FOLKERS	Professeur, Département de Pharmacie, ETH, Zürich
	M. Claude LECOMTE	Professeur, UHP, Nancy 1
	M. Dino MORAS	Directeur de Recherche, IGBMC, Illkirch

Le travail présenté dans ce mémoire a débuté au Laboratoire de Minéralogie Cristallographie et Physique Infrarouge dirigé par Monsieur le Professeur Jean Protas; qu'il soit remercié pour son accueil. Son successeur le Professeur Claude Lecomte (Laboratoire de Cristallographie et Modélisation des Matériaux Minéraux et Biologiques) m'a toujours encouragé dans mon travail ; je lui adresse toute ma gratitude pour avoir bien voulu examiner ce travail.

J'exprime ma respectueuse reconnaissance à Messieurs Gerd Folkers (Professeur) et Dino Moras (Directeur de Recherche) pour l'honneur qu'ils me font de juger cette thèse et ma profonde gratitude à Monsieur Jean-Paul Briand (Directeur de Recherche) pour avoir accepté d'évaluer ce travail en tant que rapporteur.

Je tiens à remercier Otto Dideberg (Directeur de Recherche) d'avoir accepté d'être rapporteur de mon travail et également de la bienveillance qu'il m'a manifestée à chacune de mes visites au cours desquelles il m'a toujours prodigué des conseils précieux et éclairés. Son hospitalité et sa gentillesse m'ont été des plus précieuses, je ne saurais l'oublier.

J'exprime mes remerciements les plus vifs à Monsieur le Professeur Guy Branlant qui, après m'avoir apporté de nombreuses remarques pertinentes, a accepté de juger ce travail.

André Aubry, responsable du groupe biocristallographie, a dirigé avec compréhension mon travail, m'aidant de conseils toujours judicieux, tout en me laissant une liberté de mouvement que j'ai appréciée. Je le remercie pour l'encadrement de qualité dont il m'a fait profiter, sa grande disponibilité et sa générosité.

Je tiens à remercier tout spécialement Bernard Vitoux pour ses multiples encouragements et discussions approfondies. Qu'il trouve ici l'expression de ma plus profonde gratitude.

Que tous les chercheurs et techniciens du LEGG de l'UHP Nancy I, dont la collaboration m'a été d'un grand secours, trouvent ici l'expression de ma reconnaissance.

Je remercie aussi toute l'équipe du LCM de l'IBS de Grenoble qui m'a apporté aide et amitié. J'exprime en particulier ma reconnaissance à Emile Duée qui à eu la gentillesse de me donner de précieux conseils dans l'élaboration de ce mémoire.

Enfin c'est avec joie que je remercie tous mes camarades de laboratoire, chercheurs et techniciens, pour l'esprit d'entraide et l'amitié dont ils ont su m'entourer.



TABLE DES MATIERES

TABLE DES MATIERES



INTRODUCTION

1

PARTIE A : ETUDE BIBLIOGRAPHIQUE

AI DESCRIPTION STRUCTURALE DES DOMAINES DE FIXATION DES

DINUCLEOTIDES NAD ET NADP DANS LES ENZYMES..... 3

A Ia Présentation des dinucléotides NAD/NADP..... 3

A Ia1 Introduction 3

A Ia2 Rôle physiologique..... 4

A Ib Repliement caractéristique des déshydrogénases 4

A Ib1 Topologie du repliement caractéristique 4

Introduction 4

Motif du 'Rossmann fold' 5

A Ib2 Comparaison des 'Rossmann fold' 6

Structures primaires..... 6

Topologie 7

A Ib3 Conformation globale du NAD(P) lié..... 8

A Ib4 Disposition du groupement pyrophosphate dans le site 8

A Ib5 Disposition de l'adénosine dans le site 9

Le cycle adénine..... 9

Le ribose adjacent au cycle adénine 11

A Ib6 Disposition du nicotinamide dans le site..... 11

Le cycle du nicotinamide 11

Le ribose adjacent au cycle du nicotinamide 12

AIc	<u>Les autres repliements fixant les dinucléotides à nicotinamide</u>	12
AIc1	Structures α/β planes	12
	<i>Cas de la dihydrofolate réductase</i>	12
	<i>Cas de l'isocitrate déshydrogénase et de l'isopropylmalate déshydrogénase</i>	13
	<i>Remarques sur les autres structures α/β planes fixant les dinucléotides</i>	13
AIc2	Structures en tonneau α/β	14
	<i>Cas des aldo-céto réductases</i>	14
	<i>Cas de l'enzyme 'old yellow'</i>	16
AIc3	Structures en tonneau β	16
AII	<u>LA SPECIFICITE POUR LES COFACTEURS NAD OU NADP</u>	16
AIIa	<u>Introduction</u>	16
AIIb	<u>Spécificité de cofacteur dans la glutathione réductase</u>	17
	AIIb1 Choix des mutations (Scrutton <i>et al.</i> , 1990)	17
	AIIb2 Résultats structuraux (Mittl <i>et al.</i> , 1994)	18
AIIc	<u>Spécificité de cofacteur dans l'isocitrate déshydrogénase</u>	20
	AIIc1 Choix des mutations (Chen <i>et al.</i> , 1995)	20
	AIIc2 Résultats structuraux (Hurley <i>et al.</i> , 1996)	20
AIId	<u>Dualité de cofacteur dans la dihydrodipicolinate réductase</u>	22
	Résultats structuraux	22
AIIf	<u>Conclusion</u>	23
AIII	<u>POURQUOI UTILISER LA GAPDH POUR FAIRE CETTE ETUDE ?</u>	25
AIV	<u>PRESENTATION STRUCTURALE DE LA GAPDH</u>	25
AIVa	<u>Les différentes classes de GAPDH</u>	25
	AIVa1 introduction	25
	AIVa2 La GAPDH glycolytique	26
	AIVa3 La GAPDH chloroplastique	27

AIVb <u>Structure de la GAPDH glycolytique</u>	27
AIVb1 Introduction.....	27
AIVb2 Les GAPDH cristallisées.....	30
AIVb3 Arrangement du tétramère dans le cristal	33
AIVb4 Relations de symétrie dans le tétramère.....	36
AIVb5 Description des interfaces entre les sous-unités	37
AIVb6 Structure du monomère.....	38
<i>Structure générale</i>	38
<i>Description du site NAD</i>	38
<i>Le site actif</i>	42
AIVb7 Mécanisme réactionnel.....	43
<i>Introduction</i>	43
<i>Intermédiaire hémithioacétal</i>	44
<i>Intermédiaire thioacyl et phosphorylation</i>	45
AIVb8 Transition apo-holo	48
<i>Introduction</i>	48
<i>Transition apo-holo dans les GAPDH de E. coli et de</i> <i>B. Stearothermophilus</i>	48
<i>Transition apo-holo dans la GAPDH de homard américain</i>	50
AV <u>SPECIFICITE DE COFACTEUR NAD/NADP DANS LA GAPDH</u>	50
AVa <u>Première approche : double mutant S</u>	50
AVa1 introduction	50
AVa2 Résultats enzymatiques	51
AVa3 Résultats structuraux.....	52
AVa4 Perspectives.....	53
AVb <u>Deuxième approche : triples mutants et quintuple mutant</u>	53
AVb1 Résultats enzymatiques sur les triples mutants	53
AVb2 Résultats enzymatiques sur le quintuple mutant	53
AVc <u>Objectifs du travail</u>	54

PARTIE B : MATERIELS ET METHODES

BI <u>PRODUCTION ET PURIFICATION</u>	55
BII <u>CRISTALLISATION</u>	56
BIIa <u>Principe</u>	56
BIIb <u>La diffusion de vapeur</u>	58
BIIc <u>Macrosemis</u>	58
BIId <u>Cristallisation en gel</u>	59
BIII <u>ACQUISITION ET INTEGRATION DES DONNEES DE DIFFRACTION</u>	59
BIIIa <u>Introduction</u>	59
BIIIb <u>Sources de rayons X</u>	60
BIIIc <u>Monochromateurs</u>	61
BIIId <u>Détecteurs</u>	62
BIIId1 <u>Système plaque image</u>	62
BIIId2 <u>Système FAST (ENRAF NONIUS)</u>	64
BIIIe <u>Traitement des données</u>	65
BIIIe1 <u>Indexation</u>	65
BIIIe2 <u>Intégration</u>	66
<i>Méthode de la 'shoe box'</i>	66
<i>Méthode du 'profile fitting'</i>	66
BIV <u>RECHERCHE DES PHASES</u>	67
BIVa <u>Remplacement moléculaire</u>	67
BIVa1 <u>Recherche de la rotation</u>	68
BIVa2 <u>Recherche de la translation</u>	70
BV <u>AFFINEMENT DU MODELE</u>	71
BV1 <u>Affinement en corps rigide</u>	72
BV2 <u>Affinement par gradient conjugué</u>	73
BV3 <u>Dynamique moléculaire</u>	74

PARTIE C : RESOLUTION DES STRUCTURES

CI <u>CRISTALLISATION</u>	76
CIa <u>Introduction</u>	76
CIb <u>Préparation de la protéine</u>	76
CIc <u>Forme monoclinique</u>	77
CId <u>Forme orthorhombique</u>	78
CII <u>CONDITIONS DE MESURE</u>	80
CIIa <u>Triple mutant cocristallisé avec NAD</u>	80
CIIa1 Enregistrement.....	80
CIIa2 Intégration.....	80
CIIb <u>Triple mutant cocristallisé avec NADP</u>	81
CIIb1 Enregistrement.....	81
CIIb2 Intégration.....	81
CIIc <u>Quintuple mutant cocristallisé avec NAD</u>	82
CIId <u>Quintuple mutant cocristallisé avec NADP</u>	83
CIII <u>PHASAGE</u>	84
CIIIa <u>D32G-Sn, D32G-Sp, B-Sn</u>	84
CIIIb <u>B-Sp</u>	84
CIIIb1 Remplacement moléculaire.....	84
CIIIb2 Empilement cristallin.....	85
CIV <u>AFFINEMENT DES MODELES</u>	88
CV <u>QUALITE DES MODELES</u>	94
CVa <u>Densité électronique</u>	94
CVb <u>Estimation de l'erreur sur les positions atomiques</u>	95
CVc <u>Géométrie</u>	96

PARTIE D : RESULTATS ET DISCUSSION

DI <u>Résumé de la publication</u>	99
DIa <u>Partie résultats structuraux</u>	99
DIa1 Etat général des mutants.....	99
DIa2 Les mutations D32G, B et S.....	99
DIa3 Le site du cofacteur.....	100
DIb <u>Discussion</u>	100
DIa Recherche des facteurs moléculaires jouant un rôle dans la sélectivité NAD/NADP	100
DIb2 Comparaison avec les sites NADP naturels.....	101
DIb3 Remarques sur les GAPDH chloroplastiques.....	102
DII <u>PUBLICATION</u>	103
‘A Crystallographic comparison between mutated glyceraldehyde-3-phosphate dehydrogenases from <i>Bacillus stearothermophilus</i> complexed with either NAD ⁺ or NADP ⁺ ’	103
Summary.....	104
Introduction.....	105
Results	107
<i>Quality of models</i>	107
<i>Overall behaviour of the mutants</i>	108
<i>The D32G mutation</i>	108
<i>The B mutations (L33T, T34G, D35G)</i>	110
<i>The S mutations (L187A, P188S)</i>	110
<i>The overall cofactor-binding site behaviour</i>	111
Discussion	113
<i>Molecular determinants of NAD⁺/NADP⁺ selectivity</i>	113
<i>Comparison of NAD⁺-binding features of wild-type, D32G-S and B-S GAPDHs</i>	116

<i>Comparison of the NADP⁺-binding features of D32G-S and B-S mutants....</i>	116
<i>Comparison with natural NADP-binding sites.</i>	118
<i>Remarks about the chloroplastic GAPDHs</i>	119
Materials and methods	119
<i>Preparation and crystallization.....</i>	119
<i>Data collection</i>	120
<i>Model building and refinement.....</i>	120
<i>Data deposition.....</i>	120
Acknowledgements.....	121
References	121
DIII <u>COMPARAISON STRUCTURALE DE 25 ENZYMES DE TYPE 'ROSSMANN FOLD'</u>	
.....	125
DIIIa <u>Introduction.....</u>	125
DIIIb <u>Mise en évidence de cinq familles</u>	127
DIIIc <u>Comportement du pyrophosphate dans les cinq familles.....</u>	131
DIIIc1 La famille i (G/AXG/AXXG).....	131
DIIIc2 La famille ii (GXXXGXG)	132
DIIIc3 La famille iii (GXXGXXG)	133
DIIIc4 La famille iv (GXGXXA/S)	135
DIIIc5 La famille v (GXXXXXXXXGX/A).....	135
DIIId <u>Réflexion sur la dualité de cofacteur dans la GADPH chloroplastique.....</u>	135
CONCLUSION 	138
ANNEXE 1 	141
BIBLIOGRAPHIE 	150

Abréviations

acétate s.	acétate de sodium
AMP	adénosine 5'-monophosphate
AMP 2'-phosphate	adénosine 5'-monophosphate 2'-phosphate
DTT	dithiothréitol
EDTA	acide éthylène diamine tétraacétique
FAD	flavine adénine dinucléotide
FMN	flavine mononucléotide
Hepes	4-(2-hydroxyéthyl)-piperazine-1-éthane acide sulfonique
IR	isomorphous replacement (remplacement isomorphe)
K _D	constante de dissociation
MR	molecular replacement (remplacement moléculaire)
NAD(P)	nicotinamide adénine dinucléotide (phosphate)
NAD(P) ⁺	nicotinamide adénine dinucléotide (phosphate), forme oxydée
NAD(P)H	nicotinamide adénine dinucléotide (phosphate), forme réduite
NCS	non crystallographic symmetry (symétrie non cristallographique)
NMN	nicotinamide mononucléotide
PEG	polyéthylène glycol
PDB	Protein Data Bank (banque de données des structures de protéines de Brookhaven)
Pipes	piperazine N,N'-bis-2-éthanesulfonate
RMN	résonance magnétique nucléaire
rms	root mean square (écart quadratique moyen)
SA	sulfate d'ammonium
Tris	N-tris(hydroxyméthyl)aminométhane
3 α ,20 β HSD	3 α ,20 β -hydroxystéroïde déshydrogénase
6PGDH	6-phosphogluconate déshydrogénase
7 α HSD	7 α -hydroxystéroïde déshydrogénase
17 β HSD	17 β -hydroxystéroïde déshydrogénase
ADH	alcool déshydrogénase
DHFR	dihydrofolate réductase
DHPiR	dihydrodipicolinate réductase
DHPtR	dihydroptéridine réductase

DLDH	D-lactate déshydrogénase
DPGDH	D-phosphoglycérate déshydrogénase
ENR	enoyl acyl carrier protéine réductase
FDH	formate déshydrogénase
GALE	UDP-galactose-4-épimérase
GAPDH	glycéraldéhyde-3-phosphate déshydrogénase
GR	glutathione réductase
ICDH	isocitrate déshydrogénase
InhA	protéine InhA
LADH	alcool déshydrogénase de foie
LDH	lactate déshydrogénase
LHicDH	L-2-hydroxyisocaproate déshydrogénase
LipDH	lipoamide déshydrogénase
MDH	malate déshydrogénase
MLCR	carbonyle réductase
NPX	NADH peroxidase
QOR	quinone oxydoréductase
ThR	thiorédoxine réductase
TR	trypanothioneréductase

INTRODUCTION



La fonction biologique de très nombreuses enzymes appartenant à la famille des oxydoréductases implique une interaction spécifique avec un dinucléotide à pyridine. Les oxydoréductases utilisant le nicotinamide adénine dinucléotide (NAD) ne sont en général pas impliquées dans les mêmes voies métaboliques que celles utilisant le nicotinamide adénine dinucléotide phosphate (NADP). Le coenzyme NAD intervient le plus souvent dans les voies cataboliques alors que le NADP intervient dans les voies anaboliques. Depuis une dizaine d'années, de nombreuses équipes de recherche ont tenté de comprendre pourquoi deux enzymes structuralement proches présentaient des spécificités de cofacteur différentes. L'équipe du Pr Branlant s'est donné comme objectif de travailler sur ce sujet en prenant comme enzyme modèle la glycéraldéhyde-3-phosphate déshydrogénase (GAPDH). En effet, cette enzyme possède la particularité d'utiliser le NAD ou le NADP en fonction de sa localisation subcellulaire : la GAPDH glycolytique, présente dans le cytoplasme, reconnaît spécifiquement le NAD, tandis que la GAPDH chloroplastique des végétaux supérieurs présente une dualité de cofacteur avec une préférence pour NADP.

Dans son travail de thèse, S. Rahuel-Clermont (1994) a construit et caractérisé une série de mutants de la GAPDH de *Bacillus stearothermophilus* dans le but de déterminer les facteurs moléculaires responsables de la spécificité NAD/NADP dans les GAPDH. A partir de ces résultats deux mutants ont été retenus pour notre étude cristallographique : le triple mutant D32G, L187A et P188S présentant une inversion de spécificité de NAD vers NADP par rapport à l'holoenzyme de type sauvage et le quintuple mutant L33T, T34G, D35G, L187A, P188S conçu pour mimer le site de fixation du cofacteur de la GAPDH chloroplastique. Pour chacun des mutants, les structures des complexes mutant-NAD et mutant-NADP ont été résolues, afin de comparer les structures d'une même enzyme complexée avec deux cofacteurs différents.

La première partie de ce mémoire regroupe : une description succincte des enzymes fixant les dinucléotides à pyridine ; un résumé des informations structurales obtenues sur la spécificité de cofacteur dans les oxydoréductases ; une revue bibliographique des résultats

cristallographiques obtenus sur les GAPDH et une présentation des travaux réalisés par l'équipe du Pr Branlant sur la spécificité de cofacteur dans la GAPDH.

Dans la seconde partie, nous exposons les différentes méthodes utilisées au cours de la détermination structurale des mutants de la GAPDH.

Les résultats concernant la cristallisation des mutants, l'obtention des intensités diffractées, le traitement des données, le phasage des facteurs de structure et l'affinement des modèles sont développés dans la partie C.

Le chapitre D débute par un bref résumé de la publication intitulée 'A crystallographic comparison between mutated glyceraldehyde-3-phosphate dehydrogenases from *Bacillus stearothermophilus* complexed with either NAD⁺ or NADP⁺'. La publication insérée présente une description structurale des quatre complexes étudiés en s'attachant à mettre évidence les relations structure-affinité de la GAPDH envers le cofacteur. Le travail publié est complété par une comparaison structurale de 25 enzymes de type 'Rossmann fold' soit NAD-dépendantes, soit NADP-dépendantes, soit à dualité de cofacteur.

PARTIE A

ETUDE BIBLIOGRAPHIQUE

AI DESCRIPTION STRUCTURALE DES DOMAINES DE FIXATION DES DINUCLEOTIDES NAD ET NADP DANS LES ENZYMES

A1a Présentation des dinucléotides NAD/NADP

A1a1 Introduction

La première structure publiée d'une protéine cocrystallisée avec un dinucléotide est celle de la lactate déshydrogénase (LDH) (Adams *et al.*, 1970). Depuis, environ cinquante structures de protéines à cofacteur dinucléotidique du type nicotinamide adénine dinucléotide (NAD) ou du type nicotinamide adénine dinucléotide phosphate (NADP) ont été résolues. La majorité de ces macromolécules sont impliquées dans des réactions d'oxydoréduction où le dinucléotide joue le rôle d'accepteur ou de donneur d'électrons (Stryer, 1995). Les structures chimiques du NAD et du NADP sont très proches : le NADP présente simplement un groupement phosphate supplémentaire fixé au groupe hydroxyle en 2' du ribose porteur de l'adénine (Figure A1). Malgré cette faible différence chimique, les enzymes à dinucléotide ne peuvent en général catalyser leur réaction qu'avec le cofacteur dont elles sont spécifiques (NAD ou NADP). Cependant, il existe des enzymes capables de fonctionner avec les deux cofacteurs, mais avec une légère préférence pour l'un ou pour l'autre.

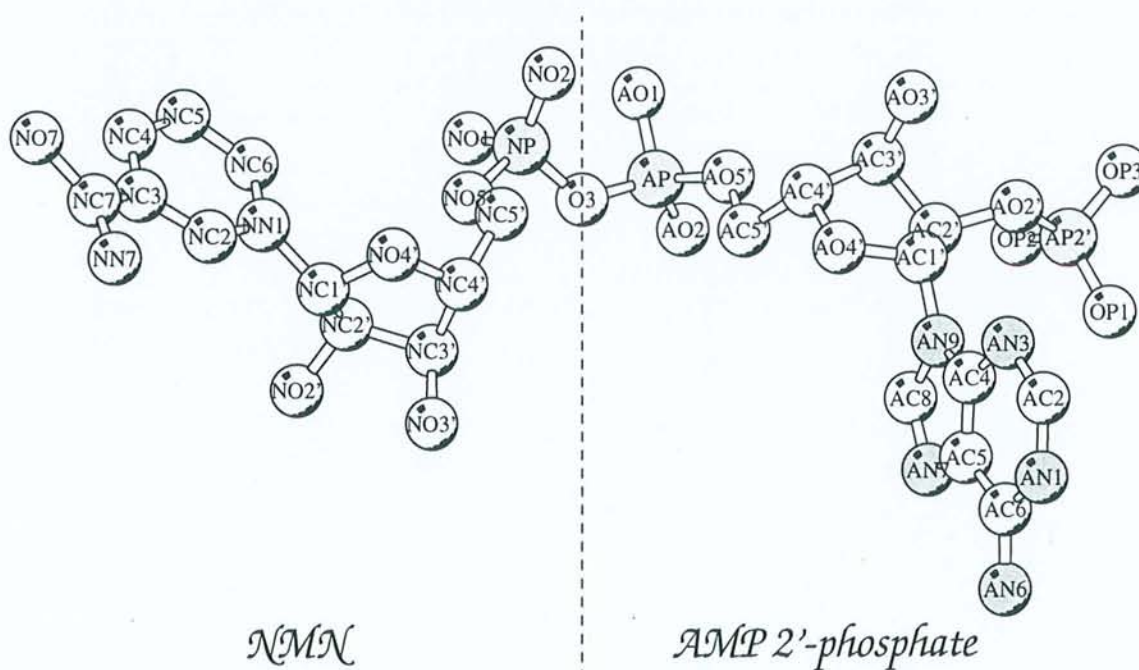


Figure A1 : Structure et nomenclature atomique du dinucléotide NADP.

La partie nicotinamide mononucléotide (NMN) et la partie adénosine 5'-monophosphate 2'-phosphate (AMP 2'-phosphate) sont représentées respectivement à gauche et à droite.

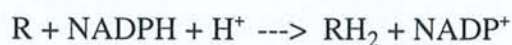
A1a2 Rôle physiologique

Les coenzymes pyridiniques prennent part à des réactions de type sensiblement différents (Stryer, 1995) :

Le coenzyme NAD est généralement impliqué dans des systèmes oxydant des substrats afin d'en retirer de l'énergie (voie catabolique). Dans l'oxydation, le cycle pyridinium du NAD^+ accepte un ion hydrure suivant la réaction:



Le coenzyme NADP participe sous forme réduite à des réactions de synthèse (voie anabolique). Dans le sens de l'oxydation de ce coenzyme, l'équation s'écrit sous la forme:



A1b Repliement caractéristique des déshydrogénases

A1b1 Topologie du repliement caractéristique

Introduction

A la fin de l'année 1973, quatre structures de déshydrogénases fixant le NAD avaient déjà été résolues : la lactate déshydrogénase (LDH), la malate déshydrogénase (MDH), l'alcool déshydrogénase (ADH) et la glycéraldéhyde-3-phosphate déshydrogénase (GAPDH). Ces quatre enzymes possèdent entre elles des différences notables : la LDH et la GAPDH sont actives sous forme d'homotétramère alors que la MDH et la ADH le sont sous forme d'homodimère. Les structures primaires de la LDH et de la MDH présentent de fortes similitudes alors qu'aucune homologie de séquence significative n'est détectée entre la LDH, la ADH et la GAPDH. Cependant quelques points communs peuvent être mis en évidence : les chaînes polypeptidiques de chaque monomère comprennent environ 350 résidus, les structures tertiaires sont composées de deux domaines différents ayant des fonctions distinctes : l'un fixe le cofacteur NAD ; le deuxième fixe le substrat et fournit les acides aminés nécessaires à la réaction catalytique. A partir des informations structurales de ces quatre enzymes, Rossmann *et al.* (1974) ont mis en évidence une topologie commune dans les trois domaines de fixation du cofacteur ayant des structures primaires différentes (figures A2, A3). Il s'est avéré par la suite

que plus de 50% des structures tridimensionnelles de protéine à cofacteur dinucléotidique présentent cette topologie appelée communément 'Rossmann-fold'.

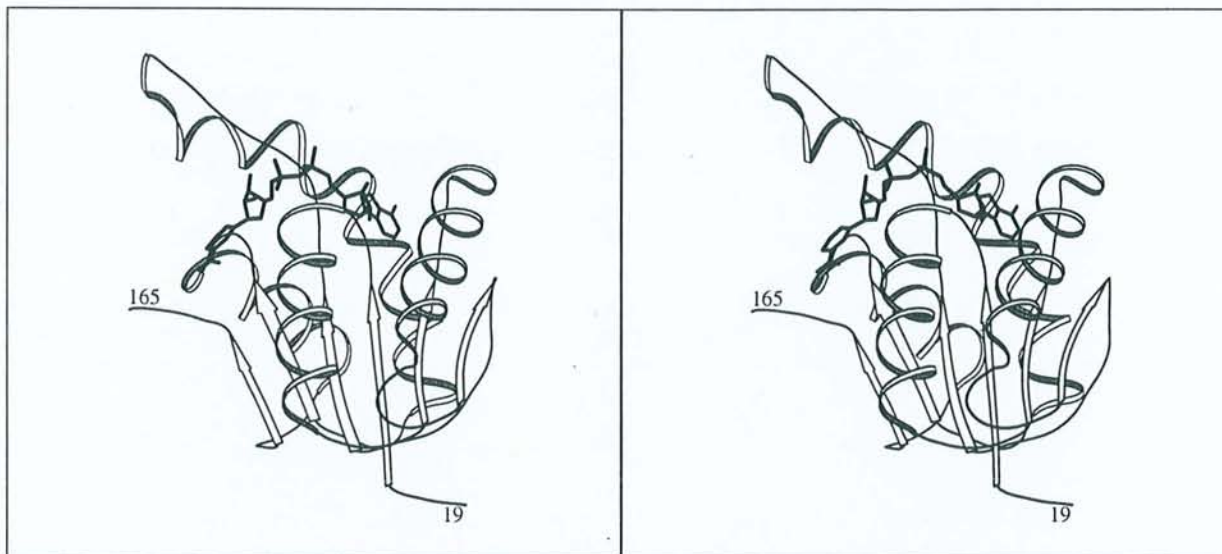


Figure A2 : Site de fixation du cofacteur de la Lactate Déshydrogénase.

Le dessin a été réalisé à partir des coordonnées atomiques du fichier pdb 3ldh (Adams *et al.*, 1970). Les structures secondaires sont représentées par les rubans (hélices α) et par les flèches (brins β), reliés par les lignes représentant les boucles.

Motif du 'Rossmann fold'

Ce repliement caractéristique contient deux unités identiques dont le motif se compose d'un feuillet β parallèle de 3 brins et de deux hélices α intercalées entre les brins du feuillet ($\beta\alpha\beta\alpha\beta$). Les deux hélices α sont disposées du même côté du feuillet. Chaque unité joue le rôle de domaine de fixation d'un mononucléotide du NAD. Le motif de l'ensemble du 'Rossmann fold' se présente sous la forme d'un feuillet β parallèle à 6 brins où les premiers brins de chaque unité sont voisins (figure A3). Ainsi les hélices α de chaque unité sont situées de part et d'autre du feuillet β . La nomenclature des brins β de la première unité est réalisée avec les lettres de A à C et celle de la deuxième avec les lettres de D à F. Les brins dans le feuillet

sont disposés suivant le motif $\beta F \beta E \beta D \beta A \beta B \beta C$. Ce feuillet β parallèle présente une torsion d'environ 100° entre les brins βF et βC . Les hélices α intercalées entre les brins βA - βB , βB - βC , βD - βE , βE - βF sont nomencaturées respectivement αB , αC , αE et αF (figure A3). La partie $\beta F \alpha F \beta E \alpha E \beta D$ constitue le site de fixation du nicotinamide mononucléotide (NMN) et la partie $\beta A \alpha B \beta B \alpha C \beta C$ celui de l'adénosine 5'-monophosphate (AMP). Cet arrangement implique la présence d'une boucle assez grande entre les brins βC et βD qui participe en général à la fixation du cycle de l'adénine.

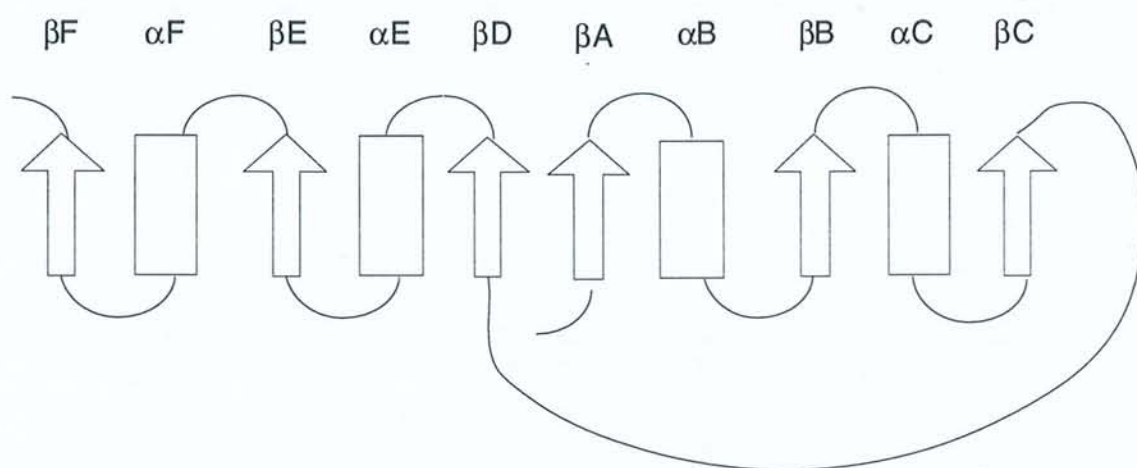


Figure A3 : Schéma de principe du repliement de Rossmann montrant l'enchaînement des hélices α (rectangles) et des brins β (flèches).

A1b2 Comparaison des 'Rossmann fold'

Structures primaires

Les structures primaires des domaines de fixation à NAD(P) ne présentent entre elles que très peu d'homologie de séquence. Il a été possible cependant de mettre en évidence trois points communs dans les séquences peptidiques des 'Rossmann fold' (Bork & Grunwald, 1990 ; Baker *et al.* 1992b) :

i) la boucle $\beta A-\alpha B$ et la partie N-terminale de l'hélice αB sont riches en résidus glycine permettant au pyrophosphate du NAD(P) de s'approcher très près de la partie N-terminale de l'hélice αB (cf § AIIb4 et AII) ;

ii) les séquences des brins βA et βB possèdent beaucoup de résidus hydrophobes ;

iii) la position C-terminale du brin βB est généralement occupée par un résidu acide dans les enzymes NAD spécifiques et la séquence peptidique de la boucle $\beta B-\alpha C$ contient le plus souvent une arginine dans celles qui sont NADP spécifiques (cf § AIIb5 et AII).

Topologie

Les structures secondaires les mieux conservées dans les enzymes dites à 'Rossmann fold' sont : les brins βA , βB , βC et βD ainsi que l'hélice αB (figure A3). Le feuillet β parallèle à 6 brins observé dans l'ADH, la MDH et la LDH est soit amputé d'un ou deux brins (par exemple le brin βE dans la thiorédoxine réductase (Waksman *et al.*, 1994) ; les brins βE et βF dans la glutathione réductase (GR ; Karplus & Schulz, 1987) et la lipoamide déshydrogénase (LipDH; Mattevi *et al.*, 1992) ; ...), soit étendu d'un ou plusieurs brins du côté du brin βF (par exemple le brin βG parallèle à βF dans la carbonyl réductase (MLCR ; Tanaka *et al.*, 1996) ou le brin βG antiparallèle à βF dans la formate déshydrogénase (FDH ; Lamzin *et al.*, 1994)). Cette extension du feuillet β ne suit pas obligatoirement directement le 'Rossmann fold' dans l'enchaînement des structures secondaires. Par exemple, dans la dihydrodipicolinate réductase (DHPiR ; Scapin *et al.*, 1995) le domaine catalytique s'intercale entre le feuillet à 6 brins et son extension constituée d'un brin β et d'une hélice α . De nombreuses variations sont observées dans la composition et le nombre des hélices α : la plupart des repliements de type 'Rossmann fold' présentent une hélice α supplémentaire (notée αD) entre les brins βC et βD ; certains sont

dépourvus de quelques hélices comme l'hélice α C dans la dihydroptéridine réductase (DHPtR ; Su *et al.*, 1993) et dans la DHPiR, ou comme les hélices α E et α F dans la GR et dans la LipDH.

AIb3 Conformation globale du NAD(P) lié

Le cofacteur NAD(P) se fixe le long des extrémités C-terminales des brins du feuillet β parallèle. Les deux nucléosides du NAD(P) se trouvent de part et d'autre du prolongement imaginaire des brins β (figure A2). Ainsi, le pyrophosphate se fixe à proximité des extrémités des éléments centraux du 'Rossmann fold' : les parties C-terminales des brins β A et β D et la partie N-terminale de l'hélice α B. A l'état lié, le NAD(P) se trouve dans une conformation étendue, où les cycles adénine et nicotinamide sont distants d'environ 14 Å (figure A1). Les riboses adoptent une conformation C2'-endo ou C3'-endo selon les enzymes (nomenclature IUPAC-IUB, 1983). Le cycle adénine se trouve en général dans une disposition anti par rapport au ribose adjacent (angle de torsion AO4'-AC1'-AN9-AC4 = $180^\circ \pm 90^\circ$). On note une exception dans la quinone oxydoréductase (QOR ; Thorn *et al.*, 1995) où l'adénine du NADP se trouve dans une disposition syn. Par contre, la disposition soit syn soit anti du cycle nicotinamide définie par l'angle de torsion NO4'-NC1'-NN1-NC2 est très variable d'une enzyme à l'autre. Elle est en relation directe avec la stéréospécificité du transfert d'hydrure qui se réalise du substrat vers le carbone NC4 du nicotinamidium (You, 1982).

AIb4 Disposition du groupement pyrophosphate dans le site

Le groupement pyrophosphate se fixe à proximité de la boucle β A- α B. Cette boucle, avec quelques résidus de la partie N-terminale de l'hélice α B, forme une séquence caractéristique d'environ 6 acides aminés dont trois sont assez bien conservés. Ces trois résidus sont en général des Glycines ou des résidus à courte chaîne latérale. L'un d'eux occupe la première place de la séquence permettant à la chaîne peptidique de s'orienter et d'amorcer la

boucle $\beta A-\alpha B$ (position P1 d'après Baker *et al.* 1992b). Le deuxième n'a pas de position fixe dans la séquence caractéristique (position P3 d'après Baker *et al.* 1992b). Son faible encombrement stérique facilite la formation d'interactions entre le pyrophosphate et la chaîne principale de l'enzyme. Le troisième occupant la dernière position du motif (position P6 d'après Baker *et al.* 1992b) optimise les contacts du brin βA avec l'hélice αB . Cette hélice, conservée dans tous les repliements de type 'Rossmann fold', joue un rôle très important dans la stabilisation du groupement pyrophosphate (Hol *et al.*, 1978 ; Wierenga *et al.*, 1985). En effet, son dipôle permanent dont le pôle positif se trouve du côté N-terminal de l'hélice, compense en partie les charges négatives portées par le groupement. Le nombre d'interactions observées entre les oxygènes du pyrophosphate et la protéine est très variable d'une enzyme à l'autre. Dans la majorité des cas, comme la boucle $\beta A-\alpha B$ se trouve à proximité du pyrophosphate, les sites NH amides de la chaîne principale de cette boucle sont impliquées dans des liaisons hydrogène avec le cofacteur (OP2A --- R10N et OP1N --- I11N dans la GAPDH (Skarzynski *et al.*, 1987) ; OP2A --- R201N, OP2A --- R201N, et OP1N --- I202N dans la FDH (Lamzin *et al.*, 1994) ; OP2N --- I19N dans la MLCR (Tanaka *et al.*, 1996) ; ...). Dans certains sites, une charge positive provenant d'un résidu basique a été observée dans le voisinage de la partie centrale du NAD(P). Celle-ci permet de renforcer la stabilisation des deux phosphates (R201 dans la FDH (Lamzin *et al.*, 1994) ; K84 dans l'UDP-galactose 4-épimérase (Thoden *et al.*, 1996)). Cependant, certaines structures n'ont révélé aucune interaction directe entre la protéine et le groupement pyrophosphate (LipDH ; Mattevi *et al.*, 1992) ; 6-phosphogluconate déshydrogénase (6PGDH ; Adams *et al.*, 1994) ; ...). Finalement, l'environnement hydrophile du groupement pyrophosphate permet la fixation de molécules d'eau structurées qui dans la plupart des cas servent de ponts avec la protéine.

A1b5 Disposition de l'adénosine dans le site

Le cycle adénine

Le cycle adénine se fixe dans une poche hydrophobe dont la profondeur est limitée par la boucle $\beta C-\beta D$ reliant le premier motif $\beta \alpha \beta \alpha \beta$ du 'Rossmann fold' au deuxième motif. Les

atomes de la chaîne principale de cette boucle servent dans certains cas de points d'ancrage au cycle adénine (N6A --- R77O dans la GAPDH (Skarzynski *et al.*, 1987) ; N1A --- A90N dans l'énolyl réductase (ENR ; Rafferty *et al.*, 1995)). En prenant comme référence la figure A4, la face gauche du cycle forme le plus souvent des contacts hydrophobes avec les chaînes latérales de résidus provenant du brin βA (Val 177 dans la LipDH) et/ou de la boucle $\beta D-\alpha E$ (Val 74 et Ala 79 dans la 6PGDH). La face droite est en général stabilisée par la chaîne latérale d'un résidu provenant de la boucle $\beta B-\alpha C$. Dans les enzymes NADP dépendantes, ce sont les atomes aliphatiques $C\beta$, $C\gamma$ et $C\delta$ de l'arginine en interaction avec le groupement 2'-phosphate qui sont impliqués (Arg33 dans la 6PGDH). Dans les enzymes NAD spécifiques, c'est plutôt la chaîne latérale d'un résidu à caractère hydrophobe qui est impliqué (Ala 202 dans la LipDH, Trp52 dans l' ENR, ...).

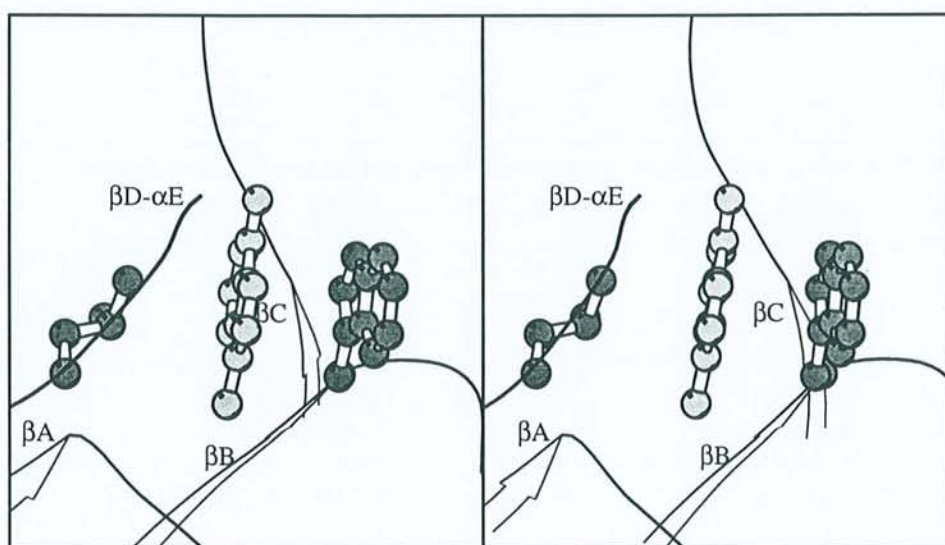


Figure A4 : Stabilisation hydrophobe du cycle adénine dans l'énolyl réductase (Rafferty *et al.*, 1995).

Les résidus hydrophobes impliqués sont un Tryptophane (Trp52) localisé dans la boucle $\beta B-\alpha C$ (côté droit) et une Leucine (L137) située dans la boucle $\beta D-\alpha E$ (côté gauche).

Le ribose adjacent au cycle adénine

Le ribose de l'adénosine est situé dans une dépression peu profonde tapissée par les boucles $\beta A-\alpha B$, $\beta B-\alpha C$ et par les parties C-terminales des brins βC et βD . Il forme le plus souvent des contacts de Van der Waals avec la boucle $\beta A-\alpha B$ (Adams *et al.*, 1994). Le 2'-phosphate est stabilisé dans la majorité des cas par une arginine localisée dans la boucle $\beta A-\alpha B$ des sites de fixation du NADP (R39 dans la DHPiR (Scapin *et al.*, 1995), R176 dans la thiorédoxine réductase (Waksman *et al.*, 1994) ; ...). Dans la plupart des enzymes à NAD, un résidu acide (Asp ou Glu) est observé au niveau C-terminal du brin βB . Les deux oxygènes du carboxylate de ce résidu caractéristique (signature P7 d'après Baker *et al.*, 1992b) interagissent avec les groupements hydroxyles 2' et 3' du ribose de l'adénosine du NAD. Les différences entre les sites de fixation du NAD et du NADP seront examinées de manière un peu plus détaillée dans le paragraphe AII intitulé 'La spécificité pour les cofacteurs NAD ou NADP'.

AIIb6 Disposition du nicotinamide dans le site

Le cycle du nicotinamide

Le deuxième motif $\beta\alpha\beta\alpha\beta$, site de fixation du NMN, est moins bien conservé que le premier dans le repliement du 'Rossmann fold', si bien que le cycle du nicotinamide ne se présente pas dans la même position d'une enzyme à l'autre. Cependant, dans toutes les structures où il a été possible de positionner avec précision le cycle nicotinamide, ce dernier se situe à proximité du site de fixation du substrat pour permettre le transfert d'hydrure. Le cycle nicotinamide est stabilisé d'une part par des résidus à caractère hydrophobe (Ile11 et Tyr317 dans la GAPDH (Skarzynski *et al.*, 1987) ; ...) et d'autre part par des acides aminés capables de former des liaisons hydrogène avec les groupes CO et NH₂ du carboxamide. Le groupement carboxamide est dévié du plan de la pyridine d'environ 30°. Lamzin *et al.* (1994) ont montré que le site de fixation du carboxamide de la FDH est modelé pour recevoir un groupement sortant du plan de la pyridine. Lorsque le nicotinamide se trouve dans une position syn par rapport au

ribose adjacent, le groupe NH_2 forme une liaison intramoléculaire avec l'un des oxygènes du pyrophosphate.

Le ribose adjacent au cycle du nicotinamide

Comme pour le site de fixation du cycle du nicotinamide, il est difficile de définir un site de fixation commun dans les oxydoréductases pour le ribose de la partie NMN du NAD(P). La majorité des protéines impliquent au moins un acide aminé dans la stabilisation du ribose. En ne comparant que les enzymes de la famille des réductases/épimérase/déshydrogénases (Labesse *et al.*, 1994) dont la structure est connue, il est intéressant de noter la conservation d'une Lys dans le site de fixation du ribose, bien que certaines ne présentent entre elles qu'un faible taux d'identité de séquence (Tanaka *et al.*, 1996 ; Rafferty *et al.*, 1995). Cet acide aminé forme des liaisons hydrogène avec les atomes $\text{NO}2'$ et $\text{NO}3'$.

A1c Les autres repliements fixant les dinucléotides à nicotinamide

A1c1 Structures α/β planes

Cas de la dihydrofolate réductase

La dihydrofolate réductase (DHFR) est une enzyme monomérique dont le repliement de la chaîne peptidique consiste en un feuillet β de huit brins (sept parallèles et un antiparallèle) et de quatre hélices α (deux de chaque côté du feuillet) (figure A5a). Sawaya & Kraut (1997) ont montré qu'il est possible de diviser la structure du monomère de la DHFR en deux sous-domaines : le plus important des deux constitue le site de fixation de la partie adénosine 2'-phosphate 5'-diphosphate du NADP ; le deuxième est appelé domaine des boucles car celles-ci jouent un rôle important dans la réaction catalytique. La partie nicotinamide du NADP est située dans une poche entre les deux sous-domaines. Il est important de remarquer que les quatre brins β et l'hélice αC présents dans le plus important des deux sous-domaines peuvent être identifiés respectivement aux quatre premiers brins et à l'hélice αB du 'Rossmann fold' (Sawaya & Kraut, 1997 ; Birktoft & Banaszak, 1984). Cependant l'ensemble du site de fixation du

cofacteur de la DHFR ne présente pas toutes les caractéristiques du 'Rossmann fold' : d'une part la séquence primaire de la DHFR ne présente pas dans la boucle β B- α C (équivalent β A- α B dans le 'Rossmann fold') les glycines caractéristiques (Mittl *et al.*, 1994) et d'autre part le cofacteur n'adopte pas la même conformation dans les deux sites (en particulier le plan moyen du ribose de l'adénosine est parallèle au feuillet β dans la DHFR alors qu'il lui est perpendiculaire dans le 'Rossmann fold') (Matthews *et al.*, 1978 ; Sawaya & Kraut, 1997).

Cas de l'isocitrate déshydrogénase et de l'isopropylmalate déshydrogénase

L'isocitrate déshydrogénase (ICDH ; Hurley *et al.*, 1991; Stoddard *et al.*, 1993) et l'isopropylmalate déshydrogénase (Hurley & Dean, 1994) sont deux enzymes homodimériques où chaque monomère, constitué d'un feuillet β mixte central entouré d'hélices α , est divisé en deux domaines. Ces deux enzymes présentent le même repliement de la chaîne polypeptidique des sites de fixation du cofacteur. L'interface des deux domaines et deux brins antiparallèles du feuillet β forment une poche fixant le coenzyme (figure A5b). La conformation du cofacteur lié est très différente de celle du NAD(P) fixé dans un 'Rossmann fold'. Les plus grandes différences concernent la partie nicotinamide avec en particulier le puckering du ribose adjacent au nicotinamide (4'endo dans l'ICDH ; 2'endo ou 3'endo dans un 'Rossmann fold').

Remarques sur les autres structures α/β planes fixant les dinucléotides

La structure du domaine catalytique des toxines ADP-ribosylantes utilisant le NAD comme substrat est constituée de deux feuillets β antiparallèles et de deux hélices α . La résolution de la structure de la toxine diphtérique cocrystallisée avec son substrat a permis de situer précisément le site de fixation du NAD (Bell & Eisenberg, 1996). Il se trouve dans une crevasse délimitée par les deux feuillets β . Le NAD n'adopte pas la même conformation dans la toxine diphtérique que dans une oxydoréductase à 'Rossmann fold'. L'angle dièdre γ (AO5'-AC5'-AC4'-AC3') de 95° donne une conformation ouverte, inhabituelle, à la partie AMP du NAD.

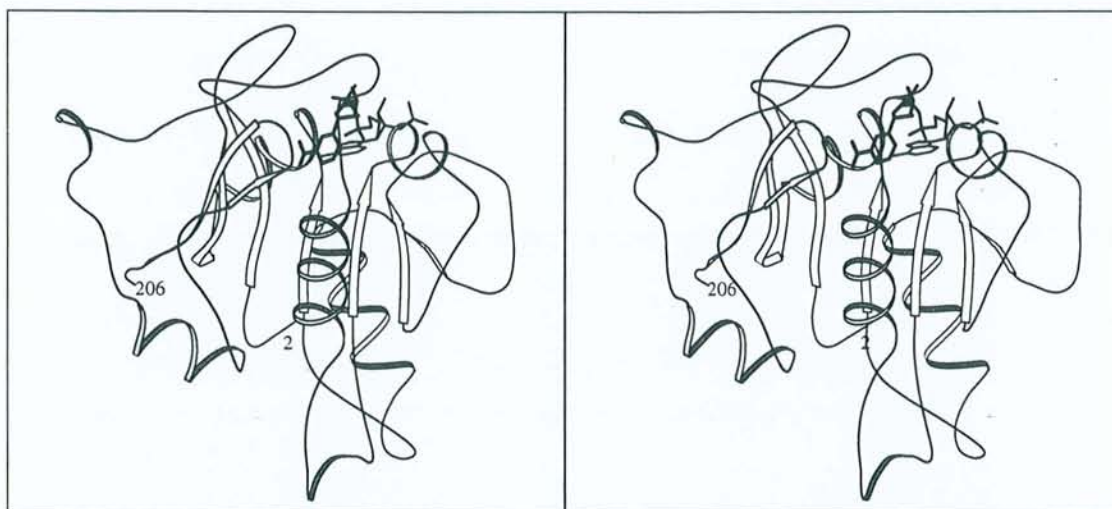
La NADH oxydase est une enzyme dimérique qui utilise NAD ou NADP comme substrat et FAD (flavine adénine dinucléotide) ou FMN (flavine adénine mononucléotide) comme cofacteur. Hecht *et al.* (1995) ont résolu deux structures de la NADH oxydase complexée respectivement avec les cofacteurs FMN et FAD sans le substrat. Chaque monomère se divise en deux domaines de tailles inégales : le plus important est constitué d'un feuillet β antiparallèle de quatre brins entouré par cinq hélices α et le second principalement de deux hélices. Le substrat et le cofacteur sont localisés à l'interface des deux domaines.

La 3-hydroxy-3-méthylglutaryl-coenzyme A réductase, utilisant le NAD comme cofacteur, est une enzyme homodimérique. Sa structure a été déterminée sous forme apo par Lawrence *et al.* (1995). Chaque monomère se décompose en deux domaines où le plus petit fixe le cofacteur. Ce domaine est constitué d'un feuillet β antiparallèle de quatre brins et de deux hélices α bordant le même côté du feuillet.

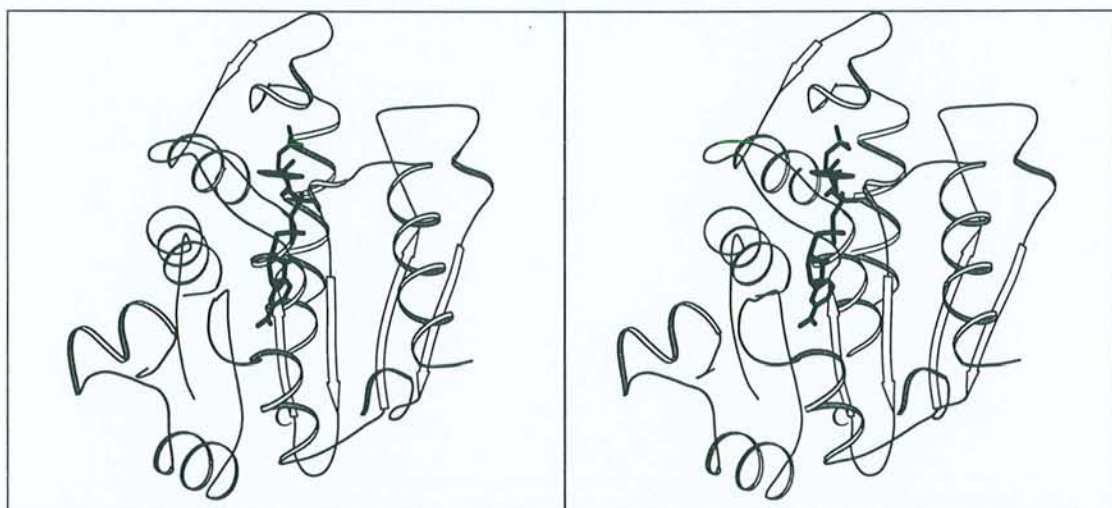
A1c2 Structures en tonneau α/β

Cas des aldo-céto réductases

La structure des enzymes de la famille des aldo-céto réductases est constituée de l'enchaînement régulier de brins β et d'hélices α alternés dont le corps se présente sous la forme d'un tonneau β parallèle de huit brins. Quatre structures d'enzyme de cette famille ont été déterminées à ce jour : l'aldose réductase (Rondeau *et al.*, 1992 ; Wilson *et al.*, 1993), l'aldéhyde réductase (El-Kabbani *et al.*, 1994 ; El-Kabbani *et al.*, 1995), la 3 α -hydroxystéroïde/dihydrodiol déshydrogénase de rat (Hoog *et al.*, 1994) et la protéine FR-1 de souris (Wilson *et al.*, 1995). Le site de fixation du dinucléotide est localisé du côté C-terminal des brins et s'étend jusqu'au centre du tonneau (figure A5c). Le cofacteur adopte une conformation étendue où le cycle adénine se positionne à la périphérie du tonneau et le nicotinamide approximativement sur l'axe du tonneau.



a) Dihydrofolate réductase (code PDB 1DYR) (Champness *et al.*, 1994)



b) Isocitrate déshydrogénase (code PDB 4ICD) (Stoddard *et al.*, 1993).

Seuls les résidus situés à une distance inférieure à 20 Å du cofacteur sont représentés.



c) Aldose réductase (code PDB 1ADS) (Wilson *et al.*, 1993).

Figure A5 : Modèle en rubans de trois enzymes fixant le NAD(P) dans des motifs différant du 'Rossmann fold'

Cas de l'enzyme 'old yellow'

L'enzyme 'old yellow' est une flavoenzyme FMN-dépendante capable de fixer le NADP. Sa structure tertiaire se présente sous la forme d'un tonneau α/β où le site de fixation du FMN est similaire au site de fixation du NADP dans les enzymes de la famille des aldo-céto réductases (Fox & Karplus, 1994). L'enzyme 'old yellow' a été cocrystallisée avec un analogue du NADP dont seule la partie nicotinamide ribose est visible dans les cartes de densité électronique. Celle-ci est située dans la poche du site actif et forme des interactions de van der Waals avec le cycle flavine du cofacteur.

A1c3 Structures en tonneau β

Les catalases de foie de boeuf et de *Proteus mirabilis* sont des enzymes tétramériques où chaque monomère est divisé en quatre domaines (Fita *et al.*, 1986 ; Gouet *et al.*, 1995). Le dinucléotide NADP est situé dans une poche à la périphérie du deuxième des quatre domaines, constitué d'un tonneau β antiparallèle à 8 brins et de cinq hélices α . La conformation adoptée par le NADP fixé est radicalement différente de celles observées dans les autres protéines : la molécule est dans une conformation repliée où la distance entre les atomes NC6 et N3A est seulement de 3.8 Å (figure A1).

AII LA SPECIFICITE POUR LES COFACTEURS NAD OU NADP

AIIa Introduction

La quantité impressionnante d'informations sur les oxydoréductases, tant d'ordre structural que d'ordre enzymatique, a permis le développement d'études sur la reconnaissance spécifique des cofacteurs NAD/NADP par ces enzymes. En effet, la plupart des oxydoréductases fixent préférentiellement l'un des deux coenzymes qui, comme nous l'avons déjà souligné, possèdent des structures chimiques très proches. Les chercheurs ont concentré

leurs efforts dans le remodelage, par mutagenèse dirigée, des domaines de fixation des dinucléotides à nicotinamide pour tenter de déterminer les facteurs moléculaires responsables de la spécificité de cofacteur NAD/NADP. La majorité des études a été réalisée sur des protéines dont le domaine de fixation du dinucléotide est un 'Rossmann fold'. Le choix judicieux des résidus à muter, guidé par l'alignement des séquences de deux enzymes structuralement proches mais ne présentant pas la même spécificité de cofacteur, a permis dans certains cas d'inverser la spécificité de cofacteur. C'est le cas de la glutathione réductase NADP-dépendante comparée à la dihydrolipoamide déshydrogénase (Scrutton *et al.*, 1990) et vice-versa (Bocanegra *et al.*, 1993), de la xylitol déshydrogénase NAD-dépendante comparée à l'alcool déshydrogénase NADP-dépendante (Metzger & Hollenberg, 1995), de la malate déshydrogénase cytosolique NAD-dépendante comparée à son homologue chloroplastique NADP-dépendante (Nishiyama *et al.*, 1993) et de la glycéraldéhyde-3-phosphate déshydrogénase cytosolique NAD-dépendante comparée à son homologue chloroplastique NADP-dépendante (Corbier *et al.*, 1990a ; Clermont *et al.*, 1993 ; Eyschen *et al.*, 1996). Il est intéressant de noter aussi le cas de l'isopropylmalate déshydrogénase NAD-dépendante comparée à l'isocitrate déshydrogénase NADP-dépendante (Miyazaki & Oshima, 1994) et vice-versa (Chen *et al.*, 1995), qui ne sont pas des enzymes à 'Rossmann fold'. Cette partie a pour but de présenter les études structurales des mutants NAD-dépendants de la glutathione réductase (Mittl *et al.*, 1994) et de l'isocitrate déshydrogénase (Hurley *et al.*, 1996), mais aussi l'étude cristallographique de la dihydrodipicolinate réductase cocrystallisée avec NAD et NADP (Scapin *et al.*, 1995 ; Reddy *et al.*, 1996).

AIIB Spécificité de cofacteur dans la glutathione réductase

AIIB1 Choix des mutations (Scrutton *et al.*, 1990)

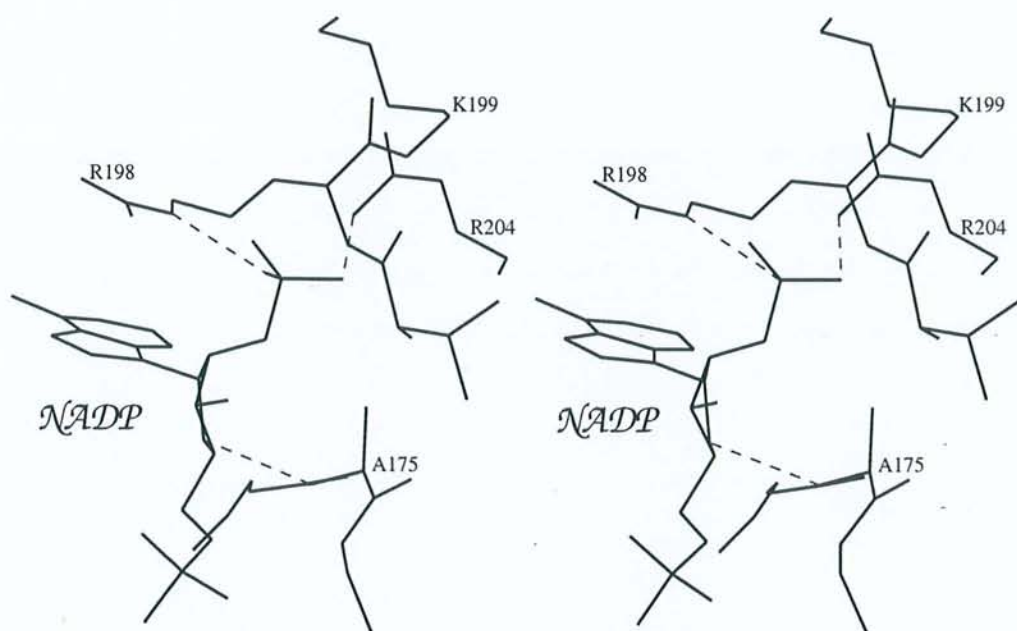
Le choix des mutations a été effectué à partir de la structure primaire de la dihydrolipoamide déshydrogénase de *E. coli* qui est NAD-dépendante. Les résidus mutés sont situés dans le domaine de fixation de la partie AMP 2'-phosphate du cofacteur. Le remplacement de l'alanine 179 (position P6 d'après Baker *et al.*, 1992b) et de l'alanine 183 par des glycines avait pour objectif, en éliminant deux groupements méthyles, de donner localement plus de liberté conformationnelle à la chaîne polypeptidique. Les mutations Val197Glu, Arg198Met,

Lys199Phe, His200Asp et Arg204Pro avaient pour but de reproduire le plus fidèlement possible l'environnement du ribose de l'adénosine du NAD observé dans la dihydrolipoamide déshydrogénase (Mattevi *et al.*, 1992). La substitution de la valine 197 en position P7 (d'après Baker *et al.* 1992b) par l'acide glutamique devait en particulier permettre la stabilisation des deux hydroxyles du ribose de l'adénosine par le carboxylate.

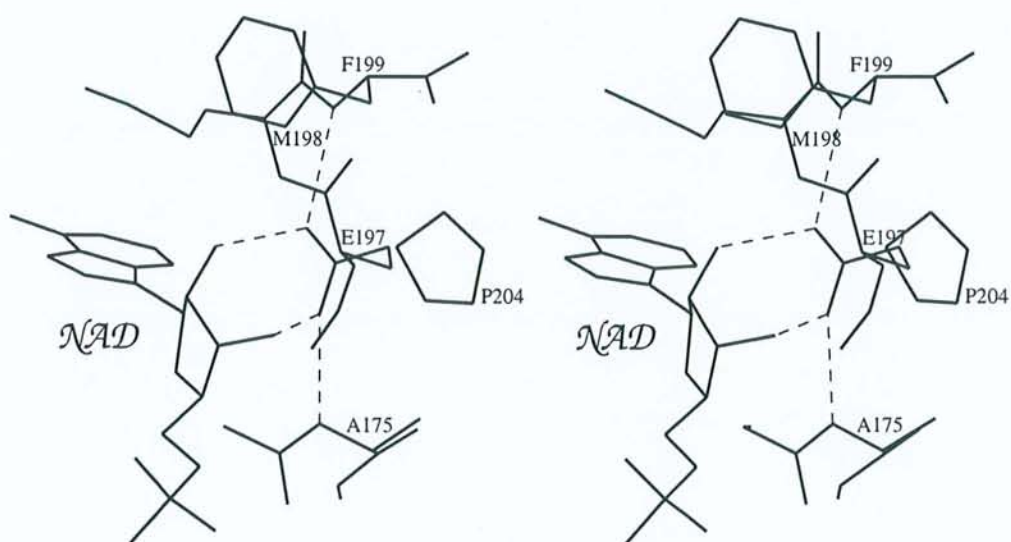
AIIb2 Résultats structuraux (Mittl *et al.*, 1994)

L'analyse structurale de la structure du septuple mutant ne révèle que des différences mineures par rapport à l'holoenzyme de type sauvage. Ces différences sont situées autour des points de mutation dans le site de fixation du cofacteur. Mittl *et al.* (1994) ont montré que les mutations Ala179Gly et Val197Glu étaient essentielles dans le changement de spécificité de cofacteur dans la glutathione réductase. Elles induisent une rotation de 90° de la liaison peptidique 174-175, modifiant ainsi les interactions de l'azote de la chaîne principale du résidu Ala175 (figure A6) : alors qu'il était impliqué dans une liaison hydrogène avec l'atome O4'A du NADP dans l'enzyme de type sauvage il forme une interaction avec le carboxylate du résidu acide 197 dans le septuple mutant.

Le changement de puckering du ribose de l'adénosine du NAD de 3'endo en 2'endo semble lié à deux phénomènes (figure A6) : d'une part il optimise la stabilisation du cofacteur par la formation de liaisons hydrogène entre le carboxylate de l'acide Glutamique 197 et les hydroxyles en 2' et 3' et, d'autre part, il évite la création de mauvais contacts entre l'oxygène AO3' et le carbone C β de l'Ala175 déplacé dans le changement conformationnel de la liaison peptidique 174-175. Les mutations Arg198Met, Lys199Phe, Arg204Pro créent un environnement non polaire autour du résidu acide 197 pour favoriser son interaction avec le cofacteur. Les deux autres mutations Ala183Gly et His200Asp ne semblent pas jouer de rôle significatif dans le changement de spécificité. D'ailleurs, la nature du résidu 200 (His dans l'enzyme de type sauvage et Asp dans le mutant) ne joue pas sur l'orientation de la chaîne latérale, qui pointe vers le solvant dans les deux cas.



forme de type sauvage



septuple mutant Ala179Gly, Ala183Gly, Val197Glu, Arg198Met, Lys199Phe, His200Asp et Arg204Pro.

Figure A6 : Sites de fixation du cofacteur de la glutathione réductase dans la forme de type sauvage complexée avec NADP (figure du haut) et dans le septuple mutant complexé avec NAD (figure du bas)

Le septuple mutant avec NAD ne mime cependant pas complètement l'enzyme de type sauvage avec NADP car son activité enzymatique est 30 fois plus faible que celle mesurée dans l'enzyme de type sauvage avec NADP. Cette différence notable peut s'expliquer par un mauvais positionnement du cycle nicotinamide dans son site de fixation, diminuant ainsi l'efficacité du transfert d'hydrure. Mittl *et al.* (1994) pensent que cette baisse d'activité peut être compensée par des effets à longue distance qui sont néanmoins très difficiles à modéliser.

AIIC Spécificité de cofacteur dans l'isocitrate déshydrogénase

AIIC1 Choix des mutations (Chen *et al.*, 1995)

Le choix des sept mutations sur l'isocitrate déshydrogénase NADP-dépendante de *E. coli* a été fait à partir de la structure primaire de l'isopropylmalate déshydrogénase NAD-dépendante de *Thermus thermophilus*. Ces deux enzymes ne possèdent que 25 % d'identité de séquence avec cependant des structures tertiaires similaires. Cinq des points de mutations (Lys344Asp, Tyr345Ile, Val351Ala, Tyr391Lys et Arg395Ser) sont situés dans le domaine de fixation de la partie AMP 2'-phosphate du cofacteur. Les deux autres points (Cys201Met et Cys332Tyr) sont distants respectivement de 38 Å et 33 Å du site de fixation de l'adénosine.

AIIC2 Résultats structuraux (Hurley *et al.*, 1996)

Hurley *et al.* (1996) classent les résidus dans trois ensembles suivant leurs effets plus ou moins directs sur le cofacteur : les résidus en interaction directe avec le coenzyme (première couche), les résidus en interaction avec les acides aminés de la première couche (deuxième couche) et les résidus éloignés du coenzyme (effet de longue distance).

La mutation K344D a deux effets de première couche : la charge négative portée par la chaîne latérale inhibe probablement la fixation du NADP en créant une répulsion électrostatique avec le groupement 2'-phosphate, alors que le carboxylate favorise la stabilisation du NAD en formant trois liaisons hydrogène avec les hydroxyles du ribose de l'adénosine. Les résidus

Tyr345 et Arg395 mutés respectivement en Ile et Ser éliminent en partie l'environnement favorable à la stabilisation du 2'-phosphate du NADP. Enfin, alors qu'il était prévisible que la mutation Tyr391Lys allait favoriser la création d'une nouvelle interaction avec le 2'-phosphate du NADP, la chaîne latérale de la lysine adopte une conformation inattendue en ne pointant pas dans la direction du site de fixation du cofacteur.

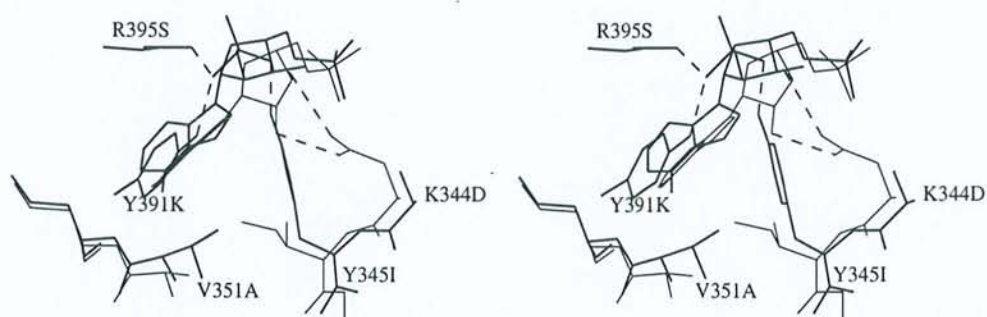


Figure A7 : Superposition du site de fixation du ribose de l'adénosine de l'isocitrate déshydrogénase de type sauvage NADP-dépendant (traits gras) avec son septuple mutant NAD-dépendant (traits fins).

La mutation Val351Ala semble jouer un rôle prédominant dans les effets de deuxième couche. L'espace libéré permet le déplacement de l'adénosine vers le résidu Asp344 dans le septuple mutant. De plus, le remplacement de la Val351, impliqué dans un amas hydrophobe dans l'enzyme de type sauvage, facilite le changement conformationnel de la chaîne principale de ce résidu. Ainsi le résidu Ile345 nouvellement introduit peut positionner correctement son carbone C δ dans le voisinage direct du site occupé par le carbone C γ 1 du résidu Val351 dans l'enzyme de type sauvage.

La structure ne permet pas d'expliquer les effets à longue distance créés par les mutations Cys201Met et Cys332Tyr. Celles-ci n'induisent aucun changement conformationnel significatif. De plus, elles ne semblent pas jouer de rôle dans la spécificité de cofacteur car les efficacités des complexes septuple mutant/NADP et septuple mutant/NAD sont respectivement 4.5 et 7.8 fois plus grandes que celles mesurées dans les complexes quintuple mutant/NADP et quintuple mutant/NAD.

AIId Dualité de cofacteur dans la dihydrodipicolinate réductase

La dihydrodipicolinate réductase est la seule enzyme présentant une dualité de cofacteur NAD/NADP à avoir été cocrystallisée avec les deux coenzymes à nicotinamide (Scapin *et al.*, 1995 ; Reddy *et al.*, 1996). La structure tertiaire de son domaine de fixation du dinucléotide est de type 'Rossmann fold'. Il est intéressant de noter que la position P7 est occupée par un résidu acide (Glu38) et la suivante par un résidu basique (Arg39).

Résultats structuraux

Les deux structures de la dihydrodipicolinate réductase cocrystallisée respectivement avec NAD et NADP sont très similaires. Les parties nicotinamides des deux coenzymes sont superposables alors que le cycle adénine est légèrement plus enfoui dans la structure complexée avec NAD. Ce changement structural est accompagné par un réarrangement conformationnel de la chaîne polypeptidique comprise entre les résidus 37 et 40, en particulier au niveau des chaînes latérales des acides aminés Glu38 et Arg39. L'acide glutamique forme deux liaisons hydrogène avec les hydroxyles du ribose de l'adénosine dans la structure cocrystallisée avec NAD alors que dans celle complexée avec NADP la chaîne latérale de ce résidu acide est légèrement plus distante de la partie adénosine, évitant ainsi une répulsion électrostatique avec le 2'-phosphate. L'arginine 39 est le seul résidu stabilisant le 2'-phosphate dans la structure cocrystallisée avec NADP. De plus, sa chaîne latérale forme avec le cycle adénine une interaction $\pi-\pi$ ('stacking') non conservée dans le complexe avec NAD.

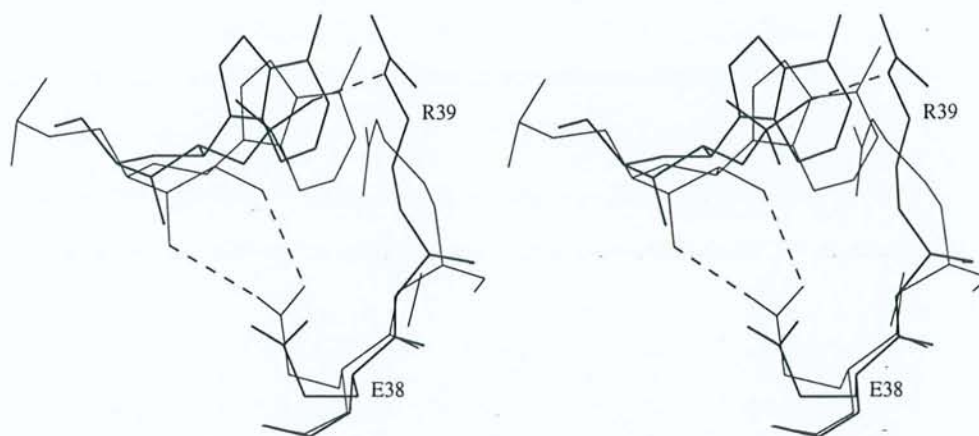


Figure A8 : Superposition du site de fixation ribose de l'adénosine des structures de la dihydronicotinamide réductase cocrystallisée respectivement avec NADP (traits gras) et NAD (traits fins).

AIIe Conclusion

Ces études structurales, ayant pour but de déterminer les facteurs moléculaires responsables de la spécificité de cofacteur NAD et NADP, permettent de confirmer certaines hypothèses concernant les effets de première couche, d'apporter des informations structurales supplémentaires sur les effets de deuxième couche et de donner de nouvelles perspectives pour les effets à longue distance.

Comme l'ont montré de nombreuses études sur les oxydoréductases présentant un 'Rossmann fold', les effets de première couche impliquent principalement deux résidus situés dans le voisinage de l'hydroxyle AO2' du ribose de l'adénosine (Scrutton *et al.*, 1990 ; Baker *et al.*, 1992b ; Clermont *et al.*, 1993 ; Eyschen *et al.*, 1996 ; Tanaka *et al.*, 1996) : un résidu acide qui est déterminant dans la stabilisation du NAD et antidéterminant dans la fixation du NADP (répulsion électrostatique) et un résidu basique intervenant dans la stabilisation du 2'-phosphate du NADP. En général, les structures primaires des enzymes strictement NAD-dépendantes possèdent le résidu acide et ne possèdent pas le résidu basique et inversement pour les enzymes strictement NADP-dépendantes. Il est intéressant de souligner le cas de la

dihydrodipicolinate réductase présentant une dualité de cofacteur NAD/NADP, car elle possède à la fois le résidu acide et le résidu basique.

Les résidus impliqués dans les effets de deuxième couche souvent partiellement anticipés apparaissent très importants dans la spécificité de cofacteur car ils optimisent le positionnement des résidus en contact avec le coenzyme (Mittl *et al.*, 1994 ; Hurley *et al.*, 1996).

Les effets à longue distance sont très difficiles à prévoir. Hurley *et al.* (1996) proposent l'utilisation d'une mutagénèse systématique pour les mettre en évidence.

AIII POURQUOI UTILISER LA GAPDH POUR FAIRE CETTE ETUDE ?

Le but de cette étude est de comprendre d'un point de vue structural la spécificité de cofacteur dans les déshydrogénases à NAD(P). La glycéraldéhyde-3-phosphate déshydrogénase (GAPDH) est un modèle intéressant car selon sa localisation subcellulaire elle est soit strictement NAD dépendante (GAPDH glycolytique), soit à dualité de cofacteur NAD(P) (GAPDH chloroplastique). Des essais de mutagenèse dirigée ont été réalisés sur la GAPDH de *B. stearothermophilus* strictement NAD dépendante dans le but de créer une enzyme capable de fixer préférentiellement le NADP. Les connaissances de la structure tridimensionnelle de cette enzyme (Skarzynski *et al.*, 1987) et des structures primaires des GAPDH chloroplastiques ont été utiles pour réaliser ce remodelage protéique. La conception et l'étude enzymatique des mutants de la GAPDH de *B. stearothermophilus* sont rassemblées dans le travail de thèse de S. Rahuel-Clermont (1994) et dans trois publications (Corbier *et al.*, 1990a; Clermont *et al.*, 1993; Eyschen *et al.*, 1996). Notre projet était de cristalliser et de résoudre la structure des mutants qui semblent donner des résultats cinétiques intéressants sur la fixation du cofacteur NADP.

AIV PRESENTATION STRUCTURALE DE LA GAPDH

AIVa Les différentes classes de GAPDH

AIVa1 introduction

Les glycéraldéhyde-3-phosphate déshydrogénases font partie de la famille des aldéhydes déshydrogénases et se répartissent en deux classes:

- La première classe rassemble les GAPDH dites phosphorylantes. Ces enzymes strictement NAD dépendantes ou à dualité de cofacteur NAD(P) catalysent réversiblement l'oxydation d'une fonction aldéhyde en acide phosphorylé, en présence de phosphate inorganique. Elle se divise en deux sous-classes se distinguant par leur fonction enzymatique et leur localisation cellulaire.

- La deuxième classe rassemble les GAPDH dites non phosphorylantes. Ces enzymes strictement NADP dépendantes catalysent irréversiblement l'oxydation d'une fonction aldéhyde en acide.

Nous nous sommes intéressés uniquement à la classe des GAPDH phosphorylantes.

AIVa2 La GAPDH glycolytique

La GAPDH glycolytique est une enzyme active sous forme d'homotétramère impliquée dans la première réaction de la dernière étape de la glycolyse (Stryer, 1995). Elle catalyse , en présence de NAD^+ et de phosphate inorganique, la phosphorylation oxydative du D-glycéraldéhyde-3-phosphate (G3P) en acide 1,3-disphosphoglycérique (1,3-dPG) (figure A9). La réaction suivante de la glycolyse utilise cet acide très énergétique pour produire l'adénosine triphosphate (ATP). La GAPDH glycolytique catalyse la réaction inverse quand elle est impliquée dans la voie métabolique de la néoglucogénèse.

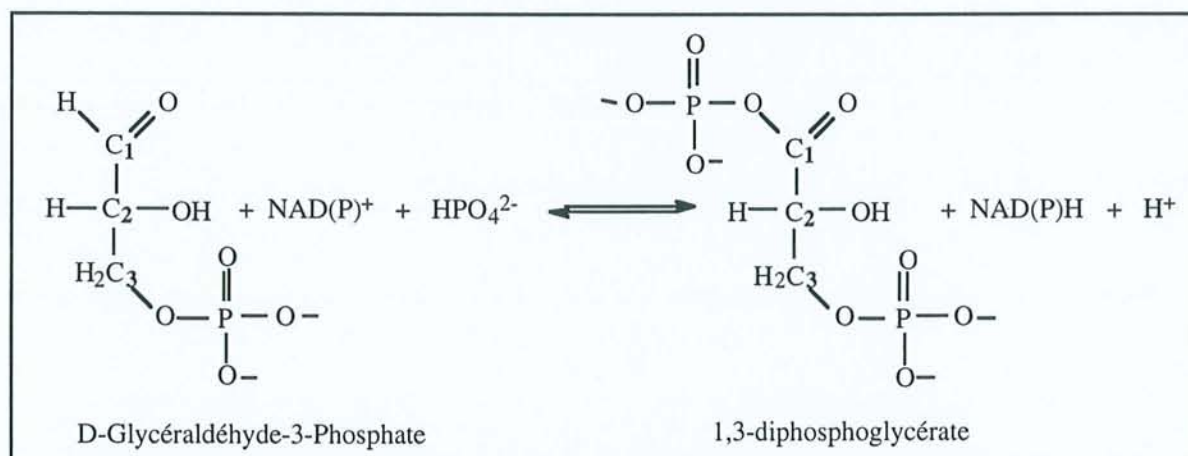


Figure A9 : Réaction catalysée par la GAPDH

La nomenclature des atomes de carbone du substrat est précisée en gras.

La GAPDH glycolytique est présente dans le cytoplasme des cellules eucaryotes, d'eubactéries et d'archaebactéries. Seules ces dernières, non halophiles, présentent une dualité de cofacteur avec une préférence pour NADP alors que les autres sont strictement NAD-

dépendantes. Les structures primaires des GAPDH glycolytiques possèdent entre elles un fort degré d'homologie (>45%) (annexe 1) à l'exception de celles d'archaebactéries non halophiles qui ne présentent que 10 à 15% d'identité avec leurs homologues eucaryotes et eubactériennes (Martin *et al.*, 1993).

AIVa3 La GAPDH chloroplastique

La GAPDH chloroplastique est impliquée dans la voie métabolique du cycle de Calvin (Stryer, 1995) lequel permet la conversion du dioxyde de carbone en glucide. Localisée dans le stroma des chloroplastes des végétaux supérieurs, la GAPDH catalyse la formation du G3P à partir du 1,3-dPG en présence de NADPH. Elle est active sous forme de tétramère contenant soit quatre sous-unités identiques (stoechiométrie A_4) soit quatre sous-unités composées de deux isoformes A et B (stoechiométrie A_2B_2) (Cerff, 1979; Ferri *et al.*, 1990). Les différentes GAPDH chloroplastiques présentent de fortes homologues de séquence avec les GAPDH glycolytiques eucaryotes et eubactériennes (env. 45%). D'un point de vue spécificité de cofacteur, elles sont différentes de leurs homologues glycolytiques car elles présentent une dualité de cofacteur NAD(P) avec une préférence pour NADP.

AIVb Structure de la GAPDH glycolytique

AIVb1 Introduction

La GAPDH glycolytique phosphorylante NAD-dépendante est une enzyme homotétramérique de 145 kD environ (figure A10.). La structure quaternaire du tétramère est définie dans un système de coordonnées orthogonales dont l'origine est située au centre du tétramère. Ce système a été défini par Rossmann *et al.* (1973) dans la lactate déshydrogénase (LDH). Les trois axes P , Q , R de ce système révèlent une symétrie 222 apparente : au monomère O , on fait respectivement correspondre par les axes P , Q , R les monomères P , Q et R . Chaque monomère, composé d'environ 330 acides aminés, comporte deux domaines : le domaine de fixation du cofacteur NAD et le domaine catalytique. Le domaine de fixation du cofacteur est composé des résidus 1 à 148 formant un repliement de Rossmann (Rossmann *et*

al., 1974) et d'une hélice α (résidus 312 à 333) se repliant sur le domaine N-terminal. Le domaine catalytique (résidus 149 à 311) est constitué d'un feuillet β antiparallèle de 7 brins, de 3 hélices α , ainsi que d'une région sans structure spatiale définie nommée S-loop (figure A11). La numérotation des résidus appliquée ci-dessus et dans tout le manuscrit est celle de la structure cristalline de la GAPDH de *B. stearothermophilus* (Biesecker *et al.*, 1977).

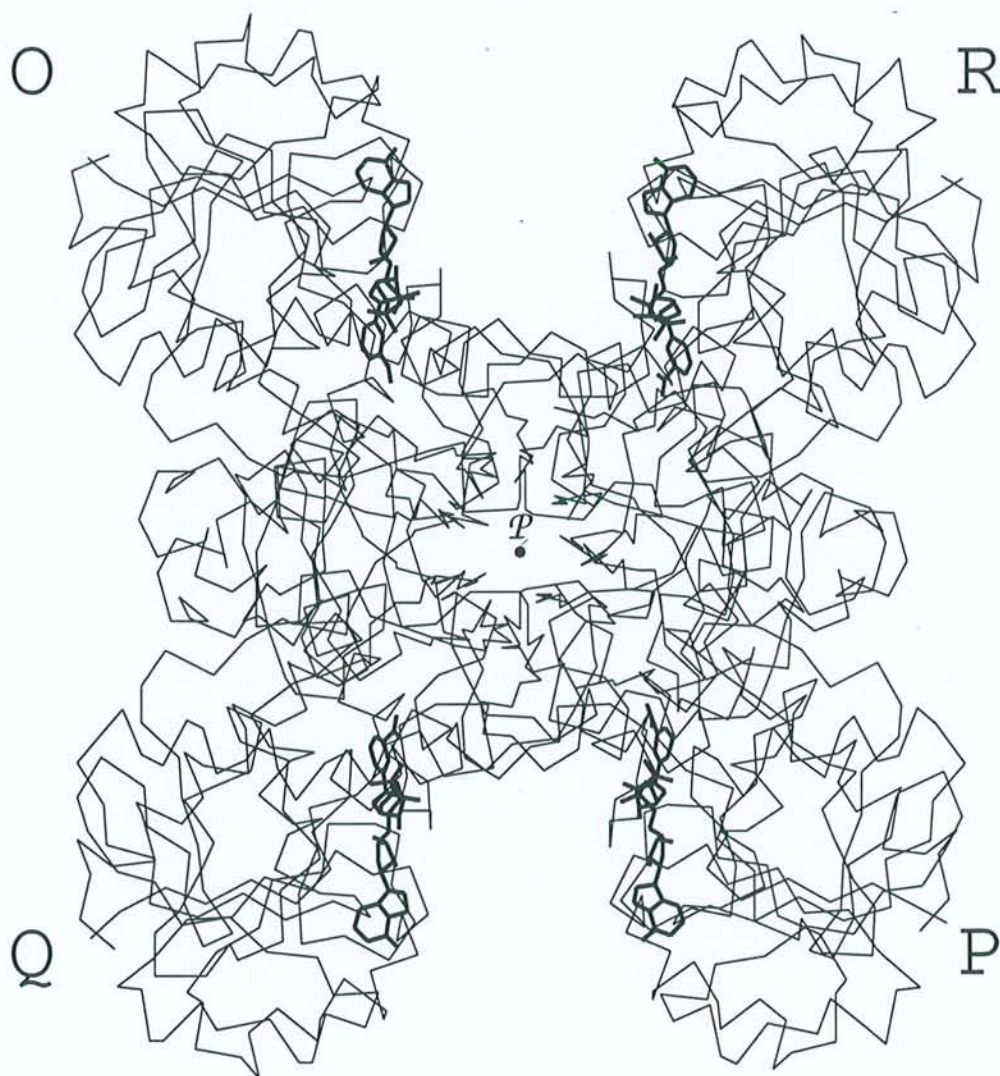


Figure A10 : Représentation de la chaîne des carbones α du tétramère de la GAPDH de *B. stearothermophilus* (traits fin), et des quatre molécules de NAD (trait épais).

L'axe binaire P est suivant la normale au plan de la figure et fait correspondre les monomères O et P. Les axes Q et R (non représentés) font correspondre respectivement le monomère O aux monomères Q et R.

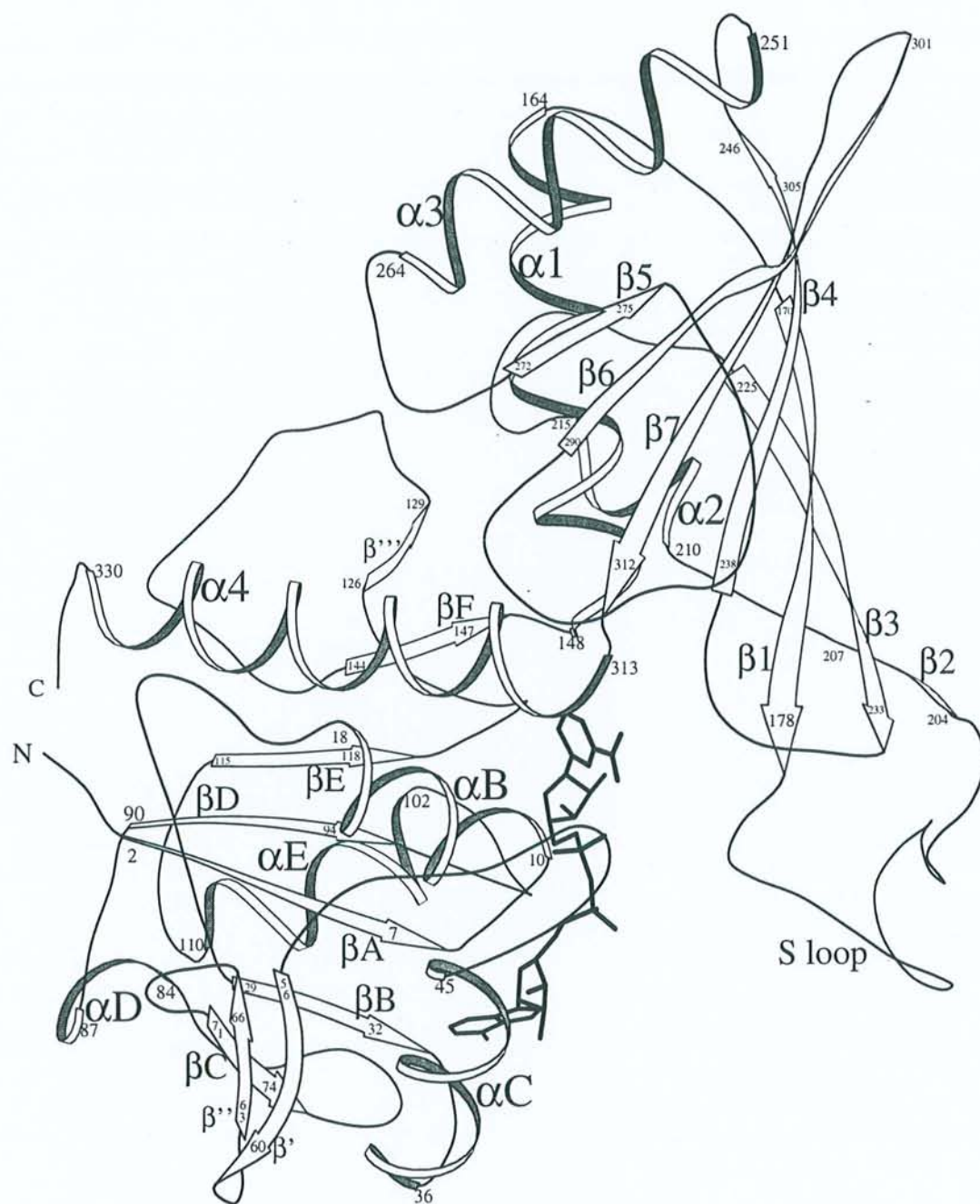


Figure A11 : Représentation en ruban d'un monomère de la GAPDH de *B. stearrowthermophilus*.

Le domaine de fixation du cofacteur est constitué des résidus de 1 à 148 ainsi que de l'extrémité C-terminale 312 à 333. Le domaine catalytique est constitué des résidus de 149 à 311.

AIVb2 Les GAPDH cristallisées

Le premier rapport structural sur une enzyme glycolytique a été fait sur la GAPDH de homard en 1964 (Watson & Banaszak, 1964). Depuis quinze GAPDH ont été cristallisées à ce jour:

- Neuf structures ont été résolues par diffraction X à haute résolution : humaine (*Homo sapiens*), (Watson *et al.*, 1972; Mercer *et al.*, 1976); homard américain (*Homarus americanus*) (Watson *et al.*, 1964; Moras *et al.*, 1975) ; *B. stearothermophilus* (Skarzynski *et al.*, 1987); langouste de Chine (*Palinurus versicolor*) (Tsou *et al.*, 1983; Lin *et al.*, 1993) ; *Thermotoga maritima* (Korndörfer *et al.*, 1995); *Trypanosoma brucei brucei* (Vellieux *et al.*, 1995); *Thermus aquaticus* (Tanner *et al.*, 1996); *Leishmania mexicana*, (Kim *et al.*, 1995); *Escherichia coli* (Duée *et al.*, 1996). Huit de ces structures sont disponibles dans la banque de données structurales des protéines de Brookhaven PDB (Bernstein *et al.*, 1977) (tableaux A1a, A1b). La première remarque intéressante concerne le degré de ressemblance entre ces neuf modèles : les structures tertiaires et quaternaires sont quasiment identiques. Les différences concernant les structures primaires et secondaires sont mineures. L'ensemble des résidus du site actif et la plupart des résidus impliqués dans la fixation du cofacteur sont conservés. Certaines de ces différences permettent toutefois d'énoncer des hypothèses sur la thermostabilité des protéines (Korndörfer *et al.*, 1995; Tanner *et al.*, 1996).

- Six ont fait l'objet d'études cristallographiques préliminaires avec détermination des paramètres de maille et même parfois identification de l'axe binaire cristallographique *P*, *Q* ou *R* : écrevisse (*Astacus*) (Campbell *et al.*, 1971); esturgeon (*Acipenser sturio*) (Holmes *et al.*, 1977); langouste méditerranéenne (*Palinurus vulgaris*) (Berni *et al.*, 1979); *Bacillus coagulans* (Lee *et al.*, 1982; Griffith *et al.*, 1983); muscle de porc (*Sus scrofa*)(Devedjiev *et al.*, 1994); *Methanothermus fervidus* (Fabry *et al.*, 1988) (tableaux A1a, A1b).

- Amelunxen & Carr (1967) ont simplement rapporté une procédure de cristallisation de la GAPDH de muscle de lapin (*Oryctolagus cuniculus*) utilisant le sulfate d'ammonium comme agent précipitant.

Tableau A1a : Récapitulatif des caractéristiques cristallographiques de huit GAPDH d'origines différentes.

	<i>H. sap.</i> ^a	<i>B. stea.</i> ^b	<i>Astac.</i> ^c	<i>H. ame.</i> ^d	<i>A. stur.</i> ^e	<i>P. vulg.</i> ^f	<i>B. coa.</i> ^g	<i>P. vers.</i> ^h
<u>Cristallisation</u>								
tampon	j	phosphate	j	j	Pipes ^k	j	Tris ^k	phosphate
pH	j	6.9	j	6.2	7.0	6.2	8.7	6.1
agent précipitant	SA ^k	SA ^k	j	SA ^k	SA ^k	SA ^k	SA ^k	SA ^k
<u>Diffraction X</u>								
groupe d'espace	C2	P2 ₁	C2	P2 ₁ 2 ₁ 2 ₁	P6 ₁ 22	C2	C222 ₁	C2
a (Å)	132.4	82.44	128.	149.0	82.2	128.4	95.7	128.11
b (Å)	97.9	124.1	100.	139.1	82.2	99.9	137.2	99.61
c (Å)	81.6	82.54	80.	80.7	458	80.3	131.9	80.69
β (°)	114.3	108.98	114	- ⁱ	- ⁱ	113.4	- ⁱ	114.4
Z ⁱ	8 (D)	8 (T)	8 (D)	16 (T)	24 (D)	8 (D)	16 (D)	8 (D)
Volume/Z (10 ³ Å ³)	120	100	117	105	112	101	108	117
axe P,Q,R	Q	aucun	un	aucun	un	Q	R	Q
cristallographique			non dét.		non dét.			
résolution (Å)	3.5	1.8	j	2.9	j	j	6.	2.7
<u>résolution de la</u>								
<u>structure</u>								
méthode	IR ^k	IR,NCS ^k	m	IR ^k	m	m	m	MR ^k
Facteur résiduel (%)	33	17.7	m	j	m	m	m	20.5
code PDB	3gpd	1gd1	m	1gpd	m	m	m	n

^a *H. sapiens*. ^b *B. stearothermophilus*. ^c *Astacus*. ^d *H. americanus*. ^e *A. sturio*. ^f *P. vulgaris*. ^g *B. coagulans*. ^h *P. versicolor*. ⁱ nombre de monomère(s) par maille, la lettre entre parenthèses indique le contenu de l'unité asymétrique : D pour dimère et T pour tétramère. ^j non indiqué dans les publications (références citées dans le texte). ^k abréviations données au début du manuscrit. ^l 90°. ^m structure non résolue. ⁿ structure non déposée dans la PDB.

Tableau A1b : Récapitulatif des caractéristiques cristallographiques de sept GAPDH d'origines différentes.

	<i>T. bruc.</i> ^a	<i>M. ferv.</i> ^b	<i>S. scrofa</i>	<i>T. aqu.</i> ^c	<i>E. coli</i>	<i>L. mex.</i> ^d	<i>T. mar.</i> ^e
<u>cristallisation</u>							
tampon	Tris ^g	phosphate	Hepes ^g	^h	Hepes ^g	phosphate	acétate s. ^g
pH	6.8	7.3	8.0	8.4	7.5	7.0	4.6
agent précipitant	SA ^g	phosphate	SA ^g	PEG ^g	trisodium	PEG-1000 ^g	SA ^g
				propanol	citrate		
<u>Diffraction X</u>							
groupe d'espace	P2 ₁ 2 ₁ 2	P4 ₁ /322	I222	P2 ₁ 2 ₁ 2 ₁	C222 ₁	P2 ₁ 2 ₁ 2 ₁	P6 ₁ 22
a (Å)	135.5	74.8	87.2	144.8	79.1	99.0	95.7
b (Å)	256.3	74.8	131.6	148.8	189.6	126.5	95.7
c (Å)	114.9	272	213.2	149.5	122.2	138.9	296.28
Z ^f	24 (T+D)	16 (D)	32 (T)	32 (2T)	16 (D)	16 (T)	24 (D)
Volume/Z (10 ³ Å ³)	166	95	76	101	114	109	113
axe P,Q,R	Q	un	aucun	aucun	R	aucun	P
cristallographique.		non dét.					
résolution (Å)	3.2	^h	3.5	2.5	1.8	2.8	2.5
<u>résolution de la</u>							
<u>structure</u>							
méthode	MR ^g	ⁱ	ⁱ	MR ^g	MR ^g	MR ^g	MR ^g
Facteur résiduel (%)	17.6	ⁱ	ⁱ	20.5	19.7	19.8	16.6
code PDB	1gga	ⁱ	ⁱ	1cer	1gad	1gyp	1hdg

^a *T. brucei brucei*. ^b *M. fervidus*. ^c *T. aquaticus*. ^d *L. mexicana*. ^e *T. Maritima*. ^f nombre de monomère(s) par maille, la lettre entre parenthèses indique le contenu de l'unité asymétrique : D pour un dimère et T pour un tétramère. ^g abréviations données au début du manuscrit. ^h non indiqué dans les publications (références citées dans le texte). ⁱ structure non résolue.

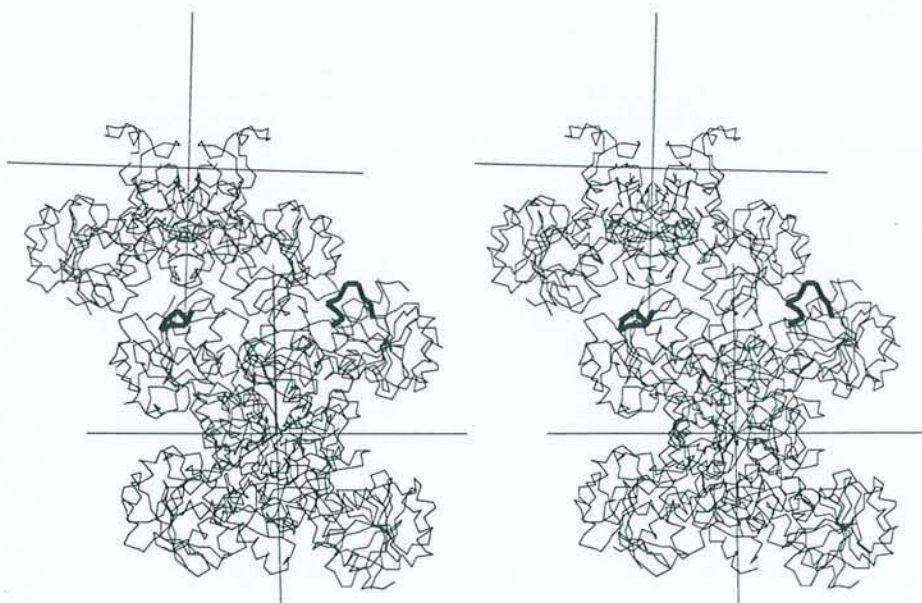
Sur l'ensemble de ces études cristallographiques, dans 75% des cas (11 sur 15) les cristaux de protéine ont été obtenus en utilisant le sulfate d'ammonium comme agent précipitant. Il est intéressant de remarquer que la GAPDH de *L. mexicana* cristallise dans une solution de PEG et de tampon phosphate où le phosphate est à une concentration proche de celle observée dans les conditions physiologiques (Kim *et al.*, 1995). On peut noter que les solutions de cristallisation des GAPDH de *E. coli* et de *T. aquaticus* ne contiennent pas d'anions phosphate et sulfate (tableau A1b).

AIVb3 Arrangement du tétramère dans le cristal

Considérons un peu plus en détail les 15 études cristallographiques répertoriées dans les tableaux A1a et b. Quatre composés cristallisent dans le groupe d'espace C2 avec un dimère par unité asymétrique (seulement deux structures sont résolues) et des paramètres de maille quasiment identiques. Au total, neuf GAPDH cristallisent avec un dimère (OP ou OQ ou OR) dans l'unité asymétrique, quatre avec un tétramère, une avec un tétramère et un dimère et une avec deux tétramères. Le rapport du volume de la maille sur le nombre de monomères par maille passe de $76 \cdot 10^3 \text{ \AA}^3$ dans la GAPDH de *S. scrofa* à $166 \cdot 10^3 \text{ \AA}^3$ dans la GAPDH de *T. brucei brucei*. Cette dernière possède de large canaux de solvant entre les tétramères expliquant le pourcentage élevé de solvant dans la maille (env. 74%) (Vellieux *et al.*, 1995).

Il est très intéressant, par exemple, de comparer l'empilement cristallin de la GAPDH de *B. stearothermophilus* avec celui de la GAPDH de *L. mexicana*. La GAPDH de *B. stearothermophilus* cristallise dans le groupe d'espace P2₁ avec un tétramère dans l'unité asymétrique. L'axe binaire *Q* non cristallographique est parallèle à l'axe hélicoïdal de la maille monoclinique. Tous les tétramères adoptent donc la même orientation dans le cristal. Chaque tétramère est en contact avec 6 tétramères voisins formant ainsi trois interfaces uniques (Skarzynski *et al.*, 1987). La figure A12a représente l'interaction tétramère-tétramère déduit de l'axe hélicoïdal (Pour simplifier la figure, seule la partie en contact du tétramère symétrique est représentée en l'occurrence un dimère). Cet empilement peut se décrire comme une association 'clé serrure' où la 'clé' est constituée par les résidus situés à la surface du domaine de fixation du cofacteur et la 'serrure' par la crevasse située entre les monomères O et Q (ou P et R).

L'empilement observé dans les cristaux de la GAPDH de *L. mexicana* semble très voisin de celui observé dans la GAPDH de *B. stearotherophilus* (figure A12b). Cependant les deux tétramères en interaction n'ont pas la même orientation. Pour en observer les différences, le tétramère symétrique de la GAPDH de *L. mexicana* a été réorienté comme le tétramère symétrique de la GAPDH de *B. stearotherophilus* de la figure A12a (figure A12c). Cette fiction montre que les boucles du premier tétramère dessinées en gras forment de mauvais contacts intermoléculaires avec le pseudo tétramère déduit de l'axe hélicoïdal observé dans la maille monoclinique de la GAPDH de *B. stearotherophilus*. Ces deux boucles sont justement différentes dans les structures primaires de ces deux GAPDH: les boucles α B- β B et β' - β'' sont plus longues respectivement de 4 et 8 résidus dans la GAPDH de *L. mexicana* (annexe 1, figures A5, A7a et b). Ces différences semblent être une des raisons principales du changement d'empilement. Le nouveau groupe d'espace est $P2_12_12_1$.

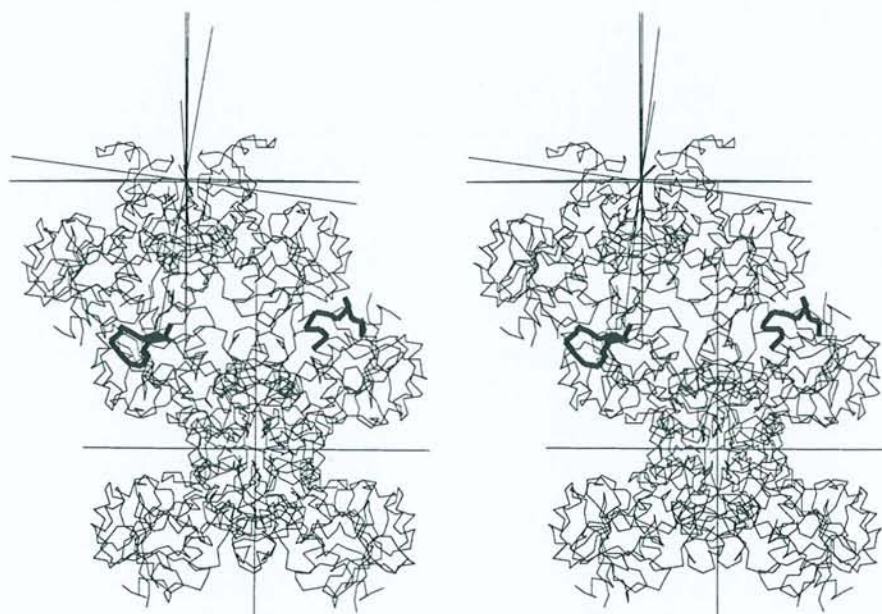


a) le tétramère de la GAPDH de *B. stearotherophilus* est représenté avec la partie inférieure de son équivalent déduit par l'opération de symétrie $-x, y+1/2, -z$.

Figure A12 : Comparaison des empilements cristallins de la GAPDH de *B. stearotherophilus* et de la GAPDH *L. mexicana*.



b) le tétramère de la GAPDH de *L. mexicana* est représenté avec la partie inférieure de son équivalent déduit par l'opération de symétrie $-x, y+1/2, -z$.



c) figure fictive où le demi tétramère de la GAPDH de *L. mexicana* a été construit comme dans le cas de la GAPDH de *B. stearothermophilus*. La figure montre une interpénétration des molécules, en particulier au niveau des boucles $\alpha B-\beta B$ et $\beta'-\beta''$ représentées en traits épais.

Figure A12 (suite) : Comparaison des empilements cristallins de la GAPDH de *B. stearothermophilus* et de la GAPDH de *L. mexicana*.

AIVb4 Relations de symétrie dans le tétramère

La GAPDH glycolytique se présente sous la forme d'un homotétramère dans lequel les quatre monomères sont reliés entre eux par trois pseudo axes binaires. Cette enzyme n'a jamais cristallisé avec un seul monomère dans l'unité asymétrique; les trois axes *P*, *Q* et *R* n'ont donc jamais coïncidé simultanément avec des axes binaires cristallographiques. Néanmoins, chacun de ces axes a été observé au moins une fois comme axe binaire vrai. Mises à part les structures des GAPDH de *H. americanus* et de *P. versicolor*, l'écart quadratique moyen résultant de la superposition des C α de deux sous-unités d'un même tétramère ne dépasse pas 0.45 Å. Sept structures présentent donc une symétrie 222 quasiment parfaite dans la limite des résolutions obtenues (tableaux A1a, A1b). Dans les cas étudiés, le défaut de symétrie est souvent lié à des résidus se trouvant à la surface des protéines. Comme l'environnement du tétramère est anisotrope dans le cristal, les chaînes latérales de ces résidus sont soit ordonnées car impliquées dans des interactions intermoléculaires, soit désordonnées car exposées au solvant. Le défaut de symétrie peut aussi être localisé dans le tétramère. Par exemple, le résidu Arg281 adopte deux conformations dans la GAPDH de *B. stearothermophilus* (une dans les monomères O et R et une différente dans les monomères P et Q). Si ce résidu avait la même conformation dans les sous-unités O et Q, les deux chaînes latérales portant une charge positive seraient trop proches l'une de l'autre par rapport à l'axe *Q* (Skarzynski *et al.*, 1987). Dans le cas de la GAPDH de *E. coli*, les différences observées entre les sous-unités O et P sont liées au protocole de cristallisation : dans les premiers essais de cristallisation (avec du citrate), le monomère O se présentait sous forme apo. Le cofacteur NAD a été introduit dans l'enzyme par trempage (soaking) pour obtenir une enzyme complètement sous forme holo. Cependant, l'analyse fine de la structure tridimensionnelle révèle des différences entre les sites de fixation des monomères O et P. Le monomère P adopte une conformation holo alors que la transition apo-holo est incomplète dans le monomère O. Ces différences structurales semblent être liées aux contraintes intermoléculaires entre tétramères dans le cristal (Duée *et al.*, 1996).

Les tétramères des GAPDH de *H. americanus* et de *P. versicolor* possèdent un seul axe binaire (axe *Q*). Les analyses structurales de ces modèles révèlent une dissymétrie marquée dans les dimères équivalents O-R et P-Q (rms $\approx$ 1.Å tableau A2). La GAPDH de *P. versicolor*,

présentant l'axe *Q* comme axe cristallographique vrai, constitue un cas particulier car cette enzyme a été cristallisée avec un dérivé fluorescent qui se comporte différemment dans les sous-unités O et R. Cette asymétrie semble être liée à la non-coopérativité de la protéine dans la fixation des quatre molécules de NAD dans le tétramère (Lin *et al.*, 1993).

Tableau A2 : Déviation à la symétrie 222 dans le tétramère de la GAPDH

	<i>H. sapiens</i>	<i>H. amer.</i>	<i>B. stearo.</i> ^c	<i>P. vers.</i>	<i>T. mar.</i>	<i>T. brucei</i> ^d	<i>T. aq.</i> ^c	<i>L. mex.</i> ^c	<i>E. coli</i>
rms Å ^a	.44	.93	.19 (2)	.88	.17	0.0	.26(1)	.26(1)	.28
rotation (°) ^b	179.9	179.5	179.7(2)	^e	179.7	180.	179.6(3)	179.7 (2)	178.8

^a écart quadratique moyen résultant de la superposition des carbones α de deux monomères. ^b rotation nécessaire pour superposer au mieux les carbones α de deux sous-unités. ^c valeur moyenne des six superpositions. La déviation standard est indiquée entre parenthèses. ^d affinement des axes non cristallographiques en mode restreint (Vellieux *et al.*, 1995). ^e valeur non calculable car la structure n'a pas été déposée dans la PDB.

AIVb5 Description des interfaces entre les sous-unités

L'interface la plus grande s'établit entre les sous-unités symétriques par rapport à l'axe *P* (environ 2000 Å²). Celle-ci est stabilisée par de nombreuses liaisons hydrogène (environ 25) et de nombreuses interactions hydrophobes impliquant des résidus du feuillet β antiparallèle du domaine catalytique. Il est intéressant de noter que la plupart de ces résidus sont conservés dans les structures primaires de la GAPDH (Skarzynski *et al.*, 1987; Duée *et al.*, 1996). L'interface entre les sous-unités symétriques par rapport à l'axe *R* (environ 1500 Å²) est formée par des résidus appartenant à l'hélice α C, à la boucle α C- β' et à la S-loop (figure A11). La troisième interface entre les monomères O et Q est la plus petite des trois (environ 500 Å²). Celle-ci est constituée par des résidus de l'hélice α C, de la boucle α C- β' interagissant avec la boucle β 5- β 6 (Figure A11).

AIVb6 Structure du monomère

Structure générale

Le monomère de la GAPDH, comme dans toutes les déshydrogénases, est composé de deux domaines : le domaine catalytique possédant une structure tridimensionnelle non commune aux autres déshydrogénases et le domaine de fixation du cofacteur présentant une structure tertiaire caractéristique des déshydrogénases : un repliement de 'Rossmann' (Rossmann *et al.*, 1974). L'analyse des structures cristallines des GAPDH montre que très peu de résidus du domaine de fixation du cofacteur forment des interactions avec des résidus du domaine catalytique. En moyenne 8 liaisons hydrogène sont observées. Il est intéressant de noter que les résidus impliqués (N146, N152, D312, N313, V237) sont conservés dans les différentes espèces (annexe 1). De nombreux ponts salins sont observés dans le domaine de fixation du cofacteur et dans le domaine du site actif, jouant un rôle dans la stabilisation des structures tertiaires de ces deux domaines. Parmi ceux-ci, certains sont conservés dans les différentes espèces (annexe 1). Par exemple, le résidu Arg10 situé au niveau de l'extrémité N-terminale de l'hélice α B forme deux ponts salins : le premier avec l'hélice α C à travers le résidu Asp47 et le second avec l'hélice α 4 à travers le résidu Glu314 (Moras *et al.*, 1975; Duée *et al.*, 1996; Skarzynski *et al.*, 1987; Korndörfer *et al.*, 1995). De même les résidus Asp241 et Lys306 du domaine catalytique forment un pont salin entre les brins β 4 et β 7.

Description du site NAD

Le cofacteur est situé dans une crevasse formée par le domaine de fixation du coenzyme (résidus 1 O à 147 O et 312 O à 333 O), par une première S-loop (résidus 178 O à 204 O), ainsi que par une seconde S-loop (résidus 178 R à 204 R) symétrique de la première par rapport à l'axe binaire *R*. Le site NAD le plus proche du monomère O est celui du monomère R (figure A10). Le cycle pyridinium du nicotinamide impliqué dans la réaction catalytique (transfert d'hydrure en position NC4) est situé dans le voisinage direct de la poche du site actif accessible au substrat. Ce cycle est stabilisé par une interaction directe entre la protéine (Asn 313 ND2) et l'oxygène du carboxamide et par des contacts hydrophobes avec les chaînes latérales des

résidus Ile11 et Tyr 317. L'oxygène NO4' du ribose du nicotinamide forme une liaison hydrogène avec l'oxygène de la chaîne latérale du résidu 119 qui est soit une sérine, soit une thréonine (annexe 1) (figures A8, A9a). Les deux autres oxygènes du ribose NO3' et NO2' n'ont aucune interaction directe avec la protéine. Cependant, l'oxygène NO2' forme une liaison hydrogène avec l'anion situé dans le site Ps (voir dans la partie site actif). Le pyrophosphate forme deux liaisons hydrogène avec les sites NH amides des résidus Arg10 et Ile11 localisés à l'extrémité N-terminale de l'hélice α B. Celle-ci stabilise avec son dipôle permanent la charge négative du pyrophosphate. Cette stabilisation est caractéristique du repliement de 'Rossmann' (Hol, 1985). Les structures de la GAPDH de *B. stearothermophilus* et de *T. aquaticus* révèlent une interaction supplémentaire protéine-pyrophosphate : le groupement amide ND2 du résidu Asn 180 forme une liaison hydrogène avec l'oxygène AO2 (figures A8, A9b). Ce résidu Asn est remplacé par une alanine ou une glycine dans la plupart des espèces de GAPDH (annexe 1). Le résidu conservé Asp32 localisé au niveau du C-terminal du brin β B est en interaction directe avec le ribose de l'adénosine. Les oxygènes OD1 et OD2 du carboxylate forment deux liaisons hydrogène, avec les groupements hydroxyles AO3' et AO2' du ribose respectivement. Ce type d'interaction est largement conservé dans les protéines possédant un repliement de Rossmann et fixant le NAD (Wierenga *et al.*, 1986). L'extrémité de la crevasse du site à NAD forme une poche hydrophobe où le cycle adénine est pris en tenaille entre les résidus 33 et 96. L'alignement des neuf séquences de GAPDH montre que la position 96 est toujours occupée par une thréonine alors que la position 33 est occupée soit par une proline soit par une leucine (annexe 1). Cette poche stabilise également le cycle avec une interaction directe entre le groupement carbonyle du résidu 77 et l'atome AN6 de l'adénine (figures A8, A9c).

La molécule de NAD adopte une conformation étendue dans les neuf structures de GAPDH résolues. Les cycles adénine et nicotinamide sont perpendiculaires aux plans moyens de leurs riboses voisins avec une conformation anti pour l'adénine et syn pour le nicotinamide. Les deux riboses adoptent le même 'puckering' C2' endo. Le nombre d'interactions directes protéine-NAD varie entre 7 et 8 dans les neuf modèles disponibles. De plus le NAD forme une

liaison hydrogène interne entre les atomes N7N et OP1N stabilisant la conformation syn du nicotinamide par rapport au ribose adjacent.

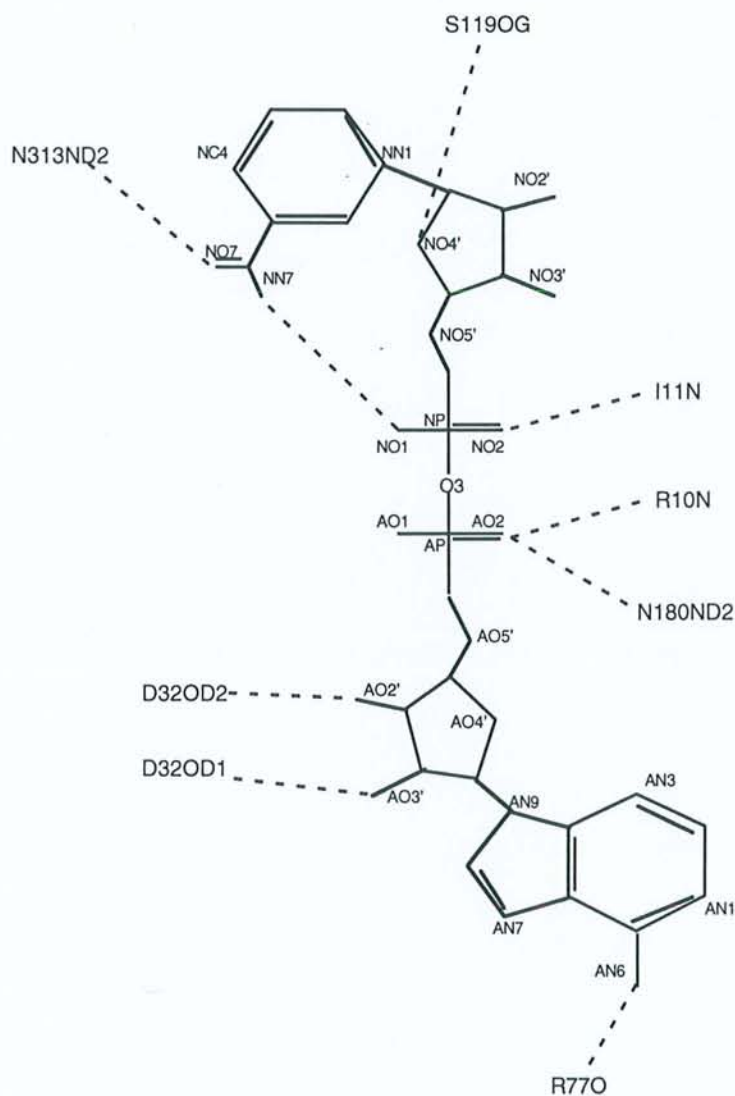
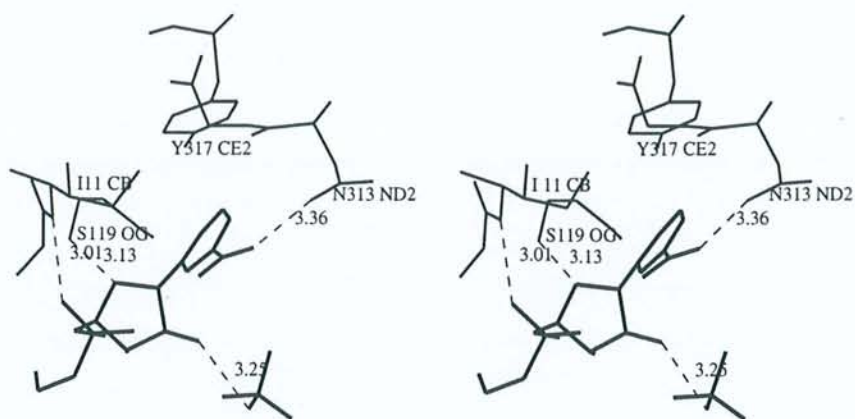
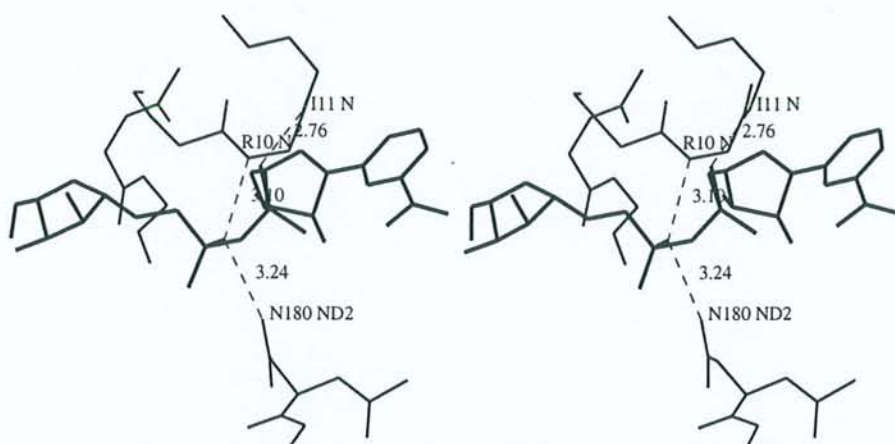


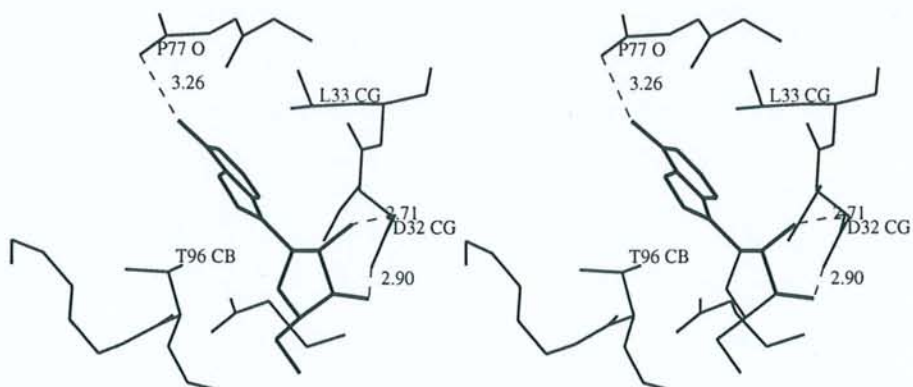
Figure A13 : Représentation schématique des liaisons hydrogène entre le NAD et la GAPDH de *B. stearothermophilus* à l'état cristallin (Skarzynski *et al.*, 1987).



a) Partie nicotinamide dans la GAPDH de *L. mexicana*.



b) Partie pyrophosphate dans la GAPDH de *T. aquaticus*.



c) Partie adénosine dans la GAPDH de *T. maritima*.

Figure A14 : Site de fixation du NAD dans la GAPDH.

Le site actif

Le site actif est situé dans le fond d'une poche étendue formée par la S-loop et les domaines catalytique et de fixation du cofacteur d'un même monomère. Ce site doit posséder une architecture particulière afin de réaliser la transformation du glycéraldéhyde-3-phosphate en acide 1,3-diphosphoglycérique (figure A9). Deux acides aminés jouent un rôle essentiel dans la réaction catalytique : la cystéine 149 dont la chaîne latérale est liée de manière covalente avec les différentes formes du substrat (hémithioacétal et thioacycle) (Harris & Perham, 1965) et l'histidine 176 dont la chaîne latérale intervient à plusieurs niveaux de la réaction (augmente la nucléophilie de la Cys149, joue le rôle de catalyseur acide-base, stabilise les intermédiaires réactionnels) (Soukri *et al.*, 1989). Le site actif peut se diviser en trois parties : la première est constituée par les deux acides aminés essentiels Cys149 et His176 et le carbone NC4 du nicotinamide impliqué dans le transfert d'hydrure (figure A15). Ces deux acides aminés sont conservés dans les GAPDH. Les deux autres parties forment deux sites anioniques dont l'un semble être occupé par le phosphate inorganique (site Pi) et l'autre par le phosphate en C3 du substrat (Site Ps). Cette notation des sites Pi et Ps a été introduite par Biesecker *et al.* (1977). Les neuf structures de GAPDH résolues dans des conditions légèrement différentes au niveau du site actif fournissent des informations très précieuses sur les différentes étapes de la réaction catalytique d'un point de vue structural. Par exemple, la GAPDH de *L. mexicana* est la seule enzyme cristallisée avec du phosphate comme anion majoritaire avec une concentration (200 mM) proche de celle observée dans les cellules (50 mM), alors que la plupart des GAPDH ont été cristallisées dans une solution contenant majoritairement l'anion sulfate à une concentration élevée (environ 1.5 M). Seules deux structures (*E. coli* et *T. brucei brucei*) cristallisées sans phosphate ni sulfate révèlent des formes apo au niveau des sites Pi et Ps. Les huit structures disponibles dans la PDB présentent des sites Ps pratiquement identiques. Les anions (phosphate ou sulfate) modélisés dans ce site sont stabilisés principalement par trois interactions : deux avec les chaînes latérales des acides aminés Thr179 et R231 (figure A15) (largement conservés dans les différentes espèces) et une avec l'hydroxyle NO2' du ribose de la partie nicotinamide du NAD. Ce site Ps est distant d'environ 7 Å du résidu catalytique Cys149. Par contre, on distingue deux types de sites Pi différents. Toutefois, les anions localisés dans

ces sites Pi forment essentiellement les mêmes interactions directes avec la protéine : avec les chaînes latérales des résidus Ser148 et Thr208 et la chaîne principale du résidu Gly209 (figure A15). Cette dernière interaction implique donc que l'anion soit très proche de la boucle $\beta 2$ - $\alpha 2$ (T207-A210). La conformation de cette boucle dans les GAPDH de *L. mexicana* et de *T. maritima* est différente de celle observée dans les sept autres modèles. Deux types de sites Pi positionnés différemment par rapport à la cystéine 149 sont donc mis en évidence (site de type I pour les sept structures et de type II pour les structures de *L. mexicana* et de *T. maritima*). La distance anion-Cys149 dans les sites de type II est inférieure d'environ 2Å à celle observée dans les sites de type I.

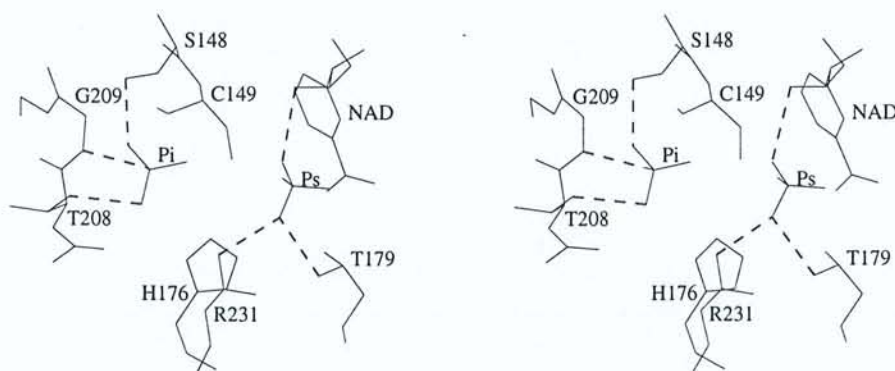


Figure A15 : Site actif de la GAPDH de *L. mexicana*.

AIVb7 Mécanisme réactionnel

Introduction

Ces études cristallographiques ont permis d'avancer plusieurs hypothèses sur la position du phosphate C-3 du substrat dans les intermédiaires hémithioacétal où le carbone C-1 du substrat est hybridé sp^3 et thioacycle où le carbone C-1 est hybridé sp^2 . Les positions des anions (sulfate ou phosphate) dans les sites Pi et Ps ont été identifiées comme celles des anions

(phosphate) impliqués dans la réaction catalytique. Le mécanisme réactionnel couramment admis (Segal & Boyer, 1953) est résumé dans la figure A16.

Intermédiaire hémithioacétal

Moras *et al.* (1975) ont proposé le premier modèle de l'intermédiaire hémithioacétal dans le site actif de la GAPDH de homard américain (figure A17a). La structure révélant une asymétrie marquée, l'His176 adopte une conformation différente dans chaque dimère O-Q et P-R. Le cycle imidazole est stabilisé par le résidu Ser238 dans les monomères P et R et par l'acide aminé Tyr311 dans les deux autres (O et Q). Le site actif de ces deux dernières sous-unités semble être mieux disposé pour représenter l'étape initiale de la réaction catalytique. La modélisation de l'intermédiaire hémithioacétal a donc été réalisée dans le site actif du monomère O. Ce modèle suppose que le phosphate C-3 du substrat soit localisé dans le site Ps (conformation Ps). Le substrat est orienté et stabilisé par au moins trois interactions: entre le groupement hydroxyle en C-1 et le résidu His176, entre le groupement hydroxyle en C-2 et le résidu Ser148 et entre le phosphate en C-3 et l'acide aminé Thr179 (figure A17a).

La résolution de la structure de la GAPDH de *B. stearothermophilus* à haute résolution (1.8Å) et les différences dans les séquences peptidiques (avec la GAPDH de Homard américain) ont conduit Skarzynski *et al.* (1987) à proposer un modèle différent de celui introduit par Moras *et al.* (1975). Ce modèle suppose que le site Pi du phosphate inorganique joue un rôle dans la fixation du phosphate en C-3 du substrat pour créer l'intermédiaire hémithioacétal (figure A17b) (conformation Pi). La cocrystallisation de l'enzyme avec le glycidol phosphate a confirmé cette suggestion car le phosphate C-3 de cet inhibiteur est localisé dans le site Pi (Skarzynski *et al.*, 1987). De plus, les résultats obtenus par Corbier *et al.* (1994) en mutagenèse dirigée semblent favoriser ce modèle.

L'observation d'un nouveau site Pi (type II) légèrement déplacé vers le résidu catalytique Cys149 dans les GAPDH de *T. maritima* et *L. mexicana* a conduit Kim *et al.* (1995) à construire l'intermédiaire hémithioacétal suivant les deux modes de fixation possibles décrits ci-dessus. En se basant sur des aspects conformationnels et énergétiques, la conformation Pi dans le site de type II est encore moins favorable que la conformation Pi dans le site de type I.

La situation la plus plausible semble être un intermédiaire hémithioacétal adoptant une conformation Ps.

Duée *et al.* (1996) proposent un modèle similaire à celui de Moras *et al.* (1975). Ils supposent que l'intermédiaire hémithioacétal se trouve dans une conformation impliquant le site Ps. Comme la structure est sous forme apo au niveau des sites Pi et Ps (cristallisation sans phosphate ni sulfate), les sites Pi et Ps sont supposés être localisés comme les sites Pi et Ps de la GAPDH de *B. stearothermophilus*. La seule différence se situe au niveau de la participation de la chaîne latérale de la sérine 148 dans la stabilisation du groupement hydroxyle en C-2 du substrat. L'intermédiaire hémithioacétal construit dans la GAPDH de homard américain suppose une interaction entre les deux groupements hydroxyles. Celle-ci n'est pas modélisée dans la GAPDH de *E. coli* où cette sérine est impliquée dans deux liaisons hydrogène avec le résidu Thr151 (Duée *et al.*, 1996; Corbier *et al.*, 1989). Le groupement hydroxyle en C-2 du substrat est stabilisé par le groupement NH de la Cys149 dans cette structure.

Intermédiaire thioacyl et phosphorylation

Tous les modèles proposés supposent que l'intermédiaire thioacyl se trouve dans une conformation Ps juste avant la phosphorylation, dans laquelle l'attaque nucléophile du phosphate inorganique sur le carbone C-1 du substrat se fait à partir du site Pi. Le modèle proposé par Skarzynski *et al.* (1987) suppose donc un passage du phosphate C-3 de l'intermédiaire thioacyl du site Pi vers le site Ps (Corbier *et al.*, 1994). Les sites actifs les plus adéquats pour la phosphorylation semblent être ceux observés dans les GAPDH de *T. maritima* et de *L. mexicana* (site Pi de type II). L'oxygène du phosphate inorganique impliqué dans l'attaque nucléophile semble être idéalement positionné (distance et angle) pour réagir avec le carbone C-1 du substrat (Kim *et al.*, 1995).

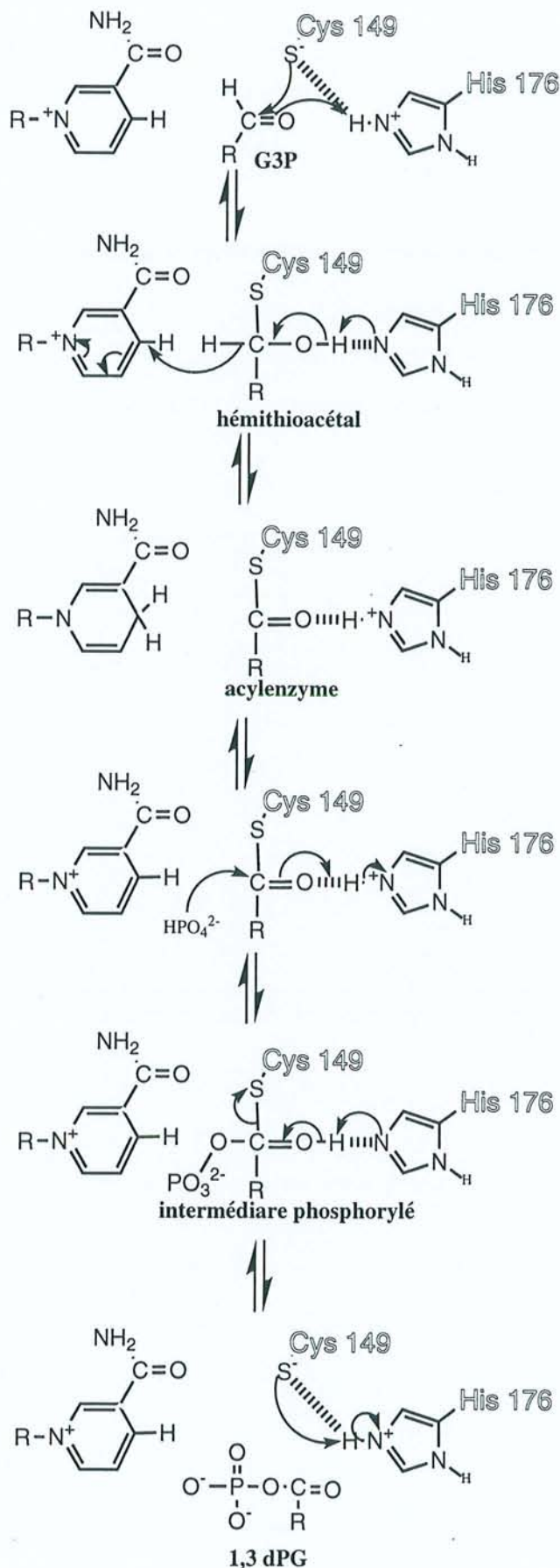


Figure A16 : Mécanisme réactionnel.

Le soufre S^γ de Cys149 à l'état thiol attaque le centre électrophile du G3P: La formation de l'hémithioacétal est assistée par His 176 protonée qui, en tant qu'acide, permet le transfert de proton sur l'atome d'oxygène du substrat.

L'hémithioacétal est stabilisé par liaison hydrogène avec His 176 neutre.

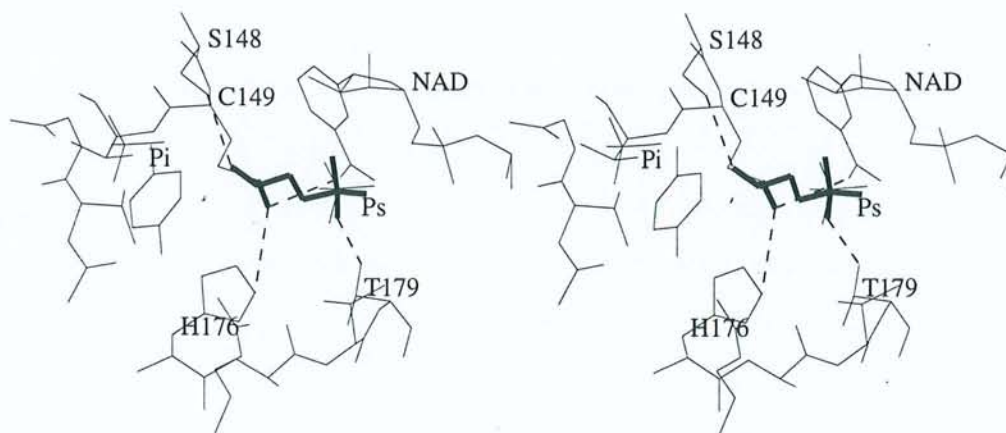
His 176 joue le rôle de base et facilite le transfert d'hydrure de l'hémithioacétal vers NAD^+ .

L'acyl-enzyme est stabilisée par liaison hydrogène avec His 176 protonée.

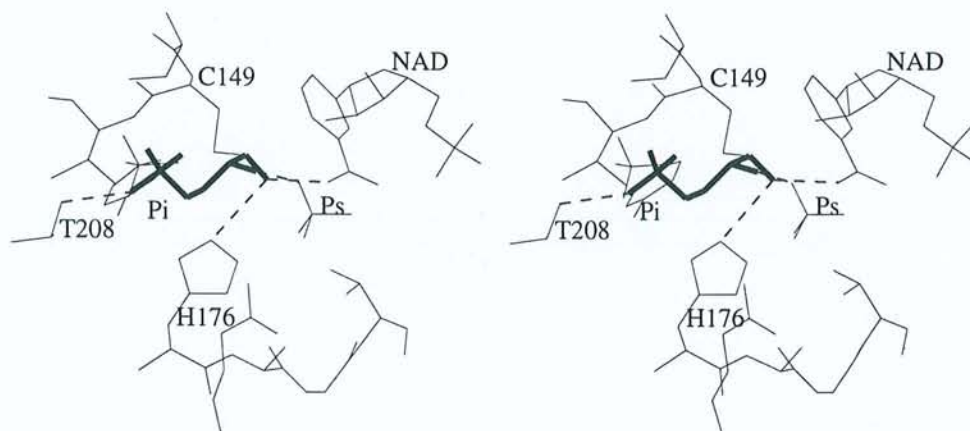
L'étape d'isomérisation de l'acyl-enzyme se fait après remplacement du NADH par un NAD^+ . L'attaque nucléophile du phosphate est assistée par His 176 protonée jouant le rôle d'acide.

La formation du 1,3 dPG est assistée par His 176 jouant le rôle de base.

Le retour du soufre S^γ de Cys 149 à l'état thiol est assisté par His 176 protonée jouant le rôle d'acide.



(a)



(b)

Figure A17 : Modélisation de l'intermédiaire hémithioacétal dans les sites actifs des GAPDH de homard américain (a) et de *B. stearothermophilus* (b). Le substrat est représenté en gras.

AIVb8 Transition apo-holo

Introduction

Des études structurales sur la fixation du cofacteur NAD ont été effectuées sur les GAPDH de homard américain, de *B. stearothermophilus* et de *E. coli* montrant des différences structurales entre les formes 'apo' et 'holo'. Les deux premières ont été cristallisées sous forme 'apo' (Murthy *et al.*, 1980; Skarzynski & Wonacott, 1988) et 'holo' (Moras *et al.*, 1975; Skarzynski *et al.*, 1987). Dans la troisième, le mutant N313T semble adopter une conformation plutôt 'apo' (Duée *et al.*, 1996). Les formes 'apo' et 'holo' des deux enzymes d'origine bactérienne semblent présenter une symétrie 222 quasi parfaite alors que la GAPDH de homard américain présente une dissymétrie marquée dans le dimère O-P sous forme 'holo' et peu de défauts de symétrie dans le tétramère sous forme 'apo'.

Transition apo-holo dans les GAPDH de *E. coli* et de *B. stearothermophilus*

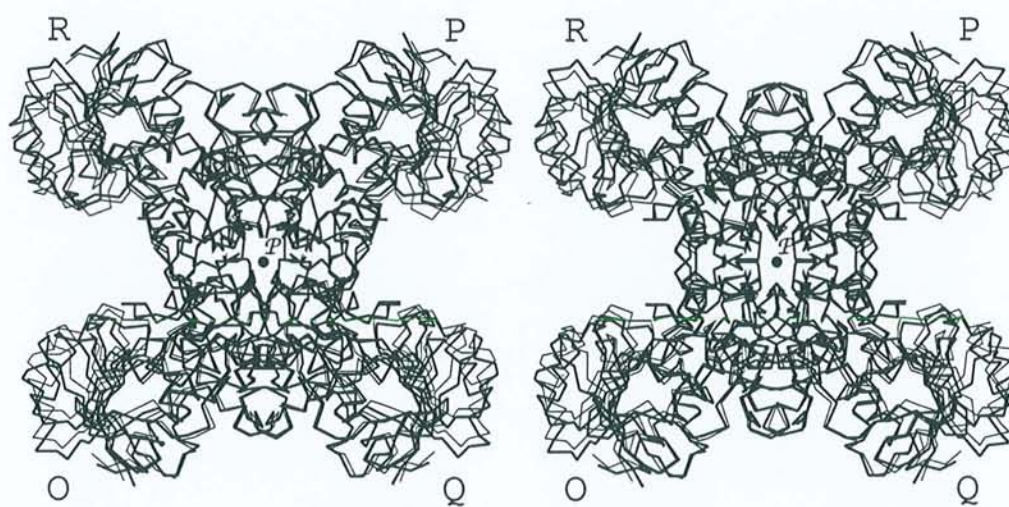
Les études structurales de ces deux GAPDH d'origine bactérienne décrivent un mécanisme commun dans la transition apo-holo. Cette transition peut se décomposer en trois mouvements :

- Le premier mouvement est une rotation autour de l'axe *P*, dans laquelle la partie fixe localisée au coeur du tétramère implique des résidus des brins $\beta 1$, $\beta 3$, $\beta 4$, $\beta 7$ du feuillet antiparallèle du domaine catalytique (en particulier un résidu impliqué dans le mécanisme catalytique : His176). Plus précisément, le tétramère se comporte comme deux dimères (O-P) et (Q-R) tournant autour de l'axe *P* d'environ 1° dans des directions opposées (figure A18a).

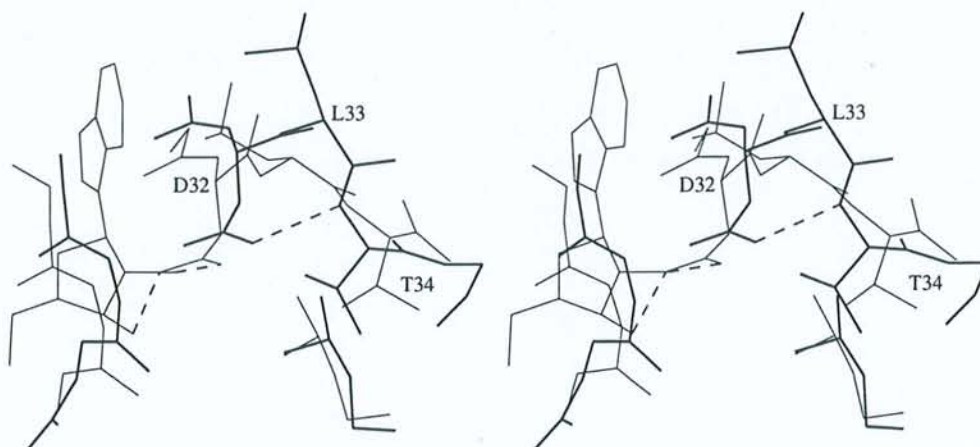
- Le deuxième mouvement peut être décrit par une rotation de chaque domaine de fixation du cofacteur : les domaines des sous-unités O et R subissent une rotation en corps rigide (environ 5°) dans des directions opposées autour de la première bissectrice des axes binaires *P* et *Q*, alors que les domaines des sous-unités P et Q subissent une rotation en corps rigide (environ 5°) dans des directions opposées autour de la deuxième bissectrice des axes binaires *P* et *Q*.

- Le dernier mouvement consiste en un réarrangement local de la région 0-69 de chaque sous-unité. Par exemple, les résidus Asp32 et Leu33 de la boucle βB - αC impliqués

dans la fixation du cofacteur subissent un changement conformationnel. Dans le sens holo vers apo ce changement crée une nouvelle interaction entre la chaîne latérale du résidu Asp32 et la chaîne principale du résidu Thr34 (Asp32-OD2....Thr34-N) (Skarzynski & Wonacott, 1988) (figure A18b). Cette conformation n'est pas observée dans le mutant N313T de la GAPDH de *E. coli*, car la transition holo vers apo n'est pas complète (Duée *et al.*, 1996).



a) rotation des dimères O-P et Q-R autour de l'axe P (perpendiculaire au plan de la figure) dans des sens opposés.



b) changement conformationnel de la boucle $\beta B-\alpha C$.

Figure A18 : Transition apo-holo dans la GAPDH de *B. stearothermophilus*. La forme holo est représentée en traits fins et la forme apo en traits épais.

Transition apo-holo dans la GAPDH de homard américain

Le mécanisme structural de la transition apo-holo dans la GAPDH de homard américain paraît très différent de celui décrit ci-dessus dans les GAPDH de *B. stearothermophilus* et de *E. coli*. Alors que dans les GAPDH d'origine bactérienne les S-loop des formes apo et holo sont très peu différentes, la superposition des carbones α des deux formes de la GAPDH de homard américain montre que les écarts les plus marqués sont localisés au niveau de la S-loop (Murthy *et al.*, 1980). De plus, certains résidus du domaine de fixation du cofacteur semblent être fixes dans cette transition apo-holo (en particulier le résidu His50) alors que le domaine de fixation du cofacteur subit une rotation en corps rigide dans la transition apo-holo des GAPDH de *B. stearothermophilus* et de *E. coli* (Skarzynski & Wonacott, 1988; Duée *et al.*, 1996).

AV SPECIFICITE DE COFACTEUR NAD/NADP DANS LA GAPDH

AVa Première approche : double mutant S

AVa1 introduction

Les glycéraldéhyde 3-phosphate déshydrogénases phosphorylantes présentent plusieurs types de spécificité de cofacteur : la GAPDH glycolytique est NAD dépendante alors que la GAPDH chloroplastique utilise aussi bien NADP que NAD comme cofacteur avec une légère préférence pour NADP (Ferri *et al.*, 1978). Dans un premier temps, deux mutations dans le domaine de fixation du cofacteur de la GAPDH de *B. stearothermophilus* (L187A et P188S) ont été réalisées pour tenter d'identifier les facteurs moléculaires responsables de la sélectivité de cofacteur dans les GAPDH (Corbier *et al.*, 1990a). Le choix de la substitution de ces deux acides aminés a été basé sur deux études complémentaires : la comparaison des structures primaires des GAPDH glycolytiques et chloroplastiques et la modélisation moléculaire du NADP dans le site de fixation du NAD de la GAPDH de *B. stearothermophilus* (Skarzynski *et al.*, 1987; code PDB 1gd1). Les résidus 187 et 188 de la GAPDH (numérotation de *B. stearothermophilus*) localisés dans la S-loop symétrique par rapport l'axe R, forment un

environnement de type quaternaire dans le site de fixation du cofacteur. Dans la GAPDH de *B. stearothermophilus*, les positions 187 et 188 sont occupées respectivement par les résidus hydrophobes leucine et proline. La modélisation du NADP dans le site de fixation du cofacteur de cette GAPDH sous forme holo a montré que ces deux résidus volumineux gênaient d'un point de vue stérique le positionnement du 2'-phosphate du NADP (figure A19). Ces deux résidus leucine et proline ont alors été respectivement remplacés par les acides aminés alanine et sérine observés dans les GAPDH chloroplastiques (annexe 1). Le double mutant de la GAPDH de *B. stearothermophilus* résultant, nommé mutant S, a fait l'objet d'études enzymatiques (Corbier *et al.*, 1990a) et cristallographiques (Branlant, Rahuel-Clermont & Wonacott, résultats non publiés).

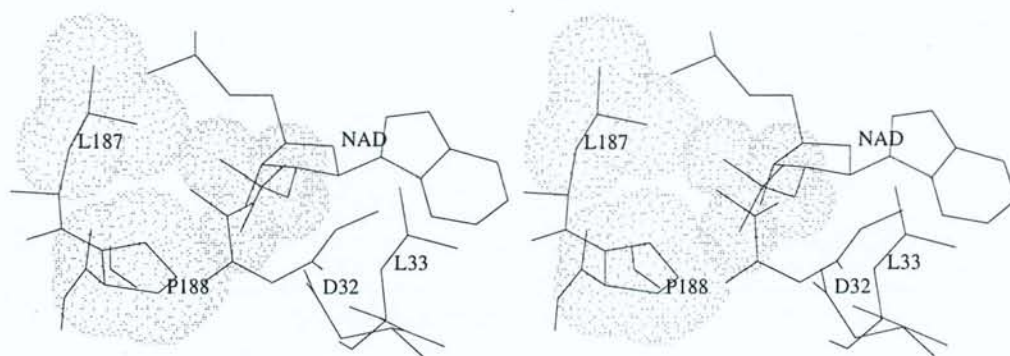


Figure A19 : Modélisation du NADP dans le site du NAD de l'holoenzyme de type sauvage de *B. stearothermophilus*.

AVa2 Résultats enzymatiques

Les études cinétiques réalisées par Corbier *et al.*, (1990a) ont montré que l'enzyme mutée deux fois était active avec les cofacteurs NAD et NADP, alors que l'enzyme de type sauvage ne l'est qu'avec NAD. Cependant, l'affinité de ce mutant S pour le NADP demeure très inférieure (d'un facteur 1000) à celle observée dans l'enzyme de type sauvage pour le NAD (Clermont *et al.*, 1993). Cette faible affinité semble due à la présence du résidu aspartique 32 dans le voisinage direct du groupement 2'-phosphate du NADP.

AVa3 Résultats structuraux

Le mutant S a été cocristallisé avec NADP (structure nommée **Sp**) dans les mêmes conditions que l'holoenzyme de type sauvage (Skarzynski *et al.*, 1987; Branlant, Rahuel-Clermont & Wonacott, résultats non publiés). Les cristaux obtenus sont isomorphes de la forme 'Holo I' de la GAPDH de *B. stearotherophilus* décrite par Leslie & Wonacott (1983). La structure a été déterminée à 2.7 Å de résolution et l'indice résiduel entre les facteurs de structure calculés et observés du modèle affiné est de 16 %. L'analyse structurale semble confirmer l'ensemble des résultats enzymatiques : d'une part le second mouvement de la transition apo-holo est incomplet (cf. § AIVb8 ; Duée *et al.*, 1996; Skarzynski & Wonacott, 1988); d'autre part le groupement 2'-phosphate du NADP n'est pas stabilisé dans le site de fixation du cofacteur (figure A20). Celui-ci pointe vers l'extérieur de la protéine en ne formant aucune interaction avec elle ni avec des molécules de solvant. De plus, la densité électronique du résidu muté S188 situé dans le voisinage direct du groupement 2'-phosphate du NADP est mal définie. Comme l'avaient supposé Corbier *et al.* (1990a), la charge négative portée par la chaîne latérale du résidu Asp32 semble liée à la non stabilisation de ce groupement phosphate. Cependant, même si cette répulsion électrostatique était prévisible, le résidu Asp32 n'a pas été muté car il est conservé dans les structures primaires des GAPDH chloroplastiques.

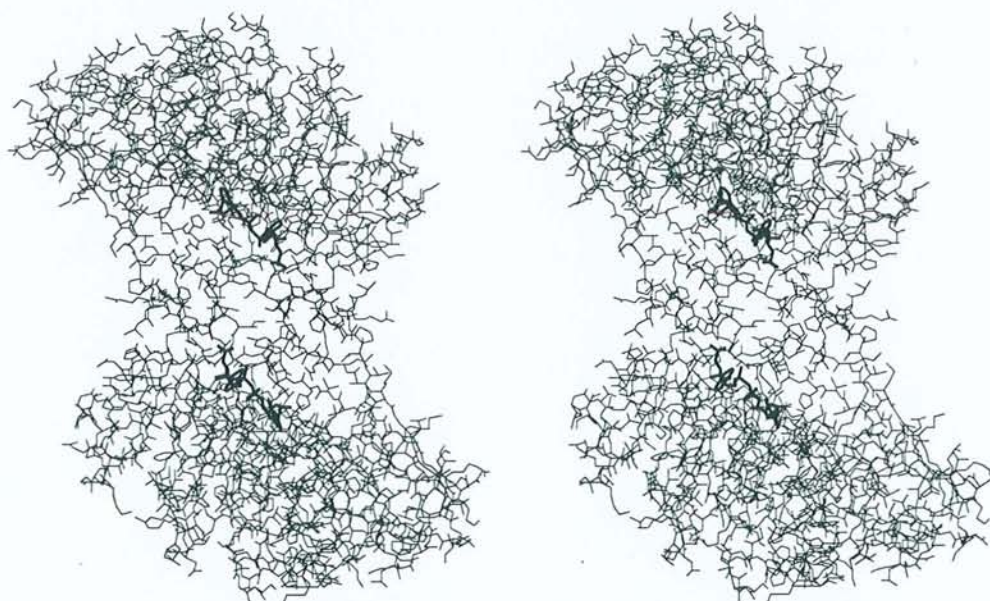


Figure A20 : Sites de fixation du NADP dans le dimère O-R du double mutant S. La protéine est représentée en traits fins et le cofacteur NADP en traits gras.

AVa4 Perspectives

Sur la base de ces résultats, deux directions de recherche ont été suivies :

- essayer d'obtenir une enzyme mutée capable de fixer le NADP de manière aussi efficace que l'enzyme de type sauvage avec le NAD. Les mutations L187A et P188S n'étant pas suffisantes, le résidu Asp32 a alors été remplacé soit par une alanine (mutant **D32A-S**) soit par une glycine (mutant **D32G-S**) (Clermont *et al.*, 1993; Eyschen *et al.*, 1996).

- mimer au mieux la GAPDH chloroplastique. Les résidus Thr33, Gly34 et Gly35 de la boucle β B- α C sont largement conservés dans les GAPDH chloroplastiques. Le résidu Thr33, moins hydrophobe que le résidu Leu33 de la GAPDH de *B. stearothermophilus*, est susceptible de former une interaction avec le 2'-phosphate du NADP. Les positions 34 et 35 occupées par des résidus à chaîne latérale dans la GAPDH de *B. stearothermophilus* gênent probablement la fixation du 2'-phosphate du NADP. Le quintuple mutant L33T, T34G, D35G, L187A et P188S (mutant **B-S**) a alors été construit (Clermont *et al.*, 1993).

AVb Deuxième approche : triples mutants et quintuple mutant

AVb1 Résultats enzymatiques sur les triples mutants

Les études enzymatiques montrent que les deux mutants D32A-S et D32G-S présentent une dualité de cofacteur NAD/NADP avec une préférence pour NADP. Les constantes de dissociation K_D du NAD et du NADP sont plus faibles dans le mutant D32G-S que dans le mutant D32A-S. L'objectif est atteint dans le mutant D32G-S : le K_D du NADP dans ce mutant est du même ordre de grandeur que celui du NAD dans l'enzyme de type sauvage. Néanmoins, l'enzyme de type sauvage est toujours plus active avec NAD que ne l'est le mutant D32G-S avec NADP (Clermont *et al.*, 1993; Eyschen *et al.*, 1996).

AVb2 Résultats enzymatiques sur le quintuple mutant

Le quintuple mutant B-S comme le double mutant S présente une dualité de cofacteur NAD/NADP avec une préférence pour NAD. Même si les trois mutations supplémentaires ont permis d'augmenter légèrement l'affinité du mutant B-S pour le NADP, celle-ci reste faible. Les

études enzymatiques montrent que les K_D du NAD dans l'enzyme de type sauvage et dans le mutant S sont similaires, alors que celui observé dans le mutant B-S est environ 5 fois plus grand (Clermont *et al.*, 1993; Eyschen *et al.*, 1996). La substitution de la séquence LTD par la séquence TGG a probablement apporté une liberté conformationnelle dans la boucle $\beta B-\alpha C$, ce qui a pour effet de diminuer l'affinité du mutant pour le NAD.

AVc Objectifs du travail

Le but de ce travail était de compléter *via* la radiocristallographie le travail de thèse de S. Rahuel-Clermont dans la recherche des facteurs moléculaires responsables de la spécificité de cofacteur des déshydrogénases à NAD(P). Nous avons choisi de cocrystalliser les mutants D32G-S et B-S avec NAD puis NADP (tableau A3) afin de comparer dans chaque site les modes de fixation du NAD et du NADP. La résolution de la structure du mutant D32G-S complexé avec NADP devait nous permettre de corréler l'affinité de ce mutant pour le NADP avec le comportement du groupement 2'-phosphate du NADP dans la structure. La détermination structurale du mutant B-S avait pour objectif d'affiner nos connaissances sur le mode de fixation du cofacteur des GAPDH chloroplastiques. En parallèle à cette étude cristallographique, Eyschen *et al.* (1996) se sont intéressés à la nature mono ou dianionique du 2'-phosphate du NADP dans les mutants présentés ci-dessus par spectroscopie RMN du phosphore (^{31}P). Les résultats obtenus par cette technique seront évoqués dans la partie 'Résultats et Discussion' de ce manuscrit.

Tableau A3 : Notations des mutants.

mutant	mutations	mutant complexé avec NAD ⁺	mutant complexé avec NADP ⁺
D32G-S	D32G	D32G-Sn	D32G-Sp
	L187A		
	P188S		
B-S	L33T	B-Sn	B-Sp
	T34G		
	D35G		
	L187A		
	P188S		

PARTIE B

MATERIELS ET METHODES

Les différentes étapes de la résolution des structures cristallines des macromolécules biologiques sont :

- La production et la purification
- La cristallisation
- L'acquisition et l'intégration des données de diffraction
- La recherche des phases
- L'affinement du modèle

BI PRODUCTION ET PURIFICATION

Les mutants de la GAPDH de *Bacillus stearothermophilus* proviennent du Laboratoire d'enzymologie et génie génétique de l'Université Henri Poincaré de Nancy I (LEGG UHP NANCY I). Ils ont été produits et purifiés par S. Rahuel-Clermont (1994). L'enzyme mutée est surexprimée dans une souche d'*Escherichia coli* où le gène chromosomique de GAPDH a été inactivé. Le gène de l'enzyme de *B. stearothermophilus* préparé par mutagenèse dirigée est inséré dans un plasmide de surexpression. Ce plasmide est introduit dans les cellules de la souche par électroporation.

L'enzyme surexprimée doit être extraite des cellules de culture, pour cela dans un premier temps le mélange biologique est soniqué. Les débris cellulaires sont éliminés par centrifugation. Le surnageant de sonication est amené à 66% de saturation en sulfate d'ammonium. Il est alors possible d'éliminer une grande partie des protéines autres que la GAPDH car elles précipitent à cette concentration en sulfate d'ammonium. La solution protéique est alors passée successivement sur deux colonnes : la première (filtration sur gel) permet de distinguer les pics protéiques des pics d'acides nucléiques et la deuxième (chromatographie d'échange d'anions) permet de séparer la GAPDH des fragments nucléiques de même taille. Les

techniques d'électrophorèse en gel et de spectrométrie de masse sont utilisées pour déterminer la qualité de la purification.

BII CRISTALLISATION

BIIa Principe

L'étape suivante dans l'étude structurale d'une macromolécule par diffraction X est l'obtention de monocristaux. Cette étape est le point fondamental de toute étude structurale car la qualité des mesures est souvent fortement liée à la qualité des cristaux. Le biocristallographe doit donc aborder ce domaine avec le plus grand soin pour obtenir des cristaux exploitables par radiocristallographie. Le protocole de cristallisation doit être reproductible car le cristallographe utilisera généralement un grand nombre de cristaux pour résoudre la structure cristalline. Pour arriver à cela, il est indispensable d'avoir toujours à l'esprit le diagramme de phases d'une protéine pour comprendre les lois de la cristallisation. Ce diagramme met en évidence la courbe de solubilité qui délimite l'état dissous de l'état sursaturé (figure B1).

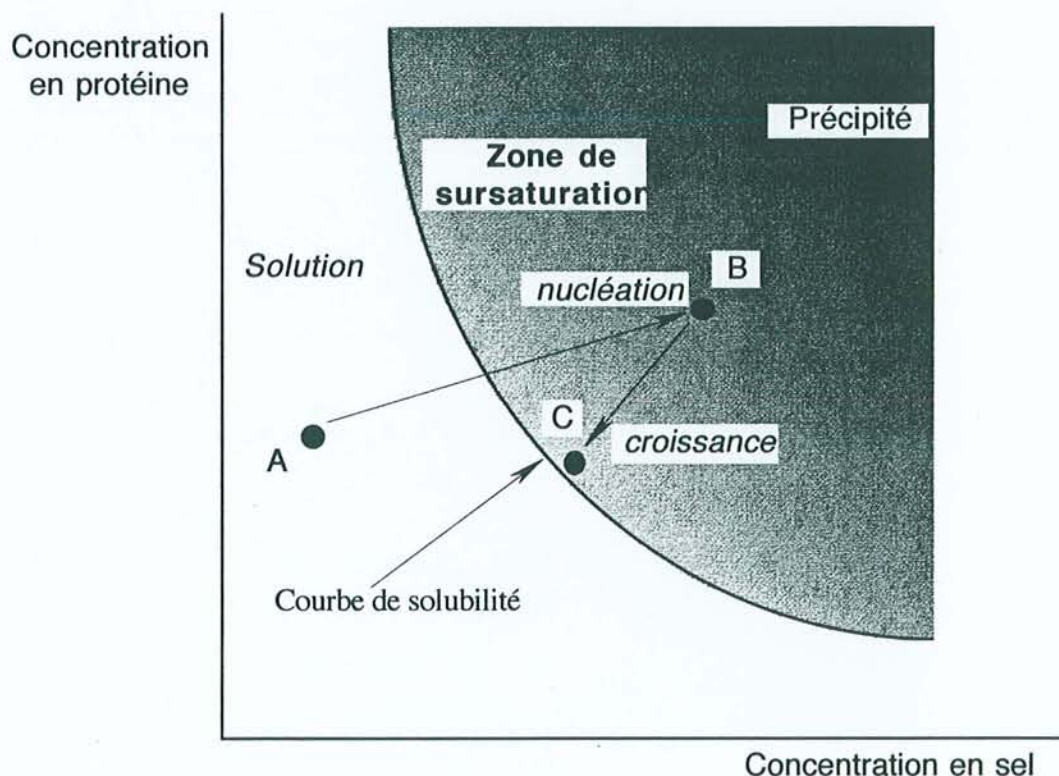


Figure B1 : Diagramme de solubilité d'une protéine.

La formation d'un cristal se déroule en deux étapes distinctes : la première a pour but de créer un germe en faisant passer le composé de la zone de solubilité (point A) à la zone de nucléation (point B) et la deuxième de faire grossir le germe cristallin en le faisant passer de cette zone de nucléation au point C, point au niveau duquel la croissance s'arrête.

Le nombre important de paramètres (pH, température, pression, force ionique, agents précipitants,...) pouvant influencer sur la cristallisation d'une protéine rend toute expérimentation très délicate. Il faut prendre soin de bien dissocier ces différents paramètres et de les faire varier un à un et pas à pas. L'idéal est de connaître le diagramme de solubilité de la macromolécule en fonction d'un maximum de paramètres. Malheureusement, dans la plupart des cas il est impossible de construire ce diagramme en raison de la faible quantité de macromolécule disponible.

Les conditions physico-chimiques de la zone de nucléation et de la zone de croissance ne sont pas identiques, ce qui rend la recherche des conditions de croissance souvent très difficile. La nucléation demande un état de sursaturation de la protéine plus important que celui de la phase de croissance (figure B1). Si l'affinement des conditions de nucléation ne permet pas l'obtention de cristaux de taille suffisante, il est possible de faire appel à des techniques de semis pour améliorer la croissance des cristaux.

La recherche des conditions initiales de nucléation d'une protéine nouvelle devrait se faire de façon systématique. Comme l'étude des variations des nombreux paramètres de cristallisation demande beaucoup de tests, cette pratique est très rarement utilisée. En fait, une matrice rassemblant cinquante conditions de cristallisation est utilisée (Jancarik & Kim, 1991). Ces cinquante tests proviennent d'une étude statistique sur toutes les conditions fructueuses de cristallisation observées en biocristallographie et représentent environ 90% des cas rencontrés. Il est vraiment très rare d'obtenir directement de beaux cristaux à partir de cette liste de conditions. Cette méthode permet simplement de trouver les conditions de départ de la nucléation. L'expérimentateur devra modifier ces conditions initiales de façon à améliorer la croissance cristalline.

BI Ib La diffusion de vapeur

La plupart des expériences de cristallisation ont été réalisées en utilisant la méthode de diffusion de vapeur en gouttes suspendues (Wlodawer & Hodgson, 1975). Une goutte de faible volume (10 μ l) contenant la macromolécule, le tampon et l'agent précipitant, est déposée sur une lamelle de verre. Cette lamelle est retournée sur un réservoir contenant 1 ml du même tampon et de l'agent précipitant plus concentré que dans la goutte (figure B2).

La différence de concentration en agent précipitant entre la goutte suspendue et le réservoir entraîne une diffusion de vapeur de la goutte vers le puits. L'équilibre est atteint quand la concentration en précipitant dans la goutte est identique à celle du réservoir. La diminution du volume de la goutte suspendue induit un accroissement de la concentration de la macromolécule. Si les conditions sont favorables, cette augmentation de concentration permet de faire passer la macromolécule d'un état dissous à un état sursaturé (figure B2).

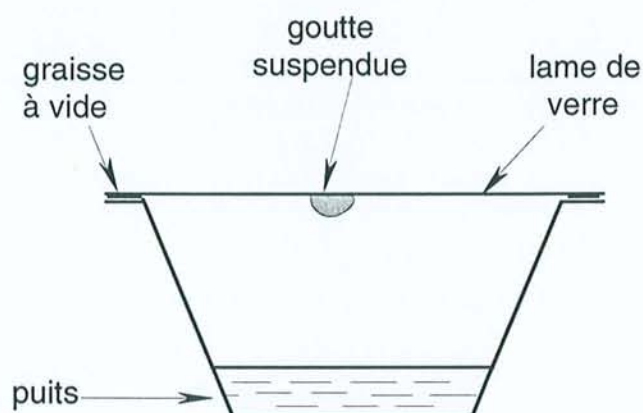


Figure B2 : Méthode de la goutte suspendue.

BI Ic Macrosemis

Il est souvent impossible d'obtenir directement des cristaux de taille suffisante dans les gouttes, soit parce que la nucléation initiale est trop abondante, soit parce que les faces des microcristaux sont polluées, ou encore parce que la zone de nucléation est trop éloignée de la

zone de croissance. La technique alors utilisée est celle du macrosemis qui permet de s'affranchir de la phase de nucléation. Elle consiste à déposer dans une nouvelle goutte sursaturée en protéine un petit cristal dans le but d'en poursuivre la croissance. Le petit cristal prélevé, entouré de germes non visibles, doit être lavé dans la solution du réservoir. Ceci permet d'éviter la croissance de microcristaux en parallèle à celle du cristal sélectionné. C'est ainsi qu'en théorie la croissance d'un seul cristal est obtenue. Ce lavage, en dépolluant les faces des cristaux, permet également d'assurer une reprise de croissance.

BIId Cristallisation en gel

Des essais de cristallisation en gel par diffusion de vapeur sont parfois utiles pour améliorer la qualité des cristaux. Le gel est composé de macropores (rayon 1000 Å environ) reliés par des canaux (rayon 100 Å environ). Le volume de solution protéique dans un macropore est très petit par rapport à celui d'une goutte suspendue en solution, ce qui évite la sédimentation et diminue la convection naturelle. Le cristal pris au piège dans chaque macropore peut ainsi grossir librement sans contrainte de voisins. Les contraintes de paroi proviennent du gel lui-même, mais, du fait de sa faible élasticité, celui-ci se fissure en formant une cavité adéquate pour le cristal. Les paramètres (pH, concentration du précipitant, ...) influençant la cristallisation en milieu liquide sont quasiment les mêmes dans un gel. Les conditions de nucléation peuvent néanmoins être légèrement différentes. Par exemple, dans le cas des gels d'agarose, le nombre de cristaux croît avec la concentration en agarose dans la goutte. Le gel d'agarose est le gel le mieux adapté dans des expériences de diffusion de vapeur en gouttes suspendues (Ducruix & Giegé, 1992).

BIII ACQUISITION ET INTEGRATION DES DONNEES DE DIFFRACTION

BIIIa Introduction

Les cristaux obtenus sont soumis à un rayonnement X dont la brillance de la source doit être suffisante pour minimiser le temps d'exposition et surtout pour obtenir des mesures à haute

résolution. La mesure des intensités diffractées, réflexion par réflexion, au moyen d'un diffractomètre automatique, est beaucoup trop longue car le rayonnement X conduit généralement à des dommages irréversibles au niveau des cristaux. Même si les nouvelles techniques de mesure au grand froid (100 K) (Rodgers, 1994) permettent d'éviter la dégradation des cristaux, le nombre très élevé de réflexions à mesurer a conduit les biocristallographes à utiliser d'autres techniques pour réaliser leurs expériences de diffraction. La méthode employée pour mesurer un maximum de réflexions en un minimum de temps est la méthode d'oscillation (Arndt, 1968). La géométrie utilisée est celle du cristal tournant. Le cristal, monté sur une tête goniométrique, est placé suivant un axe (axe ϕ) perpendiculaire au faisceau direct. Du fait des grands paramètres de maille dans les cristaux de macromolécules, une petite oscillation du cristal autour de cet axe ϕ permet de positionner beaucoup de plans réciproques en condition de diffraction. Pendant l'oscillation, les intensités diffractées par ces plans sont mesurées sur un détecteur plan perpendiculaire au faisceau direct.

BIIIb Sources de rayons X

Le cristallographe utilise un rayonnement X provenant de deux phénomènes physiques :

- le rayonnement émis par un métal bombardé par des électrons ;
- le rayonnement synchrotron émis par des électrons ou des positrons de haute énergie circulant dans un anneau.

La première méthode est accessible dans les laboratoires alors que le rayonnement synchrotron n'est disponible que dans des centres très spécialisés nécessitant des moyens très importants.

Dans l'étude cristallographique des mutants de la GAPDH, des générateurs de rayons X à anode tournante ont été utilisés. Des électrons sont accélérés par une forte différence de potentiel (45 kV) entre l'anode tournante et la cathode dont le filament de tungstène est chauffé par effet Joule. Si l'énergie des électrons est suffisamment élevée, le métal bombardé émet en plus du rayonnement de freinage un spectre de radiations X caractéristique (transition des

niveaux atomiques L et M). Ce spectre, non monochromatique, composé principalement des raies $K\alpha$ (transition du niveau L) et $K\beta$ (transition du niveau M), doit être filtré pour obtenir la raie la plus intense du métal anodique. La raie $K\alpha$ ($\lambda=1.5418 \text{ \AA}$), émise par le cuivre, est très utilisée par le biocristallographe. Sa grande longueur d'onde permet un bon compromis entre l'absorption et la diffraction à haute résolution des cristaux. En effet l'intensité diffractée est proportionnelle à λ^3 . La longueur d'onde du cuivre donne également une meilleure séparation des taches diffractées que celle des autres métaux anodiques couramment utilisés (Mo raie $K\alpha$ $\lambda=0.709 \text{ \AA}$; Ag raie $K\alpha$ $\lambda=0.561 \text{ \AA}$). Cette caractéristique est très importante dans la diffraction des cristaux des macromolécules dont les longueurs de maille excèdent souvent $100 \text{ \AA}$.

BIIIc Monochromateurs

Il existe trois méthodes classiques pour monochromatiser un spectre X :

a) Placer un filtre absorbant les radiations $K\beta$. Dans le cas du cuivre le métal absorbant utilisé est le nickel. Cette technique simple ne donne pas de résultats satisfaisants car le filtre laisse passer des radiations $K\beta$ s'il est trop mince et absorbe trop de radiations $K\alpha$ si on augmente son épaisseur.

b) Soumettre une lame cristalline au rayonnement incident (loi de Bragg) pour sélectionner la radiation $K\alpha$. Les lames les plus utilisées sont essentiellement le graphite, le silicium et le germanium. Ce système optique, très employé en radiocristallographie des petites molécules, produit un excellent rapport signal sur bruit. Cependant, le faisceau incident obtenu est toujours divergent. Cette caractéristique devient problématique pour l'étude des cristaux de macromolécules pour lesquels les spots de diffraction sont très proches les uns des autres et risquent de ce fait d'être en recouvrement.

c) Focaliser le faisceau avec deux miroirs perpendiculaires plans. Les indices de réfraction des deux miroirs dépendant de la longueur d'onde sont inférieurs à 1. pour les rayons

X. Le rayon incident est réfléchi successivement sur les deux miroirs sous un angle d'incidence critique θ_c permettant une réflexion totale. Ce système de double miroir produit un faisceau X fin, intense, ayant un angle de divergence très étroit. De plus, un bon réglage permet de produire un faisceau de profil uniforme jusqu'à cinq fois plus brillant que celui obtenu avec une lame de graphite. Ces caractéristiques sont particulièrement bien adaptées à la cristallographie des macromolécules (Harrison, 1968).

BIId Détecteurs

BIId1 Système plaque image

Le système 'plaque image' est capable de stocker de l'énergie provenant des photons X pour ensuite restituer cette énergie sous forme de lumière ($\lambda=390$ nm) (Ameniya *et al.*, 1988). Une plaque image fait environ 0.5 mm d'épaisseur. Elle est composée d'une couche de cristaux inorganiques (BaFBr:Eu^{2+}) photostimulables noyés dans un dépôt organique et déposée sur un support plastique flexible. Les atomes de brome sont introduits par dopage et forment des centres colorés.

Quand la plaque est exposée au rayonnement X une partie des ions Eu^{2+} ionisés en Eu^{3+} libère des électrons dans la bande de conduction du cristal. Ces électrons libérés vont être captés par les atomes de brome qui vont ainsi stocker de l'énergie.

L'énergie stockée doit être ensuite récupérée. Pour cela, la plaque est exposée à la lumière visible d'un laser He-Ne pour libérer les électrons piégés par les atomes de brome. Les électrons désexcitent les atomes Eu^{3+} en Eu^{2+} en émettant une lumière bleue ($\lambda=390$ nm). Ce phénomène de luminescence photostimulée est désigné par l'effet PLS (Photo-Stimulated Luminescence). Les photons sont émis proportionnellement au nombre de photons X absorbés par chaque pixel de la plaque. Ils sont collectés sur un photomultiplicateur.

La plaque, une fois lue, peut être réutilisée après effacement par une exposition à la lumière blanche. La figure B3 résume un cycle de mesure avec le système plaque image. Le temps de lecture et d'effacement est d'environ 5 minutes pour une plaque circulaire de 30 cm.

Avec ce système, lors d'une collecte, le pas d'oscillation pour un cristal de protéine est d'environ 1°.

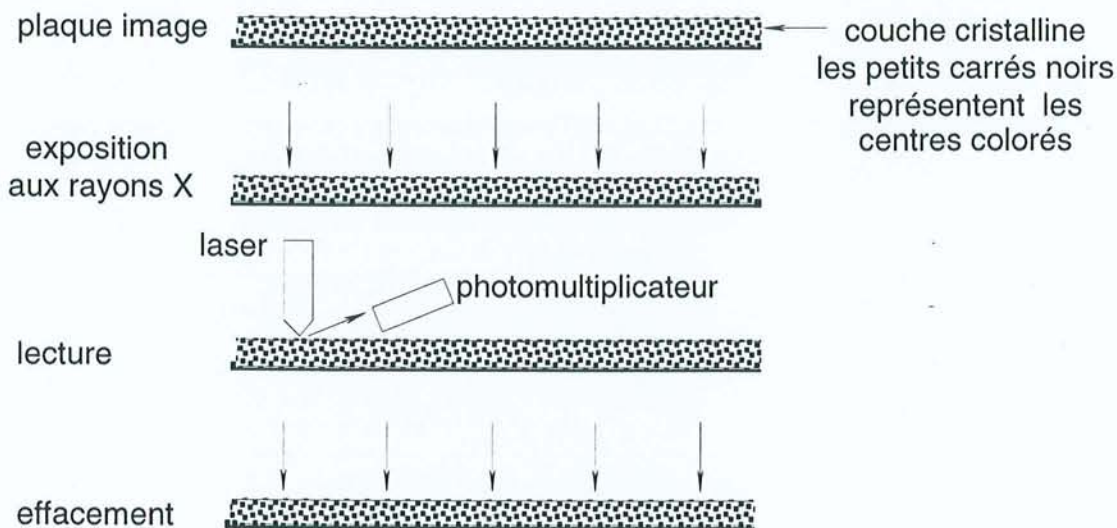


Figure B3: Principe du système plaque image.

Les caractéristiques de ce système sont très intéressantes:

- La dynamique importante des plaques images (intervalle $1:10^4-10^5$) permet de mesurer les réflexions faibles et fortes en même temps ;
- Le développement des plaques avec une surface de grande taille permet l'acquisition de données à haute résolution ;
- Ces plaques, également sensibles aux courtes longueurs d'onde, peuvent être utilisées avec un rayonnement synchrotron (Drenth, 1994).

Pour le présent travail, nous avons utilisé le système de Mar Research composé d'une plaque circulaire de 30 cm de diamètre (LCM, Institut de Biologie Structurale, Grenoble). Pendant la lecture, la plaque tourne autour de son axe tandis que le laser balaie verticalement. La dimension des pixels est de $150 \times 150 \mu\text{m}^2$.

Nous avons également utilisé dans notre laboratoire un système DIP2030-Nonius composé de deux plaques de 30 cm de diamètre. Ce système volumineux permet d'exposer une image pendant que l'autre est lue et effacée. La lecture des plaques se fait de façon spirale et la dimension des pixels est de $100 \times 100 \mu\text{m}^2$.

BIIId2 Système FAST (ENRAF NONIUS)

Le système FAST est un détecteur bidimensionnel basé sur un système vidéo où les rayons diffractés sont collectés sur un écran fluorescent. Les photons X sont convertis en photons visibles. Ces photons visibles sont amplifiés avec un intensificateur d'image, puis emmagasinés dans un tube vidéo qui transforme l'information en un signal électrique. Cette technique permet de mesurer très rapidement l'intensité des réflexions.

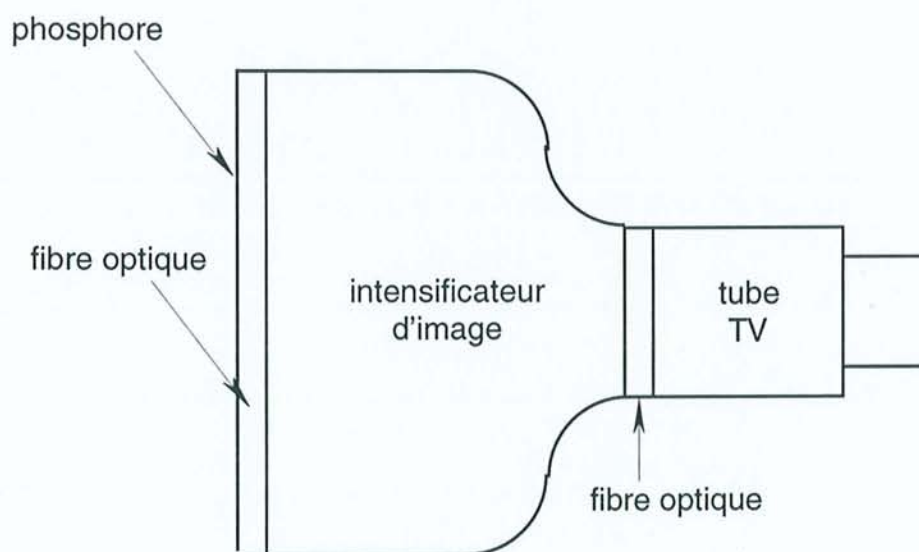


Figure B4 : Schématisation d'un système FAST.

Avec ce système, lors d'une collecte, le pas d'oscillation du cristal est d'environ 0.1° , ce qui permet d'échantillonner les spots de diffraction. Contrairement au système plaque image, l'information obtenue sur les profils d'intensité est donc ici tridimensionnelle. Le bruit

électronique limite la dynamique de l'appareil dans un intervalle de $1:10^3$. Il existe depuis quelques années un nouveau détecteur similaire dans lequel le tube vidéo est remplacé par un CCD (Charge-Coupled Device). Ce nouveau système possède une dynamique bien supérieure à celle du système FAST (Stanton *et al.*, 1992).

BIIIe Traitement des données

BIIIe1 Indexation

Les logiciels DENZO (Otwinowski, 1993) et XDS (Kabsch, 1988, 1993) sont capables d'indexer automatiquement les réflexions (en provenance d'un détecteur) pour trouver la maille du cristal.

Le logiciel localise tout d'abord les pics les plus intenses d'un cliché de diffraction. Les programmes XDS et DENZO utilisent ensuite des algorithmes différents pour déterminer la maille du cristal :

XDS utilise une méthode (Kabsch, 1993) basée sur le calcul des vecteurs reliant les pics sélectionnés suivi d'une recherche des trois plus petits vecteurs réciproques linéairement indépendants.

DENZO utilise un nouvel algorithme d'indexation travaillant dans l'espace direct, basé sur une recherche complète de tous les vecteurs de l'espace direct dont les normes appartiennent à un intervalle défini par l'utilisateur. Le programme détermine ensuite le meilleur triplet irréductible.

Les logiciels proposent ensuite une liste de solutions (14 pour DENZO et 42 pour XDS) avec un facteur d'accord pour chaque indexation possible. La vérification de la solution proposée se fait par affinement des paramètres de la géométrie du système (position du faisceau par rapport au cristal, distance cristal-détecteur, paramètres cristallins...). Les facteurs de qualité de l'affinement sont basés sur l'accord existant entre les positions des pics prédits et des pics observés. Le programme connaît alors toutes les relations géométriques existant entre le réseau direct et le réseau réciproque projeté sur la plaque du détecteur bidimensionnel.

BIIIe2 Intégration

Méthode de la 'shoe box'

La technique de la "shoe box" permet de déterminer l'intensité d'une réflexion en plaçant le spot diffracté dans une boîte bi- ou tridimensionnelle. La boîte doit être suffisamment volumineuse pour pouvoir estimer correctement le bruit de fond : le nombre de pixels utilisé pour calculer le bruit doit être supérieur au nombre de pixels occupés par le pic de diffraction. L'intensité intégrée est calculée en sommant l'intensité des pixels du pic diminuée du bruit de fond estimé. Cette approche ne donne pas de résultats dans l'intégration des réflexions faibles pour lesquelles la frontière entre le bruit de fond et le pic est difficile à déterminer.

Méthode du 'profile fitting'

Cette méthode suppose que des réflexions voisines (aussi bien faibles que fortes) possèdent le même profil normalisé (Diamond, 1969). L'intégration des réflexions se fait alors en deux étapes. Dans la première étape le logiciel prédit un profil moyen pour la réflexion à intégrer à partir des profils de réflexions fortes voisines. Dans la deuxième étape, il s'agit d'appliquer au mieux ce profil moyen sur le pic à intégrer, en minimisant le carré de la différence entre les profils prédit et observé :

$$\sum_i \frac{(M_i - P_i)^2}{V_i}$$

Equation B1

M_i est le signal observé en un pixel i de la 'shoe box'.

$P_i = B_i + C * p_i$ où :

B_i , le bruit de fond prédit, varie linéairement en fonction de la position du détecteur.

C est le paramètre à affiner (C correspond à l'intensité de la réflexion si le profil moyen est normalisé. Dans ce cas $\sum p_i = 1$).

V_i est variance du signal sur le pixel i .

L'intensité intégrée s'écrit alors sous la forme pondérée suivante:

$$I = \frac{\sum_i \frac{P_i}{V_i} (M_i - B_i)}{\sum_i \frac{P_i^2}{V_i}}$$

Equation B2

BIV RECHERCHE DES PHASES

L'expérience de diffraction permet d'obtenir les intensités des réflexions, donc les normes des facteurs de structure de la macromolécule étudiée. La connaissance des normes n'est pas suffisante pour déterminer la position des atomes dans la maille car il manque les phases des facteurs de structure. Même s'il existe de nombreuses méthodes pour déterminer ces phases, cela reste une opération très délicate et très difficile. Les méthodes les plus employées sont le remplacement isomorphe et le remplacement moléculaire. Dans l'étude des mutants de la GAPDH, seul le remplacement moléculaire a été utilisé.

BIVa Remplacement moléculaire

'La méthode du remplacement moléculaire dépend de trois aspects esthétiques, plaisants et élégants de la cristallographie: symétrie, transformation moléculaire et théorème de convolution. C'est donc un domaine où les arguments intuitifs basés sur des concepts ont une place à côté d'analyses mathématiques rigoureuses' (Blundell & Tickle, 1985).

Cette technique peut être utilisée dans deux cas de la détermination d'une structure :

- Deux protéines fonctionnellement identiques sont souvent structuralement proches.

Connaissant la structure tridimensionnelle de l'une des deux, il est possible de déterminer la structure inconnue à l'aide du remplacement moléculaire. Cette technique a été utilisée avec

succès, par exemple, dans la détermination des structures de GAPDH de *Escherichia coli* (Olivier *et al.*, 1995) et de *Leishmania mexicana* (Kim *et al.*, 1995) à partir de celle de *Bacillus stearothermophilus*. Son principe est de positionner une macromolécule connue dans l'unité asymétrique d'une macromolécule inconnue pour phaser le modèle.

- Les protéines possédant une structure quaternaire cristallisent souvent sous forme n-mérique dans l'unité asymétrique. On parle dans ce cas de symétrie non cristallographique entre les différentes sous-unités. Le remplacement moléculaire a alors pour objectif de déterminer les éléments de symétrie entre les sous-unités pour étendre les phases du modèle.

Le positionnement d'une macromolécule connue dans l'unité asymétrique d'une macromolécule inconnue peut se décomposer en deux étapes. Dans un premier temps il faut orienter la molécule connue dans le groupe d'espace de la molécule inconnue, et donc lui faire subir une rotation. Dans un deuxième temps, la molécule connue correctement orientée doit être positionnée dans l'unité asymétrique. Il faut donc lui faire subir une translation.

BIVa1 Recherche de la rotation

Cette procédure a pour but de déterminer la matrice de rotation C permettant, soit d'orienter la molécule connue dans l'unité asymétrique de la molécule inconnue, soit de superposer des sous-unités entre elles. Dans la recherche des phases d'un nouveau modèle, la distribution de Patterson des vecteurs intramoléculaires de la structure connue est calculée et la distribution de Patterson de la molécule inconnue est déterminée à partir des facteurs de structure mesurés. La superposition des deux distributions se fait ensuite en maximisant une fonction de corrélation (équ. B3) dépendant des deux fonctions de Patterson et de la matrice de rotation C . Cette fonction est appelée fonction de rotation croisée dans le cas de la recherche des phases d'un nouveau cristal (P_1 = inconnue, P_2 = connue). Cette même fonction s'appelle fonction d'autocorrélation dans la recherche de symétries non cristallographiques ($P_1 = P_2$ = inconnue).

$$R(\alpha, \beta, \gamma) = \int_V P_1(u) P_2(Cu) du$$

Equation B3

V, volume d'intégration de la fonction de Patterson, est considéré habituellement comme sphérique. Il doit inclure les vecteurs intramoléculaires **u** (recherche d'une structure inconnue) ou intra sous-unités (recherche de symétrie non cristallographique), selon le cas.

$P_1(u)$ est la fonction de Patterson de la molécule 1 ou de la sous unité 1.

$P_2(Cu)$ est la fonction de Patterson de la molécule 2 ou de la sous unité 2 dans laquelle on applique aux vecteurs **u** la matrice de rotation C construite sur α, β, γ .

Le maximum de la fonction de corrélation est recherché en calculant la fonction R dans tout l'espace des trois rotations, pour déterminer la meilleure superposition entre les deux modèles. Deux paramètres sont très importants dans le calcul de cette fonction de corrélation :

Le bon choix du rayon de la sphère du volume de la Patterson permet d'exclure les vecteurs intermoléculaires du domaine d'intégration.

L'estimation du degré de ressemblance entre les deux macromolécules voisines permet de choisir la troncature à basse et à haute résolution.

Le programme AMoRe (Navaza, 1994) à été utilisé dans cette étude. Ce logiciel apporte des améliorations dans le calcul de la fonction de corrélation, basées sur le développement des fonctions de Patterson en une somme d'harmoniques (Crowther, 1972). La fonction de corrélation s'écrit sous la forme :

$$R(\alpha, \beta, \gamma) = \sum_{l=0}^{\infty} \sum_{m, m'=-l}^l D_{m, m'}^l C_{m, m'}^l (P_1, P_2)$$

Equation B4

$D_{m, m'}^l$, sont les matrices des représentations irréductibles du groupe de rotation et sont fonction des variables angulaires,

$C_{m, m'}^l$ dépend des fonctions de Patterson des molécules 1 et 2 et de la définition du domaine sphérique d'intégration. Le calcul de $C_{m, m'}^l$ par intégration numérique (Navaza, 1987) donne une meilleure précision de la fonction de rotation qu'un calcul en expansion de type Fourier-Bessel (méthode de Crowther).

BIVa2 Recherche de la translation

Il existe plusieurs méthodes pour résoudre le problème de la translation. L'une d'elles consiste à calculer l'indice d'accord Rf (et/ou le coefficient de corrélation C) existant entre les facteurs de structure calculés de la protéine connue centrée en un point M de la maille et ceux observés de la protéine inconnue. Cette procédure systématique translate pas à pas la protéine connue dans toute la maille. Il est ainsi possible d'obtenir l'indice d'accord Rf (et/ou le coefficient de corrélation C) donnant le meilleur positionnement de la molécule connue orientée dans la maille de la structure recherchée.

Le facteur d'accord Rf s'exprime sous la forme :

$$Rf = \sum_H \|Fobs\| - k \|Fcalc\| / \sum_H \|Fobs\|$$

Equation B5

Fobs facteur de structure observé.

Fcalc facteur de structure calculé.

k facteur d'échelle.

et le coefficient linéaire de corrélation C (Fujinaga & Read, 1987) sous la forme:

$$C = \frac{\sum_H \left(|Fobs|^2 - \overline{|Fobs|^2} \right) \times \left(|Fcalc|^2 - \overline{|Fcalc|^2} \right)}{\left[\sum_H \left(|Fobs|^2 - \overline{|Fobs|^2} \right)^2 \sum_H \left(|Fcalc|^2 - \overline{|Fcalc|^2} \right)^2 \right]^{1/2}}$$

Equation B6

Ce coefficient de corrélation présente l'avantage d'être indépendant de la mise à l'échelle des facteurs de structure observés (sans facteur *k*). Cette technique de déplacement discret de la

molécule connue dans l'unité asymétrique de la protéine inconnue est peu utilisée car elle demande beaucoup de temps de calcul.

Une autre approche de ce problème de translation, moins gourmande en temps de calcul, est basée sur la maximisation d'un produit de deux fonctions de Patterson (Crowther & Blow, 1967). L'une correspond à la fonction de Patterson observée et l'autre correspond à une fonction de Patterson croisée, calculée à partir de la molécule connue préorientée ('modèle' dans l'Equation B7). Les vecteurs utilisés dans le calcul de la fonction de Patterson croisée sont les vecteurs intermoléculaires, c'est-à-dire les vecteurs reliant deux molécules cristallographiquement symétriques. La fonction de translation peut alors s'écrire sous la forme :

$$T(\mathbf{t}) = \int P_{12}(\mathbf{u}, \mathbf{t}) P(\mathbf{u}) d\mathbf{u}$$

Equation B7

$P(\mathbf{u})$ représente la fonction de Patterson observée du cristal.

$P_{12}(\mathbf{t}, \mathbf{u})$ exprime la fonction de Patterson croisée du modèle translaté. Celle-ci inclut uniquement les vecteurs intermoléculaires des molécules 1 et 2 distantes de $\mathbf{t}$ et reliées par un élément de symétrie.

Le programme AMoRe (Navaza, 1994) utilise cette fonction T de Crowther & Blow (1967) pour sélectionner un ensemble de pics bien résolus. Pour chacun des pics, il calcule un facteur d'accord R et un coefficient de corrélation C . Les solutions trouvées par le programme AMoRe sont donc accompagnées de ces deux facteurs de qualité.

BV AFFINEMENT DU MODELE

Le remplacement moléculaire donne au départ un modèle approximatif qui doit être affiné. Ce processus d'affinement a pour but de trouver le meilleur accord entre les facteurs de structure calculés et expérimentaux. Le facteur d'accord R_f (équ. B7) permet de guider

l'affinement. Dans la cristallographie des petites molécules, le nombre de paramètres à affiner ($x, y, z, B_{\text{aniso}}$) est souvent très inférieur au nombre de réflexions mesurées, et il est commun de trouver un rapport (nombre d'observations / nombre de paramètres) de l'ordre de 5. Dans ce domaine, les cristaux diffractent souvent à plus de 1 Å de résolution. La technique d'affinement utilisée dans ce système surdéterminé est celle des moindres carrés. La cristallographie des macromolécules ne permet pas aujourd'hui d'obtenir des résolutions aussi élevées, sauf cas particuliers (Dauter *et al.*, 1995) Ce même rapport dépasse rarement 3 et, par exemple, une mesure de la GAPDH à 2.2Å de résolution ne donne qu'un rapport de 1.8. Les données géométriques obtenues sur des petites molécules (liaisons, angles) permettent d'ajouter des 'observations' supplémentaires dans le processus d'affinement. De même, la symétrie non cristallographique permet d'étendre les phases du modèle. Il existe différents logiciels utilisant ces observations supplémentaires dans un affinement par moindres carrés restreint (PROLSQ ; Hendrickson, 1985) ou dans un affinement d'une pseudo énergie potentielle par gradient conjugué (Xplor ; Brünger, 1992b). Ces deux techniques conduisent toujours au minimum local le plus proche. L'examen visuel du modèle permet de 'sauter' hors de ce minimum en corrigeant les imperfections structurales importantes. Cependant, la progression est lente car elle demande un grand nombre de cycles d'affinement et d'analyses graphiques. Brünger *et al.* (1990) ont développé une méthode d'affinement très efficace utilisant la dynamique moléculaire. Cette procédure permet de 'sauter' des barrières d'énergie plus importantes et donc de limiter le nombre d'interventions manuelles.

BVa Affinement en corps rigide

Le remplacement moléculaire ne donne jamais directement une solution comportant un modèle parfaitement orienté et translaté. En considérant ce modèle comme un ou plusieurs corps rigides, cette première étape d'affinement permet de repositionner ces corps dans la maille. Une chute du facteur d'accord R_f (équ. B5) et/ou une forte augmentation du coefficient de corrélation C (équ. B6) confirment l'exactitude de la solution donnée par le remplacement moléculaire.

BV2 Affinement par gradient conjugué

Cette méthode utilisée dans le logiciel XPLOR (Brünger, 1992b) est basée sur la minimisation d'une pseudo énergie potentielle (*Etot*) composée d'un terme provenant de la mécanique moléculaire (*Emol*) et d'un autre terme, mis à l'échelle par rapport au premier, provenant de la différence entre les facteurs de structure observés et calculés (*Ecris*). L'expression à minimiser s'écrit alors sous la forme:

$$Etot = Emol + w_A Ecris$$

Equation B8

w_A (rapport de la norme du gradient de *Emol* sur la norme du gradient de *Ecris*) correspond au poids idéal donné aux contraintes cristallographiques. La minimisation de la pseudo énergie potentielle est toujours précédée d'une détermination du poids idéal w_A (sous-étape Check dans Xplor).

Le terme énergétique provenant de la mécanique moléculaire *Emol* s'exprime sous la forme d'une somme d'énergies géométrique, électrostatique et de Van der Waals:

$$\begin{aligned} Emol = & \sum_{\text{liaisons}} 1/2 K_b (b - b_0)^2 + \sum_{\text{angles}} 1/2 K_\theta (\theta - \theta_0)^2 + \sum_{\text{torsions}} 1/2 K_\zeta (\zeta - \zeta_0)^2 \\ & + \sum_{\text{dièdres}} 1/2 K_\phi [1 + \cos(n\phi + \delta)] + \sum_{\text{électrostatiques}} q_i q_j / (4\pi\epsilon_0 \epsilon_r r_{ij}) \\ & + \sum_{\text{Van-der-Waals}} [C_{12}(i, j) / r_{ij}^{12} - C_6(i, j) / r_{ij}^6] \end{aligned}$$

Equation B9

$\sum_{\text{liaisons}} 1/2 K_b (b - b_0)^2$ énergie d'oscillation harmoniques des liaisons covalentes.

K_b constante de force.

$b - b_0$ allongement de la liaison (b_0 distance de référence).

$\sum_{\text{angles}} 1/2 K_{\theta} (\theta - \theta_0)^2$ énergie d'oscillation harmonique des angles de valence.

K_{θ} constante de force.

$\theta - \theta_0$ déplacement angulaire (θ_0 angle de référence).

$\sum_{\text{torsions}} 1/2 K_{\zeta} (\zeta - \zeta_0)^2$ énergie d'oscillation des angles de torsion impropres
(angle de torsion défini sur 4 atomes non liés).

k_{ζ} constante de force.

$\zeta - \zeta_0$ déplacement angulaire (ζ_0 angle de torsion de référence).

$\sum_{\text{dièdres}} K_{\phi} [1 + \cos(n\phi + \delta)]$ énergie potentielle des angles dièdres.

K_{ϕ} barrière de torsion.

n périodicité.

ϕ angle dièdre.

δ phase déterminant le point zéro de la rotation.

$\sum_{\text{électrostatiques}} q_i q_j / (4\pi\epsilon_0 \epsilon_r r_{ij})$ énergie électrostatique.

q_i, q_j charges partielles des atomes i et j .

r_{ij} distance entre les centres atomiques i et j .

ϵ_r constante diélectrique du milieu (constante ou fonction de la distance r_{ij}).

$\sum_{\text{Van-der-Waals}} [C_{12}(i, j)/r_{ij}^{12} - C_6(i, j)/r_{ij}^6]$ énergie de Van der Waals

r_{ij} distance entre les centres atomiques i et j .

$C_{12}(i, j), C_6(i, j)$ constantes de Van der Waals (fonction du type des atomes i et j).

Le terme cristallographique *E_{crist}* s'écrit:

$$E_{\text{crist}} = \sum w_h \left[|F_{\text{obs}}(h)| - k |F_{\text{calc}}(h)| \right]^2$$

Equation B10

w_h poids individuel de chaque réflexion (fonction de l'écart type sur la mesure)

BVc Dynamique moléculaire

La dynamique moléculaire est une méthode d'étude de l'espace conformationnel d'un système moléculaire en mouvement par l'intégration de l'équation de Newton. L'énergie potentielle est connue grâce au champ de forces associé au système. Brünger *et al.* (1987) ont adapté cette technique utilisée en modélisation moléculaire à l'affinement structural des

macromolécules. Cette méthode a l'avantage d'avoir un grand rayon de convergence. Elle surclasse les affinements classiques (moindres carrés ou gradient conjugué) en permettant des sauts énergétiques jusqu'alors infranchissables.

Brünger *et al.* (1989) ont montré que le recuit simulé permet une utilisation optimale de la dynamique moléculaire dans l'affinement cristallographique. Cette méthode consiste à augmenter suffisamment la température du système pour pouvoir franchir les minima énergétiques locaux, puis à refroidir lentement ce système pour approcher le minimum énergétique. Le recuit simulé est toujours suivi d'un affinement classique de minimisation d'énergie pour atteindre le minimum global.

Cette méthode de dynamique moléculaire est particulièrement adaptée à notre étude de mutants de la GAPDH. Le modèle initial utilisé dans l'affinement est toujours celui de la protéine de type sauvage. Cette méthode doit nous permettre d'observer les changements structuraux induits par les mutations de quelques résidus. Par exemple, McRee (1993) montre très bien dans son livre 'Practical Protein Crystallography' (p. 279-284) l'importance du recuit simulé dans l'affinement d'un mutant de la zinc superoxide dismutase (SOD). Dans ce cas, seule une dynamique moléculaire à 2000°C a pu mettre en évidence la nouvelle conformation du résidu muté. La géométrie autour de ce résidu était tellement bien affinée dans la protéine de type sauvage qu'il n'était pas possible de sortir du minimum local avec les techniques d'affinement classiques.

PARTIE C

RESOLUTION DE STRUCTURES

CI CRISTALLISATION

CIa Introduction

Les conditions de cristallisation de la glycéraldéhyde-3-phosphate déshydrogénase de *B. stearothermophilus* de type sauvage ont été déterminées par Biesecker *et al.* (1977). Les cristaux grossissent sous forme de prismes monocliniques dans le groupe d'espace $P2_1$, avec un tétramère par unité asymétrique. La structure a été déterminée à haute résolution (1.8Å) (Skarzynski *et al.*, 1987).

La préparation de l'enzyme pour la cristallisation a été effectuée dans les mêmes conditions pour les mutants D32G-S et B-S complexés avec NAD et NADP. Deux formes cristallines (monoclinique et orthorhombique) ont été obtenues.

CIb Préparation de la protéine

La production et la purification des mutants D32G-S et B-S ont été réalisées dans le laboratoire d'enzymologie et de génie génétique de l'université Henri Poincaré de Nancy (Pr. G. Branlant). Les méthodes employées ont été décrites par Rahuel-Clermont (1994).

L'échantillon protéique (env. 4 mg), précipité dans le sulfate d'ammonium, est placé dans une unité de filtration pour être dialysé trois fois contre du tampon phosphate 200 mM à pH 7.2 (contenant 1 mM d'EDTA, 1 mM de DTT et 30 mM de NAD(P)) par centrifugation dans un rotor à angle fixe à 7500 g. Pour évaluer la concentration de la protéine, l'excédent en NAD(P) doit être éliminé. La solution résultante est alors dialysée au moins 4 fois contre le même tampon sans NAD(P). La concentration de la protéine est ensuite déterminée par mesure de la densité optique à 260 nm et 280 nm. Enfin, la solution est ajustée à 7 mg/ml en protéine, 2M en sulfate d'ammonium, dans un tampon phosphate 50 mM pH 6.9, EDTA 1 mM, DTT 1 mM, NAD(P) 30 mM. Les essais de cristallisation ont été réalisés au laboratoire en utilisant le principe de la diffusion de vapeur en goutte suspendue. Les gouttes de 10 µl ont été équilibrées

contre un réservoir d'un volume de 1 ml. Ce réservoir est composé de sulfate d'ammonium à 2.3M (agent précipitant), tamponné par du phosphate 50 mM pH 6.9.

C1c Forme monoclinique

Les cristallisations ont été réalisées dans une pièce thermostatée à 20°C. Pour les mutants D32G-Sn, D32G-Sp et B-Sn, les cristaux obtenus se présentent sous la forme de prismes (Photo page 78). Les conditions de cristallisation ont dû être légèrement modifiées pour chaque mutant. Le protocole suivant a été suivi :

Selon Biesecker *et al.* (1977), dans les conditions de cristallisation décrites ci-dessus, les cristaux de l'enzyme de type sauvage apparaissent au bout d'une semaine. Dans notre cas, aucun cristal n'est apparu au bout d'une semaine (2.3 M en SA à l'équilibre). Les gouttes suspendues ont alors été séparées en deux ensembles : un quart constitue l'ensemble 'test' et les trois quarts restants l'ensemble 'suiveur'. La concentration des puits de l'ensemble 'test' est d'abord augmentée de 2.3 M à 2.35 M. Si, après 4 jours, aucun germe n'est observé, la concentration en sulfate d'ammonium est augmentée de 0.05 M dans tous les puits (la concentration est donc de 2.4 M pour l'ensemble 'test' et 2.35 M pour l'ensemble 'suiveur'). Ce procédé est réitéré jusqu'à l'apparition de cristaux dans l'ensemble 'test'. Les concentrations finales de l'ensemble 'test' sont comprises entre 2.4 M et 2.55 M suivant le mutant. Les gouttes de l'ensemble 'test' n'ont donné lieu qu'à des multinucléations. Afin d'obtenir des cristaux de dimension suffisante, la technique du macrosems (Ducruix & Giegé, 1992) a alors été pratiquée en utilisant comme germe un microcristal (0.05 mm dans la plus grande dimension) prélevé dans une goutte de l'ensemble 'test'. Le germe, lavé au préalable dans des bains de sulfate d'ammonium (à la concentration des réservoirs de l'ensemble 'test'), est déposé dans une goutte de l'ensemble 'suiveur'. Les cristaux atteignent dans les deux jours des dimensions suffisantes pour les expériences de radiocristallographie.

Les mutants D32G-Sn, D32G-Sp, et B-Sn diffractent, respectivement, jusqu'à 2.5Å, 2.2Å et 2.45Å de résolution (Tableau C1). La forme cristalline de ces trois complexes est

isomorphe de celle de l'holoenzyme de type sauvage : la protéine cristallise dans le groupe d'espace $P2_1$ avec un tétramère dans l'unité asymétrique.

Tableau C1 : Caractéristiques cristallographiques des cristaux des mutants étudiés

	a (Å)	b (Å)	c (Å)	β (°)	groupe d'espace	limite de diffraction (Å)
D32G-Sn	82.95	124.68	82.97	108.63	$P2_1$	2.5
D32G-Sp	82.37	124.27	82.45	108.93	$P2_1$	2.2
B-Sn	82.87	124.80	82.92	108.72	$P2_1$	2.45
type sauvage ^a	82.44	124.10	82.54	108.98	$P2_1$	1.8
B-Sp	134.14	123.96	96.73	90.	$P2_12_12$	2.5
forme 'Holo I' ^b	132.5	126.3	99.6	90.	$P2_12_12$	3.0
Sp ^c	132.5	126.3	99.6	90.	$P2_12_12$	2.7

^a Skarzynski *et al.*, 1987 ; ^b Leslie & Wonacott, 1983 ; ^c Branlant, Rahuel-Clermont & Wonacott, résultats non publiés.

CId Forme orthorhombique

Le mutant B-Sp cristallise, à 20 °C, sous forme de prismes (Photo page suivante). Le protocole utilisé pour cristalliser ce mutant complexé avec $NADP^+$ est identique à celui décrit précédemment. Les cristaux obtenus sont de taille suffisante et diffractent jusqu'à une résolution de 2.5 Å. Ils sont isomorphes de la forme cristalline 'HOLO I' de l'enzyme de type sauvage cristallisée avec un seul cofacteur par tétramère (Leslie & Wonacott, 1983). Branlant & Wonacott (résultats non publiés) ont observé la même forme dans les cristaux du double mutant S complexé avec NADP. Le groupe d'espace de cette forme orthorhombique est $P2_12_12$. Les empilements cristallins des formes monoclinique et orthorhombique sont voisins (cf. § CIIIb2).



CII CONDITIONS DE MESURE

CIIa Triple mutant cocristallisé avec NAD

CIIa1 Enregistrement

L'acquisition des données de diffraction des rayons X sur les cristaux du mutant D32G-Sn a été réalisée avec un détecteur de type 'image plate' (Dip2000, plaque circulaire de 20 cm de diamètre) muni d'un goniomètre kappa. Pour cet enregistrement, un générateur de rayons X à anode tournante (Nonius FR591, anode de Cu, 45 kV, 90 mA) équipé d'un monochromateur vertical en graphite a été utilisé. Le cristal a été placé dans un capillaire. Les mesures ont été effectuées à température ambiante. La distance cristal-détecteur est ajustée afin d'obtenir des taches de diffraction jusqu'au bord de la plaque image. Les cristaux du mutant D32G-Sn diffractent jusqu'à une résolution de 2.5 Å, ce qui correspond à une distance de 13 cm. L'orientation du cristal par rapport au faisceau est quelconque. Durant chaque exposition de 1200 s, le cristal effectue trois oscillations de 1°. L'analyse visuelle des 180 images a montré que les 11 dernières images étaient de mauvaise qualité, probablement en raison de la dégradation du cristal sous l'action des rayons X.

CIIa2 Intégration

Le traitement des données a été réalisé à l'aide du programme Denzo pour l'intégration et du programme Scalepack (Otwinowski, 1993) pour la mise à l'échelle et la réunion des données. Neuf images (de la 161^{ème} à la 169^{ème}) affectant la qualité des mesures ont été ensuite écartées dans la réunion des données. La couverture d'un domaine angulaire effectif de 160° permet de mesurer 45468 réflexions indépendantes avec une redondance moyenne de 2.7. La complétude de ce jeu de données est de 85 %. Le Tableau C2 rend compte de la statistique des données de diffraction en fonction de la résolution. Le facteur d'accord 'Rsym' sur la moyenne des réflexions équivalentes est de 6.8 %.

Tableau C2 : statistiques de l'enregistrement des données du complexe D32G-Sn

Résolution (Å)	Intensité moyenne	Sigma moyen	Rsym %	Nombre de réflexions (*)	Complétude %
4.93	5956	355	3.8	4758	85.3
3.93	6935	371	3.8	5991	85.1
3.43	4052	260	5.2	6155	85.7
3.12	2625	201	6.9	6004	86.0
2.90	1645	161	9.6	5949	84.9
2.73	1125	141	12.9	5915	84.7
2.59	840	132	16.6	6033	83.9
2.50	664	128	20.9	4683	84.0
Total	2501	200	6.8	45468	85.0

(*) Nombre de réflexions indépendantes mesurées.

$$R_{\text{sym}} = \frac{\sum_j |I_j - \langle I \rangle|}{\sum \langle I \rangle}$$

CIib Triple mutant cocristallisé avec NADP

CIib1 Enregistrement

La mesure des données de diffraction du cristal de D32G-Sp a été faite au laboratoire LCM/IBS de Grenoble avec un détecteur de type 'image plate' (MarResearch, plaque circulaire de 30 cm) muni d'un goniomètre phi et d'une source de rayons X provenant d'une anode tournante (RU200, anode de cuivre, 40 kV, 90 mA). Comme pour les cristaux de D32G-Sn la méthode du cristal tournant a été employée pour la collecte des intensités. La distance cristal-détecteur a été fixée à 16 cm. Le pas d'oscillation du cristal a été fixé à 1° et le temps d'exposition à 1200 s. En raison de la dégradation du cristal, seules 140 images ont pu être mesurées.

CIib2 Intégration

L'intégration et la réunion des données ont été faites respectivement à l'aide du programme XDS (Kabsch, 1988, 1993) et de la procédure ROTAVATA de la chaîne de

programmes CCP4 (1994). Sur les 140 images mesurées, seules les 120 premières ont été retenues dans le traitement des données. La complétude des données entre 2.3 et 2.2 Å de résolution est de 77 % en considérant l'ensemble des réflexions mesurées, et de 65 % en ne conservant que les réflexions dont le facteur de structure est supérieur à deux fois la déviation standard estimée. La moyenne des réflexions équivalentes conduit à un facteur d'accord de 9.7 %. Le tableau C3 donne les statistiques de ces mesures.

Tableau C3 : statistiques de l'enregistrement des données du complexe D32G-Sp

Résolution (Å)	Intensité moyenne	Sigma moyen	Rsym %	Nombre de réflexions (*)	Complétude %
6.64	1382	239	6.8	2539	85.3
4.70	1007	116	6.9	4796	89.5
3.83	1251	148	8.3	6266	90.5
3.32	734	101	10.0	7232	89.4
2.97	424	69	12.9	8066	87.2
2.71	255	54	15.5	8668	84.9
2.51	185	46	17.8	9877	81.8
2.35	148	42	19.6	10558	80.8
2.19	133	42	19.4	12184	72.8
TOTAL	435	94	9.7	70186	84.1

(*) Nombre de réflexions indépendantes mesurées.

CIIC Quintuple mutant cocristallisé avec NAD

Les expériences de diffraction sur les cristaux de B-Sn ont été réalisées avec le même appareillage que pour les cristaux de D32G-Sp. Un seul enregistrement a été suffisant pour obtenir une complétude de 96 % jusqu'à 2.45 Å de résolution. Les principales caractéristiques de l'expérience sont les suivantes : distance cristal-détecteur 13 cm, temps d'exposition 1200 s,

pas d'oscillation 1°, domaine angulaire couvert 180°. Le tableau C4 rend compte des statistiques des données de diffraction en fonction de la résolution. La moyenne des réflexions équivalentes donne un facteur d'accord de 7.8 % pour une redondance moyenne de 3.1 .

Tableau C4 : statistiques de l'enregistrement des données du complexe B-Sn

Résolution (Å)	Intensité moyenne	Sigma moyen	Rsym %	Nombre de réflexions (*)	Complétude %
7.65	1682	234	8.4	1875	95.0
5.41	780	85	6.2	3557	99.3
4.42	1271	137	6.1	4532	99.7
3.83	1102	114	6.5	5355	99.5
3.42	706	84	7.4	6111	98.5
3.12	466	61	8.7	6692	98.3
2.89	305	46	10.4	7092	96.5
2.71	211	38	12.5	7228	95.2
2.55	163	33	13.8	8114	93.8
2.45	141	32	14.8	6299	81.7
Total	566	84	7.8	56855	96

(*) Nombre de réflexions indépendantes mesurées.

CIId Quintuple mutant cocristallisé avec NADP

La mesure des intensités diffractées par les cristaux de B-Sp a été réalisée au laboratoire dans les mêmes conditions que pour les cristaux de D32G-Sn. Les principales caractéristiques d'enregistrement sont les suivantes : distance cristal-détecteur 13 cm, temps d'exposition 1200 s, pas d'oscillation 1°, domaine angulaire couvert 90°. Le facteur d'accord 'Rsym' sur l'ensemble des réflexions est de 5.4%. Le tableau C5 donne les statistiques de cette mesure.

Tableau C5 : statistiques de l'enregistrement des données du complexe B-Sp

Résolution (Å)	Intensité moyenne	Sigma moyen	Rsym %	Nombre de réflexions (*)	Complétude %
5.35	6701	183	2.5	5059	85.5
4.26	8148	269	2.8	5111	89.4
3.73	6156	248	3.7	5159	91.4
3.39	4049	231	5.1	5212	92.3
3.15	2768	207	6.9	5208	92.9
2.96	1854	199	9.7	5227	93.3
2.81	1337	183	12.2	5239	93.8
2.69	1039	174	16.2	5228	94.2
2.59	826	165	20.1	5264	94.8
2.50	672	161	24.2	5294	94.8
Total	3324	202	5.4	52001	92.2

(*) Nombre de réflexions indépendantes observées.

CIII PHASAGE

Les quatre structures de mutants ont été déterminées à partir du modèle de l'holoenzyme de type sauvage (Skarzynski *et al.*, 1987; code PDB 1gd1 (Bernstein, 1977)).

CIIIa D32G-Sn, D32G-Sp, B-Sn

Comme les cristaux des mutants D32G-Sn, D32G-Sp, B-Sn sont isomorphes de ceux de l'holoenzyme de type sauvage, les coordonnées de celui-ci ont pu être directement utilisées pour obtenir les premiers jeux de phases.

CIIIb B-Sp

CIIIb1 Remplacement moléculaire

Les coordonnées des structures de la forme 'Holo I' et du mutant Sp n'étant pas disponibles dans la PDB, la méthode du remplacement moléculaire a été utilisée pour phaser les

facteurs de structure du mutant D32G-Sp à l'aide du logiciel AMoRe (Navaza, 1994). Le modèle utilisé pour résoudre cette structure de mutant est celui de l'holoenzyme de type sauvage. La gamme de résolution prise en compte dans ce calcul s'étend de 15 Å à 4 Å. Le rayon de la sphère de Patterson déterminé automatiquement par le programme est de 54.27 Å.

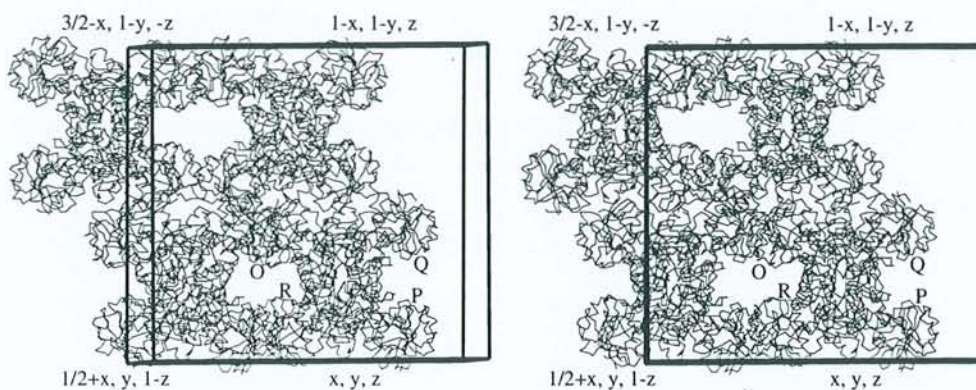
Le calcul de la fonction de rotation révèle que l'axe moléculaire Q est parallèle à l'axe cristallographique b (tableau C6), comme dans la maille cristalline de l'enzyme de type sauvage. La position du tétramère est déterminée par le calcul de la fonction de translation. Finalement, l'affinement en corps rigides permet d'affiner simultanément les fonctions de rotation et de translation. La solution se dégage sans ambiguïté avec un coefficient de corrélation de 76.6 % et un indice résiduel de 28.4 %. La position affinée du tétramère dans l'unité asymétrique en coordonnées fractionnaires est (.42, .25, .25). Les cosinus directeurs des axes moléculaires P , Q et R par rapport aux directions cristallographiques a , b et c sont donnés dans le tableau C6.

Tableau C6 : Cosinus directeurs des axes moléculaires P , Q , R par rapport à a , b , c .

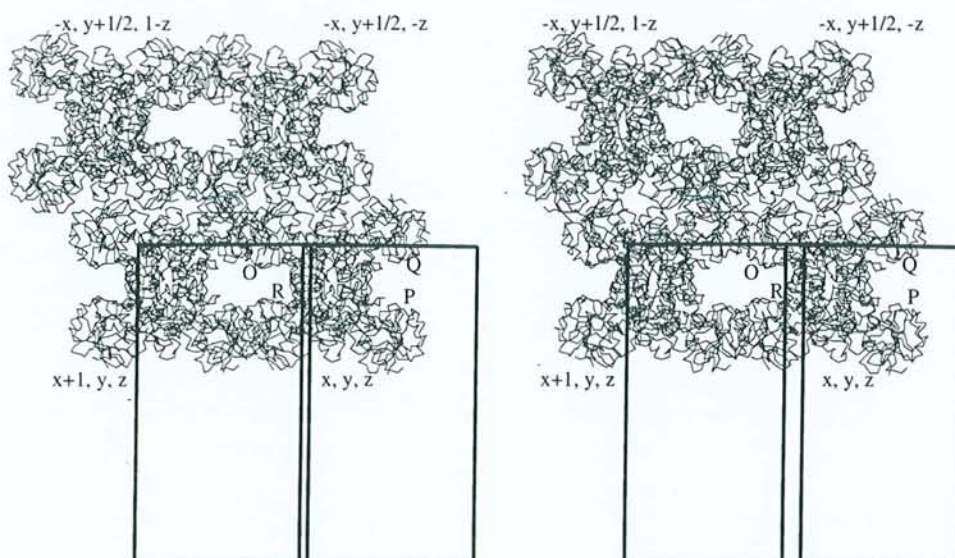
Axes	P	Q	R
a	-0.0395	-0.0904	0.995
b	0.0493	-0.995	-0.0884
c	0.998	0.0455	0.0437

CIIb2 Empilement cristallin

Chaque tétramère est en contact avec 6 tétramères voisins, formant ainsi trois interfaces uniques. Quatorze résidus sont impliqués dans les contacts intermoléculaires donnant lieu à sept liaisons hydrogène : cinq avec le tétramère symétrique déduit de l'axe hélicoïdal parallèle à a et deux avec le tétramère symétrique déduit de l'axe binaire parallèle à c (tableau C7). Les axes moléculaires P , Q et R sont quasiment parallèles respectivement aux axes cristallins c , b et a (tableau C6). Les tétramères adoptent donc tous la même orientation dans le cristal (figure C1a).



a) quatre tétramères localisés en (x, y, z) , $(1-x, 1-y, z)$, $(1/2+x, y, 1-z)$ et $(3/2-x, 1-y, -z)$ sont représentés dans la maille orthorhombique. L'axe binaire // à c est perpendiculaire au plan de la figure et l'axe hélicoïdal // à b est vertical.



b) quatre tétramères localisés en (x, y, z) , $(-x, 1/2+y, -z)$, $(1+x, y, z)$ et $(-x, 1/2+y, -z)$ sont représentés dans la maille monoclinique. L'axe hélicoïdal // à b est vertical.

Figure C1 : Comparaison des empilements cristallins du quintuple mutant B-S complexé avec NADP^+ et de l'enzyme de type sauvage complexé avec NAD^+ .

Tableau C7 : Contacts intermoléculaires dans le cristal du mutant B-Sp.

tétramère déduit de :	distance (Å)
a) axe hélicoïdal // à a	
OE2-Glu56(P) ... NZ-Lys114(R)	2.9
ND2-Asn67(P) ... O-Glu86(R)	3.0
b) axe binaire	
NZ-Lys114(O) ... OD2-Asp138(P)	2.9
O-Ala140(O) ... NH2-Arg102(P)	3.0
NE2-His141(O) ... OE1-Glu110(P)	3.2
NH2-Arg102(Q) ... O-Ala140(R)	3.1
OE1-Glu110(Q) ... NE2-His14	2.9

Leslie & Wonacott (1983) ont montré qu'il existe une relation très étroite entre les empilements cristallins du tétramère de la GAPDH de *B. stearothermophilus* dans la maille monoclinique et dans la maille orthorhombique (figure C2). Le paramètre **b** a la même dimension dans les deux mailles (tableau C1). Le passage du groupe d'espace $P2_1$ au groupe $P2_12_12$ modifie les opérateurs de symétrie reliant les tétramères voisins dans le cristal.

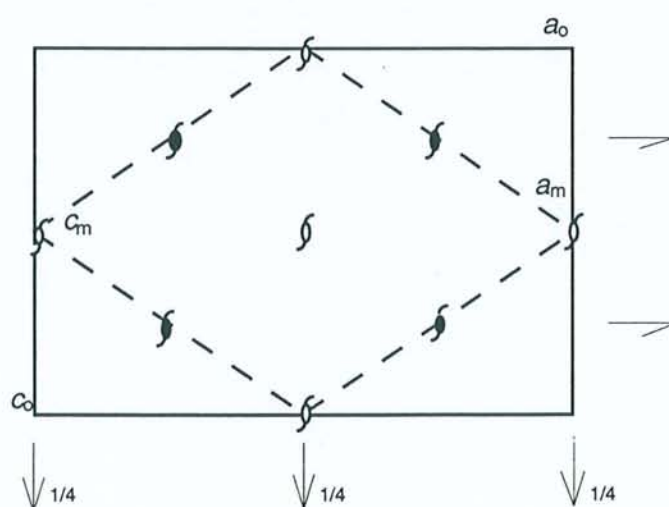


Figure C2 : Correspondances entre la maille monoclinique (traits pointillés) de l'enzyme de type sauvage et la maille orthorhombique (traits continus) du mutant B-Sp

Dans la forme monoclinique, la plus grande interface intermoléculaire dans l'empilement des tétramères est générée par l'axe hélicoïdal : le tétramère localisé en x, y, z implique dans le contact le dimère O-Q et le tétramère déduit de l'axe hélicoïdal le dimère P-R (figure C1b). L'interface générée par la translation suivant **a** implique le monomère P de l'un des tétramères et le monomère R de l'autre. De manière analogue, l'interface générée par la translation suivant **c** implique les monomères O et Q.

Dans la forme orthorhombique, la plus grande interface est générée par l'axe binaire parallèle à **c** (perpendiculaire au plan de la figure C1a). Dans ce cas, l'interface est formée par les dimères O-Q de chacun des tétramères symétriques par l'axe binaire. L'axe hélicoïdal parallèle à **a** joue le rôle des translations de la forme monoclinique en générant deux interfaces intermoléculaires entre les tétramères : une entre les monomères O et P et l'autre entre les monomères R et Q. L'axe hélicoïdal parallèle à **b** ne génère aucune interface dans la forme orthorhombique.

CIV AFFINEMENT DES MODELES

Cette étape indispensable permet de corriger les imperfections des modèles calculés à partir des phases initiales. Les quatre structures de mutants étudiées ont été affinées suivant le même protocole :

Le modèle initial dans l'affinement est celui de l'holoenzyme de type sauvage où les résidus mutés ont été systématiquement remplacés par des glycines. De plus, les molécules de NAD^+ et de solvant et les ions sulfate ne sont pas pris en compte.

Les affinements dans l'espace réciproque ont été réalisés avec le logiciel XPLOR version 3.1 (Brünger *et al.*, 1987 ; Brünger, 1992a,b). Dans la minimisation de la pseudo énergie potentielle 'Etot' (cf. § BV2), le terme dû à la mécanique moléculaire est calculé à l'aide du fichier de paramètres physico-chimiques de Engh & Huber (1991). Les modèles ont été examinés et reconstruits manuellement avec le logiciel TURBO-FRODO (Roussel & Cambillau, 1989).

Les intervalles de résolution utilisés dans les affinements sont 8-2.45 Å, 8-2.2 Å, 8-2.5 Å et 8-2.5 Å pour les structures D32G-Sn, D32G-Sp, B-Sn et B-Sp respectivement. Seules les réflexions pour lesquelles la norme du facteur de structure observé est supérieure à deux fois l'écart type estimé ($F_o > 2\sigma(F_o)$) sont utilisées dans les calculs (tableau C8).

En premier lieu, les structures D32G-Sn, D32G-Sp et B-Sn sont affinées en blocs rigides, ce qui permet de positionner et d'orienter les tétramères dans les unités asymétriques. Les facteurs d'accord R_f obtenus sont respectivement 34.8 %, 34 % et 31.5 %. Ce protocole inclus dans le programme de remplacement moléculaire AMoRe (Navaza, 1994) a déjà été appliqué sur la structure B-Sp, conduisant à un facteur R_f de 28.4 % pour les réflexions comprises entre 15 et 4 Å de résolution.

A ce stade, la procédure du recuit simulé (Brünger *et al.*, 1990) a été utilisée pour essayer de perturber la structure de l'holoenzyme de type sauvage, afin d'observer les changements structuraux éventuels induits par les différentes mutations. Cette procédure d'affinement se décompose en quatre phases : La première étape consiste à évaluer le poids W_a permettant d'équilibrer le gradient de l'énergie cristallographique 'Ecris' avec celui de l'énergie moléculaire 'Emol' (cf. § BV2). Durant la deuxième phase, un affinement par gradient conjugué est réalisé dans le but d'éliminer les mauvais contacts non liés qui pourraient poser problème au départ de la dynamique moléculaire. Le modèle ainsi 'préparé' subit alors un recuit simulé avec pour températures initiale 2000 K et finale 300 K. Le refroidissement est réalisé par sauts de 50 K toutes les 25 fs avec une incrémentation de temps de 0.5 fs. Cette procédure se termine par un affinement (120 cycles) des positions atomiques, suivi du calcul d'une carte de densité différence ($F_o - F_c$).

Le premier examen graphique permet de vérifier l'occupation des quatre sites du cofacteur de l'enzyme et le bon remplacement des molécules de NAD par des molécules de NADP dans les modèles des structures de D32G-Sp et de B-Sp (figure C3). Dans chacune des structures résolues, il a été possible, sans ambiguïté, d'introduire les coenzymes dans les modèles. La conformation initiale du cofacteur introduit est celle observée dans l'enzyme de type sauvage.

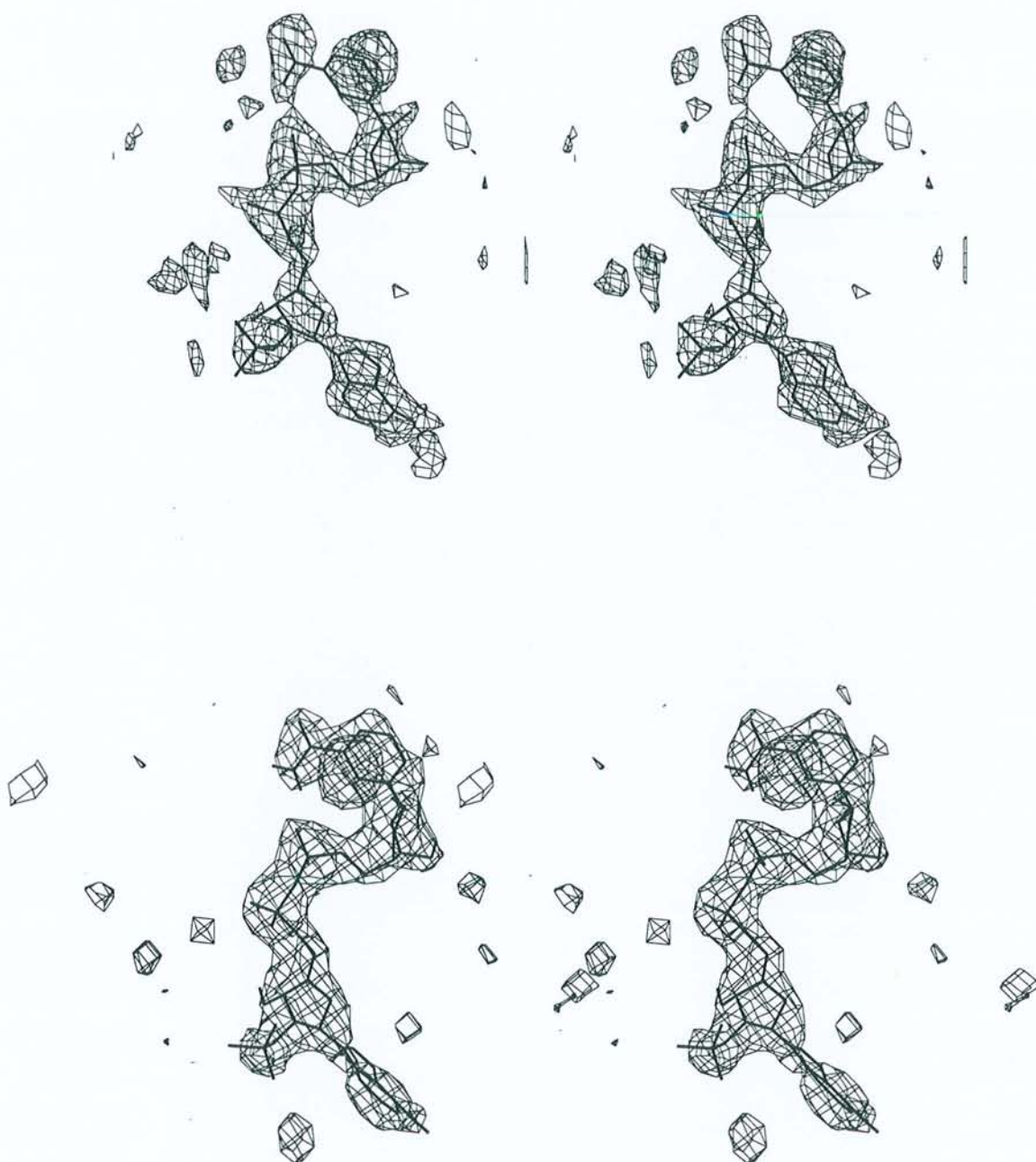


Figure C3 : Densité électronique du NADP du monomère O du mutant D32G-Sp (figure du haut) et du mutant B-Sp (figure du bas). Carte de densité électronique de type (Fo-Fc), niveau de contour 2.6 σ .

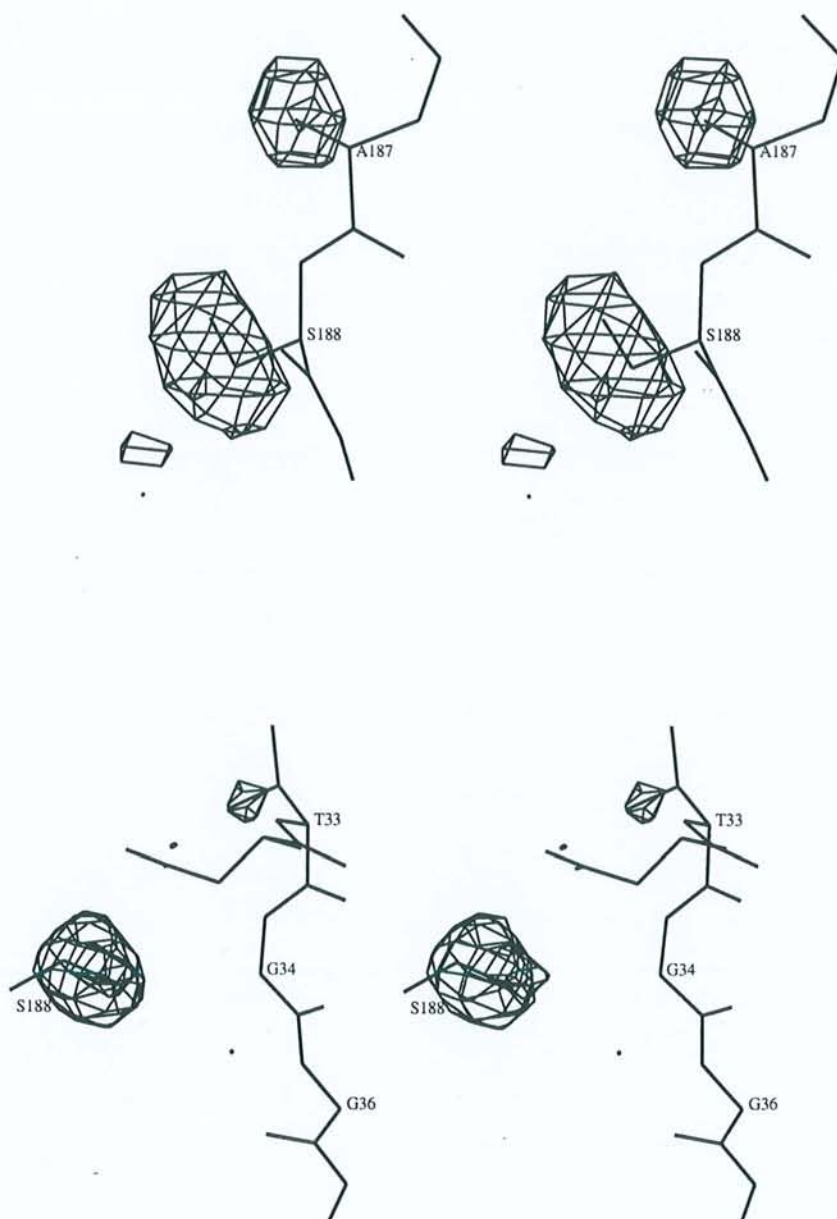


Figure C4 : Densité électronique autour des chaînes latérales des résidus S188 et Ala187 dans la structure du mutant D32G-Sn (figure du haut) et autour des chaînes latérales des résidus T33 et S188 du NADP dans la structure du mutant B-Sn (figure du bas).

Les cartes de densité électronique sont calculées en utilisant les coefficients $(|F_o| - |F_c|) \exp(i\alpha_{\text{calc}})$ où $|F_o|$ correspond à l'amplitude du facteur de structure observé et $|F_c|$ et α_{calc} sont respectivement l'amplitude et la phase du facteur de structure calculé à partir des modèles des mutants D32G-Sn et B-Sn dans lesquels les chaînes latérales des résidus mutés ne sont pas prises en compte. Le niveau de contour dans les deux cartes est de 2.8σ .

La procédure d'affinement en recuit simulé est réitérée, puis suivie d'un affinement des facteurs d'agitation thermique. Les facteurs R_f résultant de cet affinement sont inférieurs à 25% dans toutes les structures, ce qui nous autorise à calculer des cartes $(2F_o-F_c)$ (McRee, 1993, p126-127). A partir de la carte résiduelle, il est possible de construire dans chaque monomère un ion sulfate dans les sites Pi et Ps (Biesecker *et al.*, 1977). Les pics de densité supérieurs à 3.5 fois la déviation standard ($(F_o-F_c) > 3.5 \sigma$) sont attribués à des molécules d'eau lorsqu'ils sont à une distance de l'ordre de 3 Å d'atomes susceptibles de former des liaisons hydrogène. Finalement, les molécules d'eau dont le contour est supérieur à 1 σ dans la carte $(2F_o-F_c)$ de l'affinement suivant sont conservées.

Cette procédure conduit à une construction progressive des molécules d'eau à partir de l'examen simultané des cartes (F_o-F_c) et $(2F_o-F_c)$ calculées à la suite de l'affinement des positions atomiques et des facteurs d'agitation thermique du modèle. Il est intéressant de noter que la plupart des molécules enfouies dans l'enzyme de type sauvage sous forme holo et sous forme apo (Skarzynski *et al.*, 1987; Skarzynski & Wonacott, 1988) sont conservées dans les structures des quatre mutants.

L'examen des cartes (F_o-F_c) permet de confirmer en fin d'affinement les mutations des résidus Leu187, Pro188 et Asp32 respectivement en alanine, sérine et glycine dans le mutant D32G-S (figure C4). Dans le mutant B-S, les mutations Leu187Ala, Pro188Ser, Thr34Gly et Asp35Gly ont été également confirmées dans les cartes résiduelles, alors que la mutation Leu33Thr n'a pu être clairement mise en évidence, en raison du manque de densité électronique autour de la chaîne latérale de ce résidu. L'ensemble des caractéristiques des modèles en fin d'affinement est donné dans le tableau C8.

Tableau C8 : Statistiques en fin d'affinement

	D32G-Sn	D32G-Sp	B-Sn	B-Sp
statistiques sur l'ensemble des données				
intervalle de résolution (Å)	8.-2.5	8.-2.2	8.-2.45	8.-2.5
réflexions ($F_o > 2\sigma(F_o)$)	43095	65530	53718	48832
complétude ($F_o > 2\sigma(F_o)$)	80.6	78.8	90.	89.4
facteur Rf (%)	13.9	19.2	19.3	14.4
facteur Rfree (%)	21.0	25.9	26.6	21.0
réflexions utilisées dans le calcul du Rfree	4344	4197	5566	4970
statistiques sur les données à haute résolution				
intervalle de résolution (Å)	2.5-2.61	2.2-2.30	2.45-2.55	2.61-2.5
réflexions dans cet intervalle ($F_o > 2\sigma(F_o)$)	4816	6374	5518	5506
complétude des données ($F_o > 2\sigma(F_o)$)	72.9	65.0	77.	81.
facteur R	20.1	25.6	25.7	20.0
facteur Rfree	26.6	31.0	32.9	27.7
réflexions utilisées dans le calcul du Rfree	489	617	564	545

Remarque : dans l'étude des mutants de la GAPDH, le Rfree n'a été calculé qu'en fin d'affinement. 10 % des réflexions ont été prélevées de manière aléatoire du jeu de données. Un affinement des positions atomiques suivi d'un affinement des facteurs d'agitation thermique a alors été réalisé en ne prenant en compte que les 90 % restant. Dans chaque cas le Rfree se stabilise assez vite, ce qui laisse supposer que l'ensemble des réflexions utilisé pour calculer ce facteur est décorrélié du modèle. Les carbones α des structures issues de ce 'réaffinement' ont été superposés avec ceux des structures le précédant pour vérifier la stabilité des modèles. Les écarts quadratiques moyens résultants sont tous inférieurs à 0.06 Å.

CV QUALITE DES MODELES

CVa Densité électronique

Les pics de densité électronique les plus intenses dans les cartes de densité électronique (Fo-Fc) finales de tous les mutants sont localisés sur les atomes de soufre des quatre Cys149 du tétramère. Ceci montre que l'ajout de 1mM de DTT dans les solutions de cristallisation n'a pas été suffisant pour empêcher l'oxydation partielle de la chaîne latérale de la cystéine catalytique de la GAPDH.

Dans chacun des mutants, l'inspection de la carte de densité électronique (2Fo-Fc) finale montre que la densité électronique est plus diffuse dans le domaine de fixation du cofacteur que dans le domaine catalytique. Cependant, la densité électronique est continue à un niveau de contour de 1σ le long de la chaîne principale de tous les modèles. Comme dans l'holoenzyme de type sauvage (Skarzynski *et al.*, 1987), plusieurs chaînes latérales chargées et situées à la surface de la protéine sont mal définies dans les cartes de densité électronique (2Fo-Fc) finales (en particulier les résidus N61, N62, R77, E80, R98, E103 et K191). Les conformations appliquées à ces résidus dans les modèles finaux sont celles de l'holoenzyme de type sauvage.

La bonne qualité de la densité électronique autour du 2'-phosphate du NADP dans la structure du mutant D32G-Sp nous a permis de modéliser des molécules d'eau en interaction avec les oxygènes du phosphate. Il est intéressant de noter que les molécules d'eau Wat366 et Wat580 sont observées dans les quatre monomères et servent de ponts entre la protéine et le groupement 2'-phosphate (Figure C5). Par contre, peu de molécules d'eau ont été modélisées autour du 2'phosphate dans la structure du mutant B-Sp. De plus, ces molécules d'eau ne sont pas conservées dans les quatre monomères.

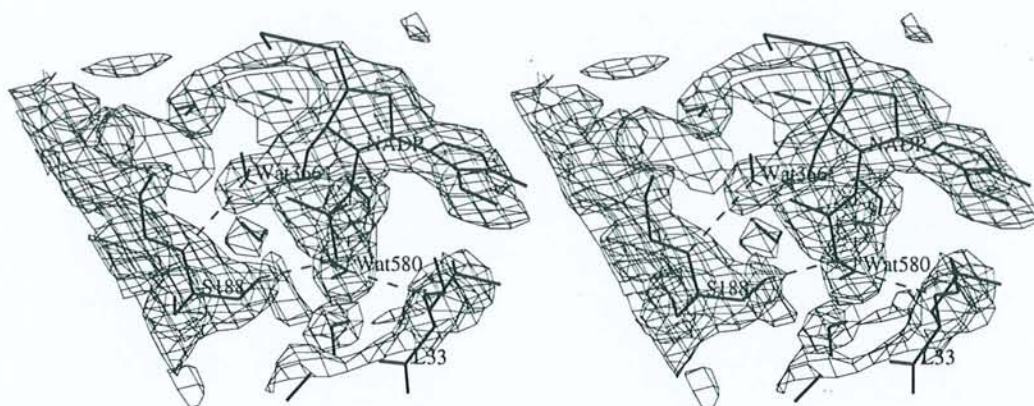


Figure C5 : Illustration de la densité électronique autour du 2'-phosphate du NADP du mutant D32G-Sp.

Carte de densité de type (2Fo-Fc), niveau de contour 1σ .

CVb Estimation de l'erreur sur les positions atomiques

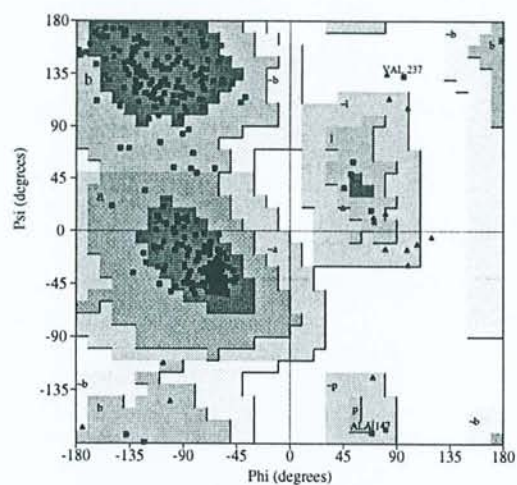
Dans un premier temps, nous avons utilisé la méthode de Luzzati (1952) pour obtenir une estimation de la limite supérieure de l'écart type des coordonnées atomiques. Cette procédure suggère un écart type de 0.22 Å dans les structures des mutants D32G-Sn et B-Sp, de 0.25 Å dans la structure du mutant D32G-Sp et de 0.30 Å dans la structure du mutant B-Sn. Comme l'ont souligné de nombreux cristallographes (Skarzynski *et al.*, 1987 ; Mittl *et al.*, 1994) la méthode de Luzzati (1952) ne donne qu'une approximation très grossière de l'erreur sur les positions atomiques. L'intervention de Cruickshank au congrès international de la cristallographie à Seattle (1996), intitulé 'Protein precision re-examined : Luzzati plots do not estimate final errors', va dans le même sens. Cruickshank a montré que les courbes théoriques permettant d'estimer la limite supérieure de l'écart type des coordonnées sont calculées à partir d'hypothèses non applicables dans la cristallographie des protéines. Alors qu'il est établi que l'erreur sur la position des atomes n'est pas uniforme dans les structures car elle dépend fortement de la valeur du facteur d'agitation thermique, l'une des hypothèses suppose que la distribution de l'erreur est identique pour chaque atome. L'écart type résultant de la

superposition des carbones α de deux monomères d'un même tétramère a alors été utilisé pour estimer l'erreur sur les positions atomiques (Fermi, 1975 ; Mittl *et al.*, 1994). Comme dans l'holoenzyme de type sauvage (Skarzynski *et al.*, 1987), l'hypothèse selon laquelle les tétramères des mutants de la GAPDH de *B. stearrowthermophilus* présentent une symétrie 222 quasi exacte a été faite. Les écarts types obtenus pour les structures des mutants D32G-Sn, D32G-Sp, B-Sn et B-Sp sont respectivement .26 Å, .29 Å, .34 Å et .26 Å.

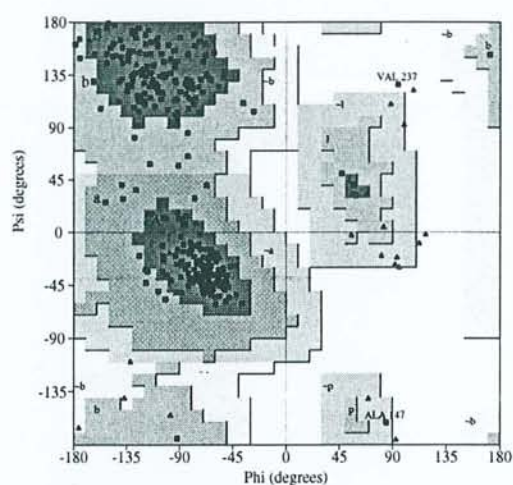
CVc Géométrie

Les diagrammes de Ramachandran (Ramakrishnan *et al.*, 1965) des quatre structures de mutants sont semblables à celui observé dans l'holoenzyme de type sauvage. Les résidus Ala147 et Val237 sont localisés dans des régions de conformation non favorable dans les quatre structures étudiées (figure C6). Ces deux acides aminés sont néanmoins bien définis dans les cartes finales de densité électronique (2Fo-Fc). Le logiciel PROCHECK de Laskowski *et al.* (1993) a été utilisé pour déterminer le pourcentage de résidus situés dans une région de conformation très favorable. Il est de 88.8 %, 88.3 %, 88. % et 89.9 % dans les structures des mutants D32G-Sn, D32G-Sp, B-Sn et B-Sp respectivement.

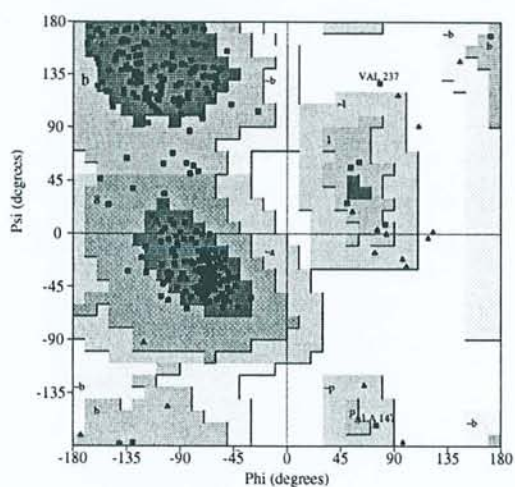
Le tableau C9 résume l'ensemble des caractéristiques des fichiers de coordonnées des structures des mutants D32G-Sn, D32G-Sp, B-Sn et B-Sp déposés dans la banque de données des structures de protéines PDB (Berstein *et al.*, 1977) respectivement avec les codes 1dbv, 2dbv, 3dbv et 4dbv.



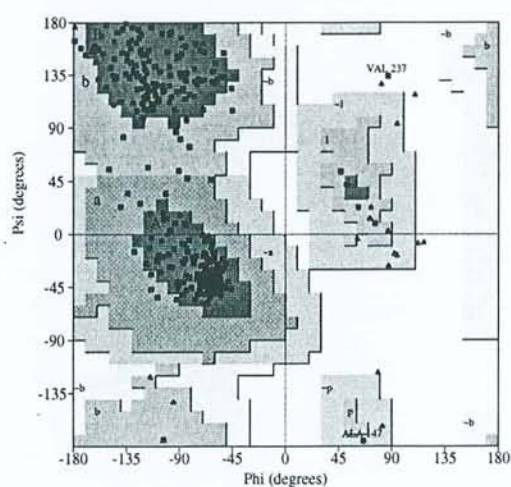
D32G-Sn



D32G-Sp



B-Sn



B-Sp

Figure C6 : Diagramme des angles dièdres de la chaîne principale d'une sous-unité des mutants

D32G-Sn (en haut à gauche), D32G-Sp (en haut à droite), B-Sn (en bas à gauche) et B-Sp (en bas à droite).

Ces diagrammes ont été tracés à l'aide du logiciel PROCHECK (Laskowski *et al.*, 1993). Les triangles correspondent aux glycines et les carrés aux autres résidus. Les résidus Ala147 et Val237 sont localisés dans des zones de conformation défavorable.

Tableau C9 : Caractéristiques des fichiers de coordonnées

Mutant	D32G-Sn	D32G-Sp	B-Sn	B-Sp
Code PDB	1dbv	2dbv	3dbv	4dbv
Nombre d'atomes introduits dans les modèles ^a				
atomes protéiques	10068	10068	10052	10052
atomes des cofacteurs et des anions sulfate	216	232	216	232
molécules d'eau	384	393	291	341
facteur d'agitation thermique moyen ($\text{\AA}^2$):				
global ^a	27.	27.	28.	30.
pour la chaîne principale ^a	23.	24.	24.	24.
pour les cofacteurs ^a	24.	22.	20.	26.
pour les molécules d'eau ^a	35.	34.	31.	37.
Ecart type par rapport aux valeurs idéales :				
Longueurs de liaisons ($\text{\AA}$)	.009	.012	.012	.009
Angles ($^\circ$)	1.64	1.83	1.79	1.69
Angles dièdres ($^\circ$)	25.8	26.1	26.1	24.9
Angles dièdres impropres ($^\circ$)	1.49	1.62	1.67	1.45

^a Atomes d'hydrogène exclus.

PARTIE D

RESULTATS ET DISCUSSION

DI RESUME DE LA PUBLICATION

DIa Partie résultats structuraux

DIa1 Etat général des mutants.

Les structures des quatre complexes D32G-Sn, D32G-Sp, B-Sn et B-Sp révèlent une conformation plutôt 'holo' que 'apo' par comparaison avec les deux structures de l'enzyme de type sauvage avec et sans NAD⁺. Une analyse plus fine permet de montrer que les feuillets $\beta 1$, $\beta 3$, $\beta 4$ et $\beta 7$ des quatre complexes forment un corps structural invariant par rapport à l'holoenzyme de type sauvage. De plus, dans les mutants D32G-Sp et B-Sn ce corps invariant comprend également l'hélice $\alpha 1$ qui contient le résidu catalytique Cys149. Cette caractéristique semble en accord avec les bandes de Racker mesurées dans ces deux complexes. En effet, les bandes de Racker sont comparables à celle mesurée dans l'holoenzyme de type sauvage alors qu'elles sont diminuées de moitié dans les complexes D32G-Sn, B-Sp. Ceci permet de distinguer les complexes D32G-Sp, B-Sn des complexes D32G-Sn, B-Sp en fonction de leurs caractère plus ou moins 'holo'.

DIa2 Les mutations D32G, B et S

- La mutation D32G n'induit pas de changement conformationnel au niveau de la chaîne principale de la boucle $\beta B-\alpha C$. L'espace libéré par la chaîne latérale de l'acide aspartique est occupé dans les deux complexes D32G-Sn et G32G-Sp par une molécule d'eau (W581) qui sert de pont entre l'atome AO2' du NAD(P)⁺ et le groupe NH de la leucine 33.

- Le remplacement des résidus Leu33, Thr34 et Asp35, respectivement par une thréonine, une glycine et une glycine donne une plus grande liberté conformationnelle à la boucle $\beta B-\alpha C$ qui se traduit principalement par une rotation de 55° de la liaison peptique 33-34. La chaîne latérale de l'acide aspartique 32 se trouve alors impliquée dans un liaison hydrogène supplémentaire avec le groupe NH du résidu Gly34.

- La double mutation S (L187A, P188S) ne perturbe pas la disposition de la S-loop. La sérine 188 nouvellement introduite semble jouer un rôle dans la stabilisation du 2'-phosphate du NADP⁺ dans le mutant D32G-Sp. Par contre, aucune interaction n'est observée entre la chaîne latérale de cette sérine et les oxygènes du 2'-phosphate dans le complexe B-Sp.

DIa3 Le site du cofacteur

Les mutations réalisées dans le site de fixation de la partie AMP du cofacteur, n'induisent pas de changement structural notable dans le site de fixation de la partie NMN. Dans les quatre structures, les nicotinamides et leurs riboses adjacents sont superposables aux mêmes fragments observés dans la structure de l'holoenzyme de type sauvage.

Les écarts observés dans la superposition des parties AMP du cofacteur semblent être liés au caractère plus ou moins holo de chaque mutant. Les riboses des adénosines des cofacteurs dans les complexes D32G-Sn, B-Sn et B-Sp adoptent le même 'puckering' que celui observé dans l'holoenzyme de type sauvage (C2'-endo). Par contre le ribose de l'adénosine du NADP⁺ dans la structure D32G-Sp est de type C1'-exo.

DIb Discussion

DIb1 Recherche des facteurs moléculaires jouant un rôle dans la sélectivité NAD/NADP

Le mutant D32G-S est caractérisé par une inversion de sélectivité de NAD⁺ vers NADP⁺ par rapport à l'holoenzyme de type sauvage. L'observation de la structure de D32G-Sn suggère que la baisse d'affinité de ce mutant pour NAD⁺ est principalement liée à la suppression de la chaîne latérale du résidu Asp32. Dans le complexe D32G-Sp, cette mutation D32G est déterminante pour éviter une répulsion électrostatique entre le carboxylate de ce résidu acide et le 2'-phosphate du NADP⁺. Cette mutation permet d'obtenir une bonne affinité du triple mutant pour NADP⁺, qui découle également de la stabilisation du 2'phosphate par le groupe hydroxyle de la sérine 188.

Le mutant B-S construit à partir de la séquence de la GAPDH chloroplastique ne présente pas d'inversion de sélectivité de NAD^+ vers NADP^+ . L'augmentation du K_D du NAD^+ dans ce mutant par rapport à celui observé dans l'holoenzyme de type sauvage semble lié à deux effets : le contact de van der Waals adénine/résidu 33 apparaît affaibli dans le mutant B-S par rapport à celui observé dans l'holoenzyme de type sauvage ; l'interaction supplémentaire observée dans la structure B-Sn entre l'atome OD2 de l'acide aspartique et le groupe NH de la Gly34 diminue probablement l'intensité de l'interaction entre l'atome de OD2 et l'hydroxyle AO2' du NAD^+ . La structure B-Sp montre que la faible affinité de ce quintuple mutant pour NADP^+ est essentiellement liée au résidu Asp32. Le 2'-phosphate déstabilisé par ce résidu acide ne forme aucune interaction directe avec la protéine.

D1b2 Comparaison avec les sites NADP naturels

La comparaison de sept structures d'enzymes à 'Rossmann fold' cocrystallisées avec NADP a permis de montrer que dans tous les cas le 2'-phosphate est stabilisé par au moins une arginine qui de surcroît interagit favorablement avec une face de l'adénine par contact $\pi-\pi$. Dans la majorité des cas cette arginine est localisée dans la boucle $\beta\text{B}-\alpha\text{c}$. Il est intéressant de noter que dans trois cas sur sept le site de fixation du cofacteur implique au moins un résidu à chaîne hydroxylée dans la stabilisation du groupement 2'-phosphate. Cependant dans tous les cas, le résidu qui apparaît le plus important dans la reconnaissance du NADP est l'arginine. Dans le complexe D32G-Sp, c'est la sérine 188 qui semble jouer ce rôle car c'est le seul acide aminé en interaction avec le 2'-phosphate.

L'observation des sites de fixation du cofacteur NADP dans des enzymes ne présentant pas le repliement de 'Rossmann' permet d'aboutir aux mêmes conclusions : dans tous les cas une arginine est impliquée dans la stabilisation du 2'-phosphate et du cycle adénine. Cependant, des études de mutagenèse dirigée réalisées sur la dihydrofolate réductase et sur l'isocitrate déshydrogénase ont montré que des résidus à chaînes hydroxylées jouaient un rôle important dans la stabilisation du 2'-phosphate. Ceci renforce l'idée selon laquelle la sérine 188 joue

probablement un rôle dans la fixation du groupement phosphate dans les GAPDH chloroplastiques.

Dib3 Remarques sur les GAPDH chloroplastiques

Les GAPDH chloroplastiques présentent une dualité de cofacteur NAD/NADP avec une préférence pour NADP. Elles possèdent au niveau C-terminal du feuillet β B le résidu acide aspartique en position 32. Les études structurales des complexes Sp et B-Sp ont montré que ce résidu acide joue un rôle antidéterminant pour la fixation du NADP⁺ en inhibant la stabilisation du phosphate. Les cinq mutations L33T, T34G, D35G, L187A et P188S dans le mutant B-S sont donc insuffisantes pour mimer la fixation du NADP⁺ dans les GAPDH chloroplastiques. Les effets de deuxième couche ou à longue distance, difficiles à mettre en évidence, jouent probablement un rôle dans la reconnaissance du cofacteur dans les GAPDH chloroplastiques.

**A Crystallographic Comparison Between Mutated
Glyceraldehyde-3-phosphate Dehydrogenases from
Bacillus stearothermophilus Complexed with Either
NAD⁺ or NADP⁺**

Claude Didierjean, Sophie Rahuel-Clermont, Bernard Vitoux
Otto Dideberg, Guy Branlant and André Aubry

A Crystallographic Comparison Between Mutated Glyceraldehyde-3-phosphate Dehydrogenases from *Bacillus stearothermophilus* Complexed with Either NAD⁺ or NADP⁺

Claude Didierjean¹, Sophie Rahuel-Clermont², Bernard Vitoux¹
Otto Dideberg³, Guy Branlant² and André Aubry^{1*}

¹Laboratoire de Cristallographie et Modélisation des Matériaux Minéraux et Biologiques, IFR Protéines, Université Henri Poincaré Nancy I (URA CNRS 809), BP 239 54506 Vandoeuvre-lès-Nancy Cedex, France

²Laboratoire d'Enzymologie et de Génie Génétique, IFR Protéines, Université Henri Poincaré Nancy I (URA CNRS 457), BP 239 54506 Vandoeuvre-lès-Nancy Cedex, France

³Laboratoire de Cristallographie Macromoléculaire, Institut de Biologie Structurale Jean Pierre EBEL, (CNRS-CEA), 41 Avenue des Martyrs 38027 Grenoble Cedex 1 France

Mutations have been introduced in the cytosolic glyceraldehyde-3-phosphate dehydrogenase (GAPDH) from *Bacillus stearothermophilus* in order to convert its cofactor selectivity from a specificity towards NAD into a preference for NADP. In the B-S mutant, five mutations (L33T, T34G, D35G, L187A, P188S) were selected on the basis of a sequence alignment with NADP-dependent chloroplastic GAPDHs. In the D32G-S mutant, two of the five mutations mentioned above (L187A, P188S) have been used in combination with another one designed from electrostatic considerations (D32G). Both mutants exhibit a dual-cofactor selectivity at the advantage of either NAD (B-S) or NADP (D32G-S). In order to analyse the cofactor-binding site plasticity at the molecular level, crystal structures of these mutants have been solved, when complexed with either NAD⁺ (D32G-Sn, resolution 2.5 Å, *R* = 13.9%; B-Sn, 2.45 Å, 19.3%) or NADP⁺ (D32G-Sp, 2.2 Å, 19.2%; B-Sp, 2.5 Å, 14.4%). The four refined models are very similar to that of the wild-type GAPDH and as expected resemble more closely the holo form than the apo form.

In the B-S mutant, the wild-type low affinity for NADP⁺ seems to be essentially retained because of repulsive electrostatic contacts between the extra 2'-phosphate and the unchanged carboxylate group of residue D32. Such an antideterminant effect is not well compensated by putative attractive interactions which had been expected to arise from the newly-introduced side-chains. In this mutant, recognition of NAD⁺ is slightly affected with respect to that known on the wild-type, because mutations only weakly destabilize hydrogen bonds and van der Waals contacts originally present in the natural enzyme. Thus, the B-S mutant does not mimic efficiently the chloroplastic GAPDHs, and long-range and/or second-layer effects, not easily predictable from visual inspection of three-dimensional structures, need to be taken into account for designing a true "chloroplastic-like" mutant of cytosolic GAPDH.

In the case of the D32G-S mutant, the dissociation constants for NAD⁺ and NADP⁺ are practically reversed with respect to those of the wild-type. The strong alteration of the affinity for NAD⁺ obviously proceeds from the suppression of the two wild-type hydrogen bonds between the adenosine 2'- and 3'-hydroxyl positions and the D32 carboxylate group. As expected, the efficient recognition of NADP⁺ is partly promoted by the removal of intra-subunit electrostatic repulsion (D32G) and inter-subunit steric hindrance (L187A, P188S). Another interesting feature of the reshaped NADP⁺-binding site is provided by the local stabilization of the extra 2'-phosphate which forms a hydrogen bond with the side-chain hydroxyl group of the newly-introduced S188. When compared to the presently known natural NADP-binding clefts, this result clearly demonstrates that an absolute need for a salt-bridge involving the 2'-phosphate is not required to switch the cofactor selectivity from NAD to

*Corresponding author

NADP. In fact, as it is the case in this mutant, only a moderately polar hydrogen bond can be sufficient to make the extra 2'-phosphate of NADP⁺ well recognized by a protein environment.

© 1997 Academic Press Limited

Keywords: GAPDH; NAD(P); X-ray structures; cofactor selectivity; Rossmann fold

Introduction

In living organisms, oxidoreductive processes are catalysed by enzymes which require various kinds of cofactor to achieve their functions. An overview of oxidoreductases known at the present time reveals that two of the most widely used coenzymes are nicotinamide adenine dinucleotide (NAD) and nicotinamide adenine dinucleotide phosphate (NADP) (Stryer, 1995). NADP differs from NAD by the presence of an extra phosphate group which esterifies the 2'-hydroxyl position of the ribose ring in the adenosine moiety. Because of this difference, NAD and NADP assume dissimilar intermolecular interactions with enzymes (Scrutton *et al.*, 1990): the free hydroxyl groups of the NAD adenosine part generally behave as proton donors in hydrogen bonds, whereas the 2'-phosphate function of NADP acts as a strong salt-bridge acceptor, in agreement with its predominant dianionic state at physiological pH (Stryer, 1995).

In relation with the distinct metabolic roles assigned to NAD (catabolic processes) and NADP (anabolic processes) in living organisms (Stryer, 1995), the vast majority of NAD(P)-dependent oxidoreductases exhibits a strong preference for either NAD or NADP. During the three last decades, crystallographic experiments have clearly revealed that a strong discrimination between NAD and NADP does not proceed from the presence of structurally distinct domains in enzymes: the most widely encountered dinucleotide-binding domain, traditionally known as the Rossmann fold (Rossmann *et al.*, 1974), has been demonstrated to be present in NAD as well as NADP-dependent dehydrogenases and reductases (Rossmann *et al.*, 1975; Pai, 1991; Mathews, 1991; Schulz, 1992; Lesk, 1995). The same assessment holds true for the more recently discovered (1) cofactor-binding motif shared by NADP-dependent isocitrate dehydrogenase (Hurley *et al.*, 1991; Stoddard *et al.*, 1993) and NAD-dependent 3-isopropylmalate dehydrogenase (Hurley & Dean, 1994) and (2) (β/α)₈ barrel present in NADP-dependent aldose reductases (Rondeau *et al.*, 1992; Wilson *et al.*, 1992), NADP-dependent aldehyde reductases (El-Kabbani *et al.*, 1994, 1995) and dual-cofactor selective 3 α -hydroxy-steroid/dihydrodiol dehydrogenase (Hoog *et al.*, 1994).

Thus, molecular determinants responsible for the selective recognition of either NAD or NADP have been suggested to lie in the close vicinity of the adenosine ribose, where they would be able to

locally modulate electrostatic and steric properties of the inner surface of the cofactor-binding cleft (Wierenga *et al.*, 1986; Hanukoglu & Gutfinger, 1989; Scrutton *et al.*, 1990; Bork & Grunwald, 1990; Baker *et al.*, 1992). Since the beginning of the nineties, much attention has been paid to identify these molecular fragments with the aid of site-directed mutagenesis. To this purpose, many experiments have attempted to switch the coenzyme selectivity of a wild-type enzyme by the introduction of rationally designed single or multiple-point mutations. On the basis of the sequence alignments of pairs of structurally related enzymes, but with opposite cofactor preferences, convenient mutations have been achieved and pertinent results have been obtained in the case of (1) NADP-dependent glutathione reductase *versus* NAD-dependent dihydrolipoamide dehydrogenase (Scrutton *et al.*, 1990) and *vice versa* (Bocanegra *et al.*, 1993), (2) NAD-dependent xylitol dehydrogenase *versus* NADP-dependent alcohol dehydrogenase (Metzger & Hollenberg, 1995), (3) NAD-dependent malate dehydrogenase *versus* its NADP-dependent counterpart (Nishiyama *et al.*, 1993), (4) NAD-dependent 3-isopropylmalate dehydrogenase *versus* NADP-dependent isocitrate dehydrogenase (Miyazaki & Oshima, 1994) and *vice versa* (Chen *et al.*, 1995) and (5) NAD-dependent cytosolic glyceraldehyde-3-phosphate dehydrogenase (GAPDH) *versus* NADP-dependent chloroplastic GAPDH (Corbier *et al.*, 1990a; Clermont *et al.*, 1993; Rahuel-Clermont, 1994; Eyschen *et al.*, 1996).

As expected, these studies have reliably confirmed that residues closely surrounding the adenosine 2'-position have striking effects on cofactor-selective recognition. On the one hand, preference towards NAD seems mainly related to the presence of an anionic side-chain able to simultaneously accept the 2' and 3'-hydroxyl groups through hydrogen bonding (Fan *et al.*, 1991; Chen *et al.*, 1991; Grimshaw *et al.*, 1992; Bernard *et al.*, 1995). On the other hand, acceptance of the NADP 2'-phosphate is often achieved by attractive electrostatic interactions with the positively-charged residues lysine or arginine (Huang *et al.*, 1990; Haefner-Gormley *et al.*, 1992; Sem & Kasper, 1993a,b; Chen *et al.*, 1994; Yaoi *et al.*, 1994; Levy *et al.*, 1996). In addition to these first-order phenomena, two recently-resolved crystal structures of mutants complexed with the "unnatural" coenzyme have emphasized the occurrence of more subtle effects on steric packing interactions, which are dictated by side-chains located in the

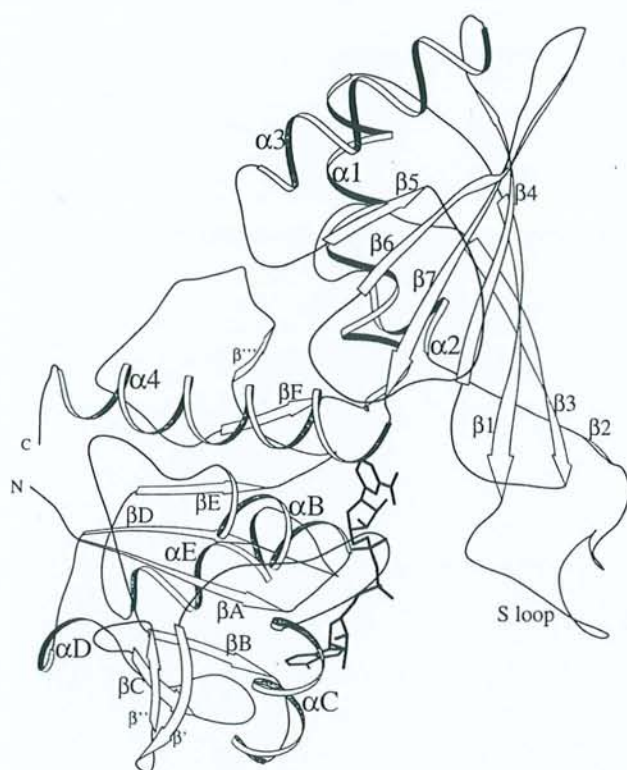


Figure 1. Localization of secondary structures in the schematic representation of the GAPDH subunit from *B. stearothermophilus*.

second layer of the binding site (Mittl *et al.*, 1994; Hurley *et al.*, 1996).

The contribution reported here concerns the ubiquitous phosphorylating glyceraldehyde-3-phosphate dehydrogenase (GAPDH) which catalyses the reversible oxidative phosphorylation of glyceraldehyde-3-phosphate in the presence of NAD(P)⁺ and inorganic phosphate (Harris & Waters, 1976). The cofactor selectivity of GAPDH has been shown to depend on the enzyme's subcellular localization: the glycolytic enzyme (EC 1.2.1.12) present in the cytoplasm of eubacteria and eukaryotes is strictly NAD-specific, whereas both NAD and NADP are recognized by chloroplastic GAPDHs (EC 1.2.1.13), with a preference for NADP (Cerff, 1978; Ferri *et al.*, 1978). Cytoplasmic and chloroplastic GAPDHs function as tetramers containing four identical subunits (glycolytic enzymes) (Harris & Waters, 1976) or two closely related isoforms A and B with A₄ or A₂B₂ stoichiometries (chloroplastic enzymes) (Ferri *et al.*, 1978, 1990; Cerff, 1979).

Crystal structures have been solved for many glycolytic GAPDHs isolated from mesophilic organisms such as man (Watson *et al.*, 1972; Mercer *et al.*, 1976), american and chinese lobsters (Buehner *et al.*, 1974; Moras *et al.*, 1975; Lin *et al.*, 1993), protozoan parasites (Vellieux *et al.*, 1995; Kim *et al.*, 1995) and *Escherichia coli* (Duée *et al.*, 1996), and from thermophilic eubacteria such as *Bacillus stearothermophilus* (Biesecker *et al.*, 1977;

Table 1. Mutant notations

Mutant	Mutations	NAD ⁺ /mutant complex	NADP ⁺ /mutant complex
D32G-S	D32G L187A P188S	D32G-Sn	D32G-Sp
B-S	L33T T34G D35G L187A P188S	B-Sn	B-Sp

Skarzynski *et al.*, 1987), *Thermotoga maritima* (Korn-dörfer *et al.*, 1995) and *Thermus aquaticus* (Tanner *et al.*, 1996). All these three-dimensional structures appear to be closely related to each other and contain a similar native fold enabling the direct superimposition of secondary structures, structural domains and whole subunits for comparative purposes. In particular, all these tetrameric GAPDHs possess a quasi 222 symmetry characterized by three orthogonal axes denoted *P* (between O and P subunits), *Q* (between O and Q subunits) and *R* (between O and R subunits), respectively (Rossmann *et al.*, 1973). Furthermore, the domains responsible for cofactor binding typically are of the Rossmann-fold type, with a six-stranded parallel β -sheet flanked on each side by two helices, thus corresponding to a $\beta\alpha\beta\alpha\beta$ topology repeated twice along the primary structure (Figure 1). From the comparison of the relevant primary structures, which reveals an identity score of about 50%, it is highly probable that the NADP-dependent chloroplastic enzymes adopt the same fold as the one observed for glycolytic GAPDHs (Martin & Cerff, 1986; Shih *et al.*, 1986). On this basis, we have considered the well-resolved crystal structure of *B. stearothermophilus* glycolytic GAPDH as an appropriate starting model to test the effects of amino acid substitutions on cofactor selectivity.

Among the series of GAPDH mutants which have been designed, purified and characterized in previous papers (Corbier *et al.*, 1990a; Clermont *et al.*, 1993; Eyschen *et al.*, 1996), two of them, namely D32G-S and B-S (Table 1), appear to be the most attractive ones because of their biochemical and physicochemical properties. They share the same mutations at positions 187 and 188, where leucine and proline are respectively replaced by alanine and serine, two invariant residues in practically all known chloroplastic GAPDH sequences. This specific chloroplastic feature is presumed to prevent the occurrence of steric hindrance between the NADP 2'-phosphate and the so-called S-loop motif which is situated in the close vicinity of bound NAD(P) and participates in the *R*-related quaternary contacts in the tetramer. In addition, the D32G-S mutant contains a single point mutation not related to chloroplastic sequence: glycine replaces aspartic acid 32 and thus prevents

Table 2. Dissociation constants, kinetic parameters and molar extinction coefficients of NAD(P) for wild-type and mutant GAPDHs

	NAD ⁺				NADP ⁺			
	K_D^a (μ M)	k_{cat}^b (s ⁻¹)	k_{cat}/K_D (μ M ⁻¹ s ⁻¹)	Racker band ^c (M ⁻¹ cm ⁻¹)	K_D^a (μ M)	k_{cat}^b (s ⁻¹)	k_{cat}/K_D (μ M ⁻¹ s ⁻¹)	Racker band ^c (M ⁻¹ cm ⁻¹)
Wild-type	0.9 ^d	70 ^d	77 ^d	1100 ^e	35 ^d	— ^f	— ^f	0 ^e
S mutant	1.1 ^d	70 ^d	63 ^d	1100 ^e	36 ^d	15 ^d	0.4 ^d	600 ^e
D32G-S mutant	17 ^g	6 ^e	0.35 ^e	620 ^e	1.3 ^g	19 ^e	14.6 ^e	1060 ^e
B-S mutant	4.4 ^g	80 ^d	18 ^d	1100 ^e	20.8 ^g	18 ^d	0.8 ^d	720 ^e

^a Dissociation constants were determined by affinity labelling with 3-(chloroacetyl)pyridine adenine dinucleotide at 22 °C in 0.1 M phosphate buffer, pH 8.5 (Corbier *et al.*, 1990a).

^b Kinetic parameters were determined at 22 °C in 40 mM triethanolamine, 50 mM phosphate and 0.2 mM EDTA buffer (pH 8.9).

^c Molar extinction coefficients of the Racker band were determined in 0.1 M Tris and 0.02 M EDTA buffer (pH 8). The Racker band corresponds to an absorbance at 360 nm due to a charge transfer complex between the cofactor nicotinamidium ring and residue C149 (Racker & Krinsky, 1952; Mougin *et al.*, 1988).

^d Clermont *et al.* (1993).

^e Rahuel-Clermont (1994).

^f Not active.

^g Eyschen *et al.* (1996).

repulsive electrostatic interactions from occurring between the NADP 2'-phosphate and the protein environment. The B-S mutant more closely resembles the chloroplastic GAPDH in the sense that, besides the 187-188 modifications, it contains three supplementary point mutations in positions 33-34-35, thus allowing the specific chloroplastic fingerprint DTGG to be present in the loop 32-35 of the binding cleft.

Both mutants appear to be less catalytically efficient than the NAD-dependent wild-type enzyme and possess dual-cofactor selectivities with opposite preferences for either NAD⁺ (B-S mutant) or NADP⁺ (D32G-S mutant) (Table 2). In order to correlate these properties with structural particularities elucidated at the molecular level, we have solved the crystal structures of mutants B-S and D32G-S complexed with either NAD⁺ or NADP⁺ (Table 1), thus enabling a direct comparison of NAD and NADP-binding features on exactly the same sites. Furthermore, our results will be also discussed in the light of the previously resolved

crystal structure of the S mutant (bearing only mutations in positions 187 and 188) in complex with NADP⁺ (denoted Sp; resolution 2.7 Å, $R = 16\%$) (G. B., S. R.-C. & A. J. Wonacott, unpublished results).

Results

Quality of models

The obtained high-resolution limits are in the range of 2.5 to 2.2 Å, R -factors are in the range of 14 to 20% and final free R -factors lie in the domain 21 to 27% (Table 3). The geometrical quality of the four refined structures is given by the root mean square (rms) deviations from ideal bond lengths and bond angles (Table 3). Ramachandran plots (not shown) appear to be practically superimposable to that of the wild-type enzyme, with residues A147 and V237 located in a disallowed region of the (ϕ, ψ) space. The high-resolution refined model of the holo wild-type enzyme reveals an almost

Table 3. Refinement statistics and deviations from ideal geometry for the four mutants

	D32G-Sn	D32G-Sp	B-Sn	B-Sp
Resolution range (Å)	8-2.5	8-2.2	8-2.45	8-2.5
R -factor (%)	13.9	19.2	19.3	14.4
R -free (%)	21.0	25.9	26.6	21.0
Number of non-hydrogen atoms:				
Protein atoms	10068	10068	10052	10052
Coenzyme and sulphate atoms	216	232	216	232
Water molecules	384	393	291	341
Average temperature factors (Å ²):				
All protein atoms ^a	27	27	28	30
Main-chain ^a	23	24	24	24
Cofactor ^a	24	22	20	26
Water molecules ^a	35	34	31	37
rms from ideal geometry:				
Bond lengths (Å)	0.009	0.012	0.012	0.009
Angles (°)	1.64	1.83	1.79	1.69
Dihedral angles (°)	25.8	26.1	26.1	24.9
Improper dihedral angles (°)	1.49	1.62	1.67	1.45

^a Excluding hydrogen atoms.

perfect non-crystallographic 222 symmetry in the tetramer (Skarzynski *et al.*, 1987). In accordance with the suggestion given by Mittl *et al.* (1994) in their crystallographic study of glutathione reductase, the atomic position errors in our structures were therefore estimated by the deviations of the individual subunits from ideal 222 molecular symmetry, so giving rise to rms values of 0.26 Å (D32G-Sn), 0.29 Å (D32G-Sp), 0.34 Å (B-Sn) and 0.26 Å (B-Sp), respectively. In each of the four structures, the average temperature factor lies in the same range with an average value of 29 Å² for all protein heavy atoms and an average value of 24 Å² for the main-chain heavy atoms only. The details concerning the B-factors of the NAD(P) molecules will be discussed below. The important buried water molecules already observed in both the holo and apo forms of the wild-type enzyme (Skarzynski *et al.*, 1987; Skarzynski & Wonacott, 1988) are also present in our refined structures.

Overall behaviour of the mutants

The crystal structure of the wild-type GAPDH from *B. stearothermophilus* has been determined in both apo form and holo form (Skarzynski & Wonacott, 1988; Skarzynski *et al.*, 1987). The apo → holo transition can be roughly described as a hinge-bending movement of small amplitude (ca 4°) involving the reorientation of each cofactor-binding domain with respect to its subunit-related catalytic domain. On this basis, a first issue to examine concerns the "holo-like" or "apo-like" behaviour of each mutant, binding either NAD⁺ or NADP⁺. To this purpose, two least-square superimpositions involving all 1336 α -carbon atoms of each complete tetrameric mutant were realized with the wild-type holo and apo structures, respectively. It ensues that both mutants complexed with either NAD⁺ or NADP⁺ can be considered to adopt "holo-like" conformations: whereas the rms values observed in the superimpositions with holo wild-type structure are less than 0.40 Å for all mutants, mean deviations obtained with respect to the apo wild-type structure are very similar to the one mentioned by Skarzynski & Wonacott (1988) as regards the apo/holo wild-type comparison (1.08 Å).

In order to perform a more precise comparison of the mutated GAPDHs with the holo wild-type enzyme, the superimposition method recently described by Duée *et al.* (1996) on the *E. coli* GAPDH was used: for each comparison, an invariant core shared by the two involved molecules is defined on the basis of intra- and inter-subunit C^α–C^α distances showing deviations less than a cut-off of 0.25 Å. The same cut-off was selected in the cases of the Sp, D32G-Sn, D32G-Sp, B-Sp mutants, but a less selective value (0.3 Å) was applied to the B-Sn complex because of the lower quality of its refined three-dimensional model (Table 3). In agreement with Duée *et al.*'s observations concerning the apo/holo comparison of wild-type GAPDH from *B. stearothermophilus*, the procedure gives a rigid

core of approximately 60 residues (15 in each subunit) when applied to our mutants, with the holo wild-type GAPDH as a common reference. Furthermore, all of the invariant cores appear to be located in the same regions of the quaternary structure of GAPDH, with a strong implication of the β 1, β 3, β 4 and β 7 strands of the β -sheet present in the catalytic domains (Figure 1). In addition, some residues belonging to the α 1 helix of the same domains, are also incorporated in the core of the D32G-Sp and B-Sn structures: this fact gives some evidence that the catalytic C149 residue present at the N terminus of the α 1 helix remains practically unperturbed with respect to its position in the wild-type enzyme and therefore accounts for a reinforced holo-like character of both these complexes with respect to the two others. Such a proposal is in good agreement with intensities of the Racker band (Table 2). This absorbance at 360 nm, which is due to a charge transfer complex between the NAD(P)⁺ pyridinium ring and residue C149 (Racker & Krinsky, 1952; Mougin *et al.*, 1988), is indicative of a well-positioning of the thiol group relative to the cofactor. In accordance with their practically perfect holo-like structural features, D32G-Sp and B-Sn mutants display a Racker band value very similar to that measured on the holo wild-type GAPDH (Clermont *et al.*, 1993; Rahuel-Clermont, 1994; Eyschen *et al.*, 1996). Finally, calculations were also performed in order to determine the rotation angles of both catalytic and cofactor-binding domains with respect to their reference positions in the holo wild-type GAPDH (Duée *et al.*, 1996). In agreement with observations mentioned above, no significant rotations of the catalytic domains are detected in the five structures of GAPDH mutants, including the Sp one. The analysis is slightly more complex for the cofactor-binding domains: concerning the Sp (G. B., S. R.-C. & A. J. Wonacott, unpublished results), D32G-Sn and B-Sp structures, the rotation value is about 1°, whereas it appears to be quite negligible in the D32G-Sp and B-Sn cases (Figure 2). This small but noticeable difference supports the previous suggestions relative to the practically perfect holo-like character of D32G-Sp and B-Sn complexes.

The D32G mutation

In the holo wild-type GAPDH from *B. stearothermophilus* (Skarzynski *et al.*, 1987), residue D32 is located at the C-terminal part of the strand β B belonging to the first β z β z β motif of Rossmann fold (Figure 1). Its two carboxylate oxygen atoms are strongly hydrogen-bonded to the 2' and 3'-hydroxyl groups of the adenosine ribose ring of NAD⁺. Moreover, the carboxylate oxygen interacting with the 3'-position is also engaged in a second hydrogen bond involving the main-chain NH group of F8 (Figure 3). Substitution of D32 by a glycine residue has been designed in the D32G-S

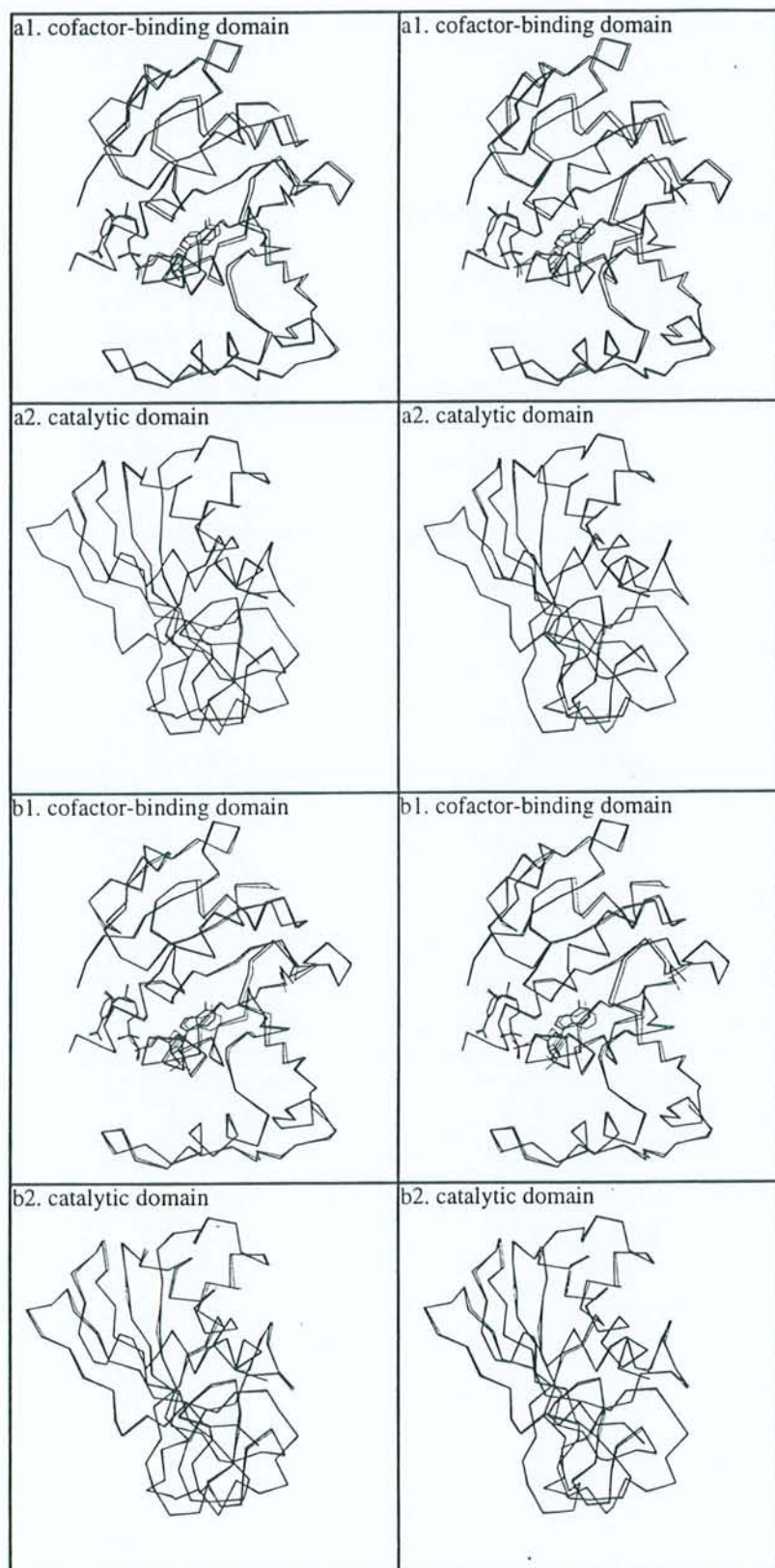


Figure 2. Application of the superimposition method of Duée *et al.* (1996) to the comparison of the D32G-S mutant (thin lines) with the holo wild-type GAPDH (thick lines). Part a1 and a2 refer to subunit O of the D32G-Sn complex. Part b1 and b2 refer to subunit O of the D32G-Sp complex.

mutant so as to circumvent repulsive electrostatic effects which probably account for the low affinity of wild-type GAPDH towards NADP⁺. In the two refined structures of the D32G-S triple mutant

complexed with either NAD⁺ (D32G-Sn) or NADP⁺ (D32G-Sp), the secondary structures surrounding the modified position are not disrupted at all and the newly introduced glycine adopts a

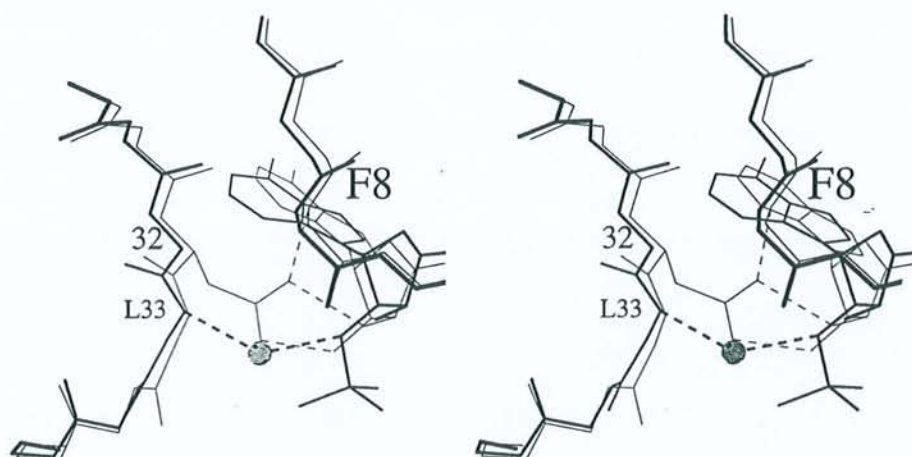


Figure 3. Superimposition of the adenosine-binding subsites in the D32G-Sp mutant (thick lines) and the holo wild-type GAPDH (thin lines). Removal of the D32 side-chain enables a water molecule (W580) to act as a bridge between the adenosine 2'-oxygen atom and the L33 main-chain NH group.

similar main-chain conformation to that experienced by D32 in the wild-type GAPDH. In particular, the G32 NH group appears to be engaged in the same hydrogen bond as the one detected in the wild-type structure (N-Asp/Gly32----O-Lys74). Removal of the side-chain of D32 releases a substantial space which is occupied by at least one water molecule (W580) acting as a bridge between the main-chain NH group of L33 and the 2'-oxygen of the NADP⁺ adenosine moiety in the D32G-Sp structure (Figure 3 and Table 4). In the same mutant complexed with NAD⁺ (D32G-Sn), this water molecule behaves similarly in three of the four monomers.

The B mutations (L33T, T34G, D35G)

In the cofactor-binding domain of GAPDH, the inner surface of the binding cleft essentially comprises residues located in the loops separating each β strand from the next helix along the Rossmann fold (Figure 1). One of these loops goes from the C $^{\alpha}$ carbon of residue 32 to the C $^{\alpha}$ carbon of residue 35. In the crystal structure of holo wild-type GAPDH (Skarzynski *et al.*, 1987), the side-chain of L33 experiences tight van der Waals contacts with the π electronic cloud of the adenine ring in a "stacking-like" manner (Figure 4) (Chakrabarti & Samanta, 1995). One water molecule (W341) acts as a bridge between the main-chain oxygen atoms of residues 34 and 74. No particular interaction is observed for the T34 side-chain. At least, the carboxylate group present in D35 is directly hydrogen-bonded with the main-chain NH of T39 in a kind of interaction typical of α -helix N-capping (Harper & Rose, 1993; Seale *et al.*, 1994). In the B-S mutant, replacement of the LTD triad by TGG enables the loop 32-35 to contain the sequence predominantly detected in these positions for chloroplastic GAPDHs (the other possibility is SGG). Due to the lack of electron density it has not been possible to

position the side-chain of T33 in the B-S mutant bound with either NAD⁺ or NADP⁺. Concerning positions 34 and 35, the conformational freedom which can be reasonably expected from the introduction of glycine residues is effectively observed in the protein main-chain, with significant variations of the torsional angles ϕ and ψ , as mentioned in Table 5. In particular, the 33-34 peptide plane is rotated by about 55°, so giving rise to the formation of a hydrogen bond between its NH group and the carboxylate moiety of the unsubstituted D32 residue (Figure 4). It is interesting to note that the same conformation is observed in the apo wild-type enzyme loop (Skarzynski & Wonacott, 1988). Concomitantly, the water molecule W341 mentioned above is removed. Finally, in the B-Sp structure, one water molecule (W590) fills the free space resulting from the side-chain deletion in position 35, and stands 4 Å apart from the main-chain NH of T39. Unfortunately, no electron density could be observed in the same region in the case of the B-Sn structure, because of its low quality.

The S mutations (L187A, P188S)

In the holo wild-type GAPDH, residues L187 and P188 are located in the middle part of the S-loop of each monomer. They do not participate in the formation of any extensive hydrophobic cluster but rather seem to locally favour an optimal binding of the NAD⁺ cofactor because of their presence in the cleft of the R-related adjacent subunit. Consequently, replacing Leu and Pro by Ala and Ser, respectively, does not produce any conformational transition of the surrounding amino acids. Concerning the behaviour of the new side-chains, the most informative results are supplied in position 188 by the serine residue. In three of the four structures reported here, namely D32G-Sn, B-Sn and B-Sp, the mean thermal factor value for

Table 4. Water molecules interacting with the NAD(P)⁺ adenosine moiety in crystallized wild-type and mutant GAPDHs

		Subunit ^c				Distance ^d (Å)	B-factor ^e (Å ²)
		O	P	Q	R		
O-Wat388 ... N1A-NAD(P)	Wild-type ^a	x	x	x	x	2.9(1)	11
	Sp ^b	x	x	x	x	2.6(3)	31
	D32G-Sn	x	x	x	x	3.0(1)	25
	D32G-Sp	x	x	x	x	3.2(2)	13
	B-Sn	x	x	x	x	3.1(2)	18
O-Wat366 ... O3'A-NAD(P)	B-Sp	x	x	x	x	3.0(2)	20
	Wild-type ^a	x	x	x	x	2.9(1)	21
	Sp ^b		x	x	x	3.4(2)	46
	D32G-Sn	x	x	x		2.7(1)	22
	D32G-Sp	x	x	x	x	2.8(2)	20
O-Wat416 ... OP1A-NAD(P)	B-Sn		x	x	x	2.9(3)	32
	B-Sp	x			x	3.0(1)	36
	Wild-type ^a	x	x	x	x	2.9(1)	18
	Sp ^b						
	D32G-Sn	x				3.5	48
O-Wat430 ... OP1A-NAD(P)	D32G-Sp				x	3.3	24
	B-Sn						
	B-Sp			x	x	3.4(1)	26
	Wild-type ^a	x	x	x	x	2.9(1)	27
	Sp ^b						
O-Wat373 ... OP2A-NAD(P)	D32G-Sn						
	D32G-Sp						
	B-Sn						
	B-Sp	x				2.9	53
	Wild-type ^a	x	x	x	x	2.7(1)	13
O-Wat580 ... O2'A-NAD(P)	Sp ^b	x	x	x	x	2.9(2)	20
	D32G-Sn	x	x	x	x	2.8(2)	11
	D32G-Sp	x	x	x	x	2.7(3)	15
	B-Sn	x	x	x		2.6(1)	15
	B-Sp	x	x	x	x	3.0(1)	17
O-Wat581 ... O3'A-NAD(P)	D32G-Sn	x		x		3.0(1)	43
	D32G-Sp	x				3.0	23
O-Wat582 ... OP3'A-NADP	D32G-Sp	x				2.8	17
O-Wat583 ... OP1'A-NADP	Sp ^b			x		3.4	41
O-Wat584 ... OP2'A-NADP	D32G-Sp			x	x	3.1(1)	48
	D32G-Sp		x	x		2.9(1)	37
O-Wat585 ... O2'A-NADP	B-Sp	x	x			2.8(1)	37
O-Wat366 ... OP2'A-NADP	Sp ^b		x			3.3	47
	D32G-Sp	x	x	x	x	2.5(1)	20

^a Skarzynski *et al.* (1987).^b G. B., S. R.-C. & A. J. Wonacott (unpublished results).^c x, Means that this water molecule is observed in this subunit.^d Average distances for four subunits. Standard deviations are in parentheses.^e Average temperature-factors for four subunits.

the oxygen atom of the side-chain is about 40 Å², whereas this value is only 25 Å² in the D32G-S mutant complexed with NADP⁺ (D32G-Sp). Moreover, the three former structures do not exhibit any electron density feature reflecting a single position for the S188 hydroxyl group, while its rotameric state can be unambiguously fixed in the later structure. In each subunit of the D32G-Sp complex, the hydroxyl group of S188 is appropriately oriented to form a hydrogen bond with one of the three oxygen atoms of the extra 2'-phosphate group present in NADP⁺ (Table 6). Given that this interaction takes place at the interfaces connecting subunits at the quaternary-structure level, this hydrogen bond occurs between S188 of monomer R and the NADP⁺ molecule assigned to monomer O, and *vice versa* (R-related sub-

units). The same relation applies in the case of subunits P and Q.

The overall cofactor-binding site behaviour

Except for NADP⁻ in the D32G-Sp complex, the conformations of all NAD⁻ and NADP⁻ cofactors bound to the GAPDH mutants are quite similar to that adopted by NAD⁻ in the holo wild-type GAPDH (Skarzynski *et al.*, 1987): the adenine and nicotinamide rings are practically perpendicular to the mean planes of their respective riboses, with an *anti* conformation of the adenine moiety and a *syn* orientation for the nicotinamide fragment. All the ribose rings of Sp (G. B., S. R.-C. & A. J. Wonacott, unpublished results) D32G-Sn, B-Sn and B-Sp structures adopt at C2'-*endo* puckering mode iden-

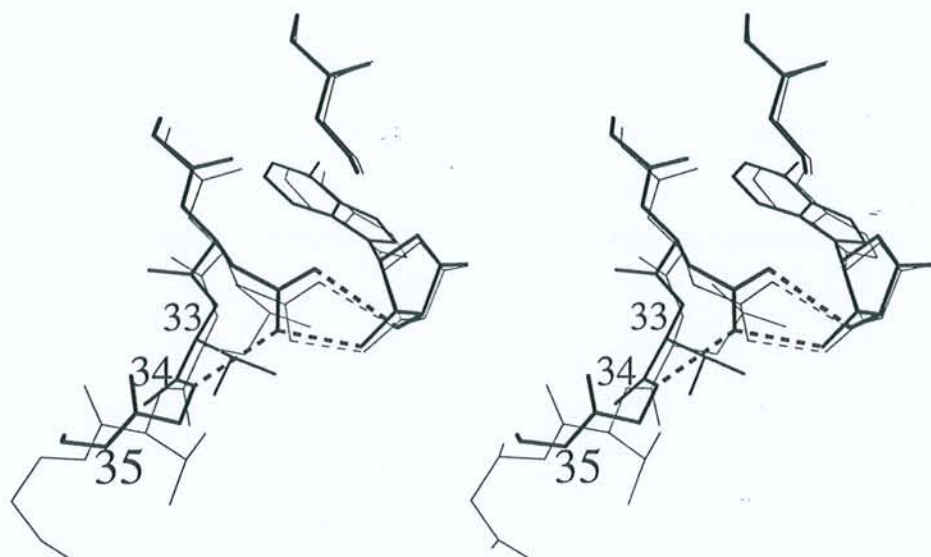


Figure 4. Superimposition of the adenosine-binding subsites in the B-Sn mutant (thick lines) and the holo wild-type GAPDH (thin lines). The B mutations induce a reorientation of the 33-34 peptide plane and the formation of a hydrogen bond between the D32 carboxylate group and the L33 main-chain NH group.

tical to that observed within NAD⁺ when bound to the wild-type GAPDH. In the D32G-Sp complex, the same result holds true in the case of the puckering behaviour of the nicotinamide ribose of NADP⁺, whereas a C1'-*exo* puckering mode is observed for the adenosine moiety. As a consequence, the adenine ring translates as a whole with a shift of about 0.4 Å (Figure 3).

The detailed examination of the thermal vibration factors in the NADP⁺ cofactors shows values for the 2'-phosphate noticeably higher in the Sp (G. B., S. R.-C. & A. J. Wonacott, unpublished results) and B-Sp structures than in the D32G-Sp one (60, 50 and 25 Å², respectively). In the later case, this value is very close to the mean *B*-factor value for the whole NADP⁺ cofactor (Table 3). On the contrary, in the Sp and B-Sp complexes, the thermal motion of the 2'-phosphate group clearly appears to be specifically increased with respect to that of the rest of the cofactor. In fact, in both these mutants, a somewhat free rotation of the tetrahedral phosphate likely occurs along the O2'-P bond, whereas the same group is rigidly embedded in a very well-defined hydrogen-bond network in the D32G-Sp structure (Figure 5).

As expected from the molecular modelling predictions (Corbier *et al.*, 1990a; Clermont *et al.*, 1993), the protein vicinity of the nicotinamide moiety is not altered at all in any of the designed mutants. Short non-covalent distances give evidence for an identical hydrogen-bond network between the protein and the nicotinamide fragment in all the complexes. In particular, the *syn*-stabilizing interaction between the side-chain of N313 (ND2) and the O7N oxygen of the NAD(P) carboxamide (Corbier *et al.*, 1990b; Duée *et al.*, 1996) remains always present. In addition, a water bridge (W380), very similar to that observed in the holo wild-type GAPDH, is also retained in all cases between the carboxamide part of nicotinamide (N7N) and the side-chain of E314 (OE2). In this region, the very well-solved electron density clearly supports the presence of hydrophobic interactions between the nicotinamide ring and the side-chains of I11 and Y317.

Apart from those which involve amino acid residues changed upon mutations, all non-covalent interactions present between the adenosine moiety and the protein remain unaffected in all the mutants. As mentioned above, in the case of the

Table 5. Conformational angles^a of the 31-35 peptide backbone for the holo and apo wild-type GAPDHs and for the B-S mutant bearing either NAD⁺ (B-Sn) or NADP⁺ (B-Sp)

Residue number	31		32		33		34		35	
	φ(°)	ψ(°)	φ(°)	ψ(°)	φ(°)	ψ(°)	φ(°)	ψ(°)	φ(°)	ψ(°)
Holo wild-type ^b	-130(3)	139(1)	-152(2)	-161(3)	-115(4)	28(4)	-145(3)	-172(4)	-92(5)	168(3)
Apo wild-type ^c	-123(4)	123(4)	-166(7)	141(3)	-78(3)	-16(9)	-70(7)	146(4)	-67(7)	150(6)
B-Sn	-129(6)	135(2)	-164(2)	164(7)	-76(6)	-19(8)	-78(5)	158(8)	-76(4)	166(8)
B-Sp	-128(1)	132(2)	-159(4)	166(4)	-76(4)	-24(6)	-80(10)	161(3)	-80(3)	163(3)

^a Average values for four subunits. Standard deviations are in parentheses.

^b Skarzynski *et al.* (1987).

^c Skarzynski & Wonacott (1988).

Table 6. Selected non-covalent distances ^a between the NAD(P)⁺ adenosine moiety and residues of the cofactor-binding site in crystallized wild-type and mutant GAPDHs

	Wild-type ^b	Sp ^c	D32G-Sn	D32G-Sp	B-Sn	B-Sp
N6A-NAD(P) ... O-Arg77	3.2(1)	3.2(2)	3.4(1)	3.2(1)	3.4(2)	3.3(2)
O2'A-NAD(P) ... OD2-Asp32	2.6(1)	3.4(2)			3.0(3)	4.1(2)
O3'A-NAD(P) ... OD1-Asp32	2.6(1)	2.6(2)			2.8(2)	2.8(1)
OP2A-NAD(P) ... N-Arg10	3.2(1)	2.9(2)	2.8(1)	3.0(1)	3.0(1)	3.0(1)
OP2A-NAD(P) ... ND2-Asn180	2.9(1)	3.3(2)	3.1(2)	3.3(2)	3.2(2)	3.2(1)
OP3'A-NADP ... OG-Ser188		4.5(2)		3.0(1)		5.1(4)

^a Average values for four subunits. Standard deviations are in parentheses.^b Skarzynski *et al.* (1987).^c G. B., S. R.-C. & A. J. Wonacott (unpublished results).

D32G-Sp structure, an additional hydrogen bond occurs between the 2'-phosphate group of NADP⁺ and the hydroxyl of S188 (Figure 5). Among the water molecules which mediate cofactor/protein interactions in the holo wild-type structure (Skarzynski *et al.*, 1987), some are practically conserved in all mutants (W373, W388, W366), but others are clearly missing in our structures (W416, W430; Table 4, Figures 5 and 6). It should be also remembered that the D32G mutation provokes the occurrence of a new water molecule (W580) which inserts between the adenosine 2'-oxygen atom of either NAD⁺ or NADP⁺, and the main-chain NH group of L33. In the case of the D32G-Sp structure, some extra water molecules also appear, which are directly linked to the adenosine part of the cofactor (Table 4, Figure 5).

Discussion

Molecular determinants of NAD⁺/NADP⁺ selectivity

As explained in the Introduction, the purpose of this study was to determine what are the main molecular determinants in the Rossmann-fold domain responsible for the NAD⁺/NADP⁺ discrimination. At the biochemical level, *K_D* values are reliable indicators of the relative affinities of wild-type GAPDH and mutants towards both cofactors.

As shown in Table 2, the D32G-S mutant is characterized by a complete reversion of the *K_D* values towards NAD⁺ and NADP⁺ when compared to the wild-type enzyme (Eyschen *et al.*, 1996). Inspection of the three-dimensional crystal structure of this mutant complexed with NAD⁺ strongly suggests that the lower affinity for NAD⁺ is essentially due to the suppression of the two hydrogen bonds which occur between the D32 carboxylate and both 2'- and 3'-positions of the NAD⁺ adenosine in the holo wild-type structure (Figure 3). In the case of NADP⁺ (D32G-Sp complex), the better affinity observed on this mutant with respect to the wild-type enzyme can be in part assigned to the loss of antideterminant effects correlated with the removal of D32 (electrostatic effects) and L187, P188 (steric effects) side-chains. Yet a more interesting feature supplied by the crystal structure concerns the stabilization of the extra 2'-phosphate

group of NADP⁺ by a supplementary hydrogen bond involving the hydroxyl group of S188 (Figure 5). Despite the moderate strength of this kind of interaction when compared to a strong salt bridge (with arginine for example), it can be inferred that such a hydrogen bond certainly plays a critical role in the achievement of a NADP⁺-related *K_D* lying in the μ molar range. Using site-directed mutagenesis, three recently published studies support this hypothesis and emphasize the importance of a hydroxylated side-chain close to the 2'-phosphate for the selective recognition of NADP⁺ in dihydrofolate reductase (Thomas *et al.*, 1994), isocitrate dehydrogenase (Yaoi *et al.*, 1994) and cinnamyl alcohol dehydrogenase (Lauvergeat *et al.*, 1995).

In the B-S mutant, mutations were introduced in order to mimic the chloroplastic enzyme in the adenosine-binding cleft and so to rationally switch to a dual-cofactor selectivity with a preference for NADP. Nevertheless, values reported in Table 2 unambiguously reveal that a better affinity towards NAD⁺ is maintained. At first glance, discrimination between NAD⁺ and NADP⁺ does not appear to proceed from the new residues introduced in the 33-35 peptide chain, since this loop behaves quite similarly in the presence of either NAD⁺ or NADP⁺ (undefined position for the T33 lateral chain, and conformational reorientation of the 33-34 peptide plane with respect to the position adopted in the wild-type structure). Thus, as in the holo wild-type GAPDH, selectivity effects are mainly provided by the presence of the D32 side-chain which acts as both a positive determinant for NAD⁺ (through hydrogen bonding; Figure 4) and an antideterminant for NADP⁺ (through electrostatic repulsion). Moreover, in this mutant, introduction of the chloroplastic sequence in positions 187-188 of the S-loop does not enable the occurrence of a new 2'-phosphate/protein hydrogen bond. So the S mutations only very weakly modulate the original NAD⁺ *versus* NADP⁺ preference observed in the wild-type enzyme. At this stage, it can be likely presumed that a significant feature is missing in our mutant for creating a NADP preference as in the chloroplastic enzyme. As discussed below, the fine structural comparison of the natural and mutated cofactor-binding sites, in conjunction

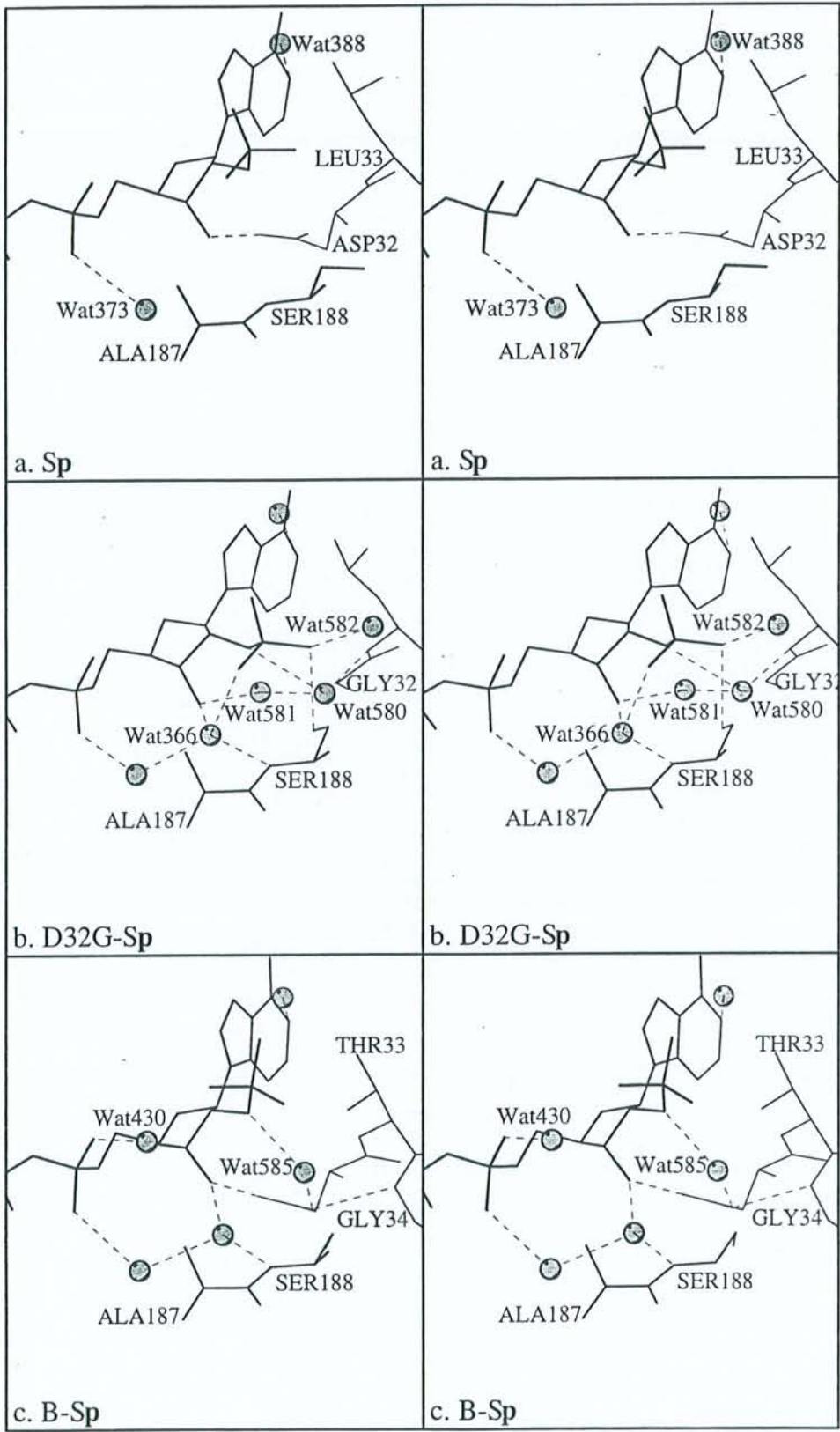


Figure 5. Hydrogen-bond networks in the adenosine-binding subsite of NADP⁺/GAPDH complexes: monomer O of Sp mutant (a; G. B., S. R.-C. & A. J. Wonacott, unpublished results), D32G-Sp mutant (b) and B-Sp mutant (c).

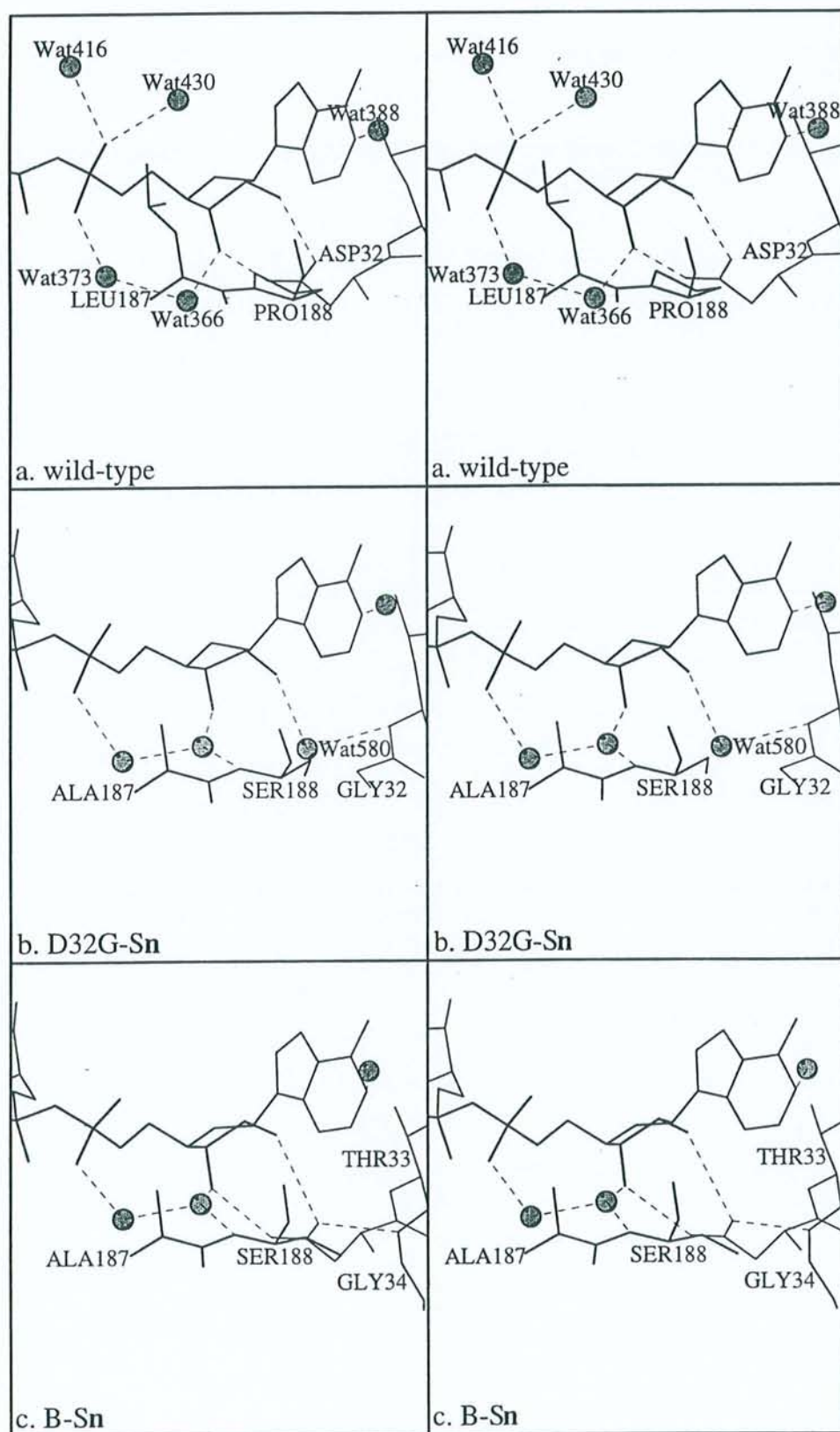


Figure 6. Hydrogen-bond networks in the adenosine-binding subsite of NAD⁺/GAPDH complexes: monomer P of wild-type GAPDH (a; Skarzynski *et al.*, 1987), D32G-Sn mutant (b) and B-Sn mutant (c).

with primary-structure considerations, will perhaps shed some light on this point.

Comparison of NAD⁺-binding features of wild-type, D32G-S and B-S GAPDHs

In the wild-type GAPDH complexed with NAD⁺, the adenosine part of the cofactor is firmly bound to the protein surface because numerous hydrogen bonds and van der Waals contacts occur with the inner surface of the binding cleft (Figure 6). The K_D values measured for NAD⁺ when bound to the D32G-S and B-S mutants clearly reveal a loss of affinity (Table 2). As it is not the case when NAD⁺ interacts with the S mutant which is only modified in positions 187 and 188 (G. B., S. R.-C. & A. J. Wonacott, unpublished results), it can be speculated that this effect essentially arises from the mutations introduced in positions 32 to 35 of the polypeptide chain.

As regards the D32G-S mutant for which the loss of affinity is the most dramatic (about 20 times), all of our observations convincingly support the hypothesis that the lack of hydrogen bonding between the removed D32 carboxylate and the two hydroxyl groups of the ribose is sufficient to explain this behaviour. In the case of the B-S mutant, whose affinity towards NAD⁺ is only moderately affected, it can be suggested that this result proceeds from the conjunction of two weak destabilizing effects: on the one hand, replacement of the bulky side-chain of L33 by the side-chain of a threonine residue actually weakens van der Waals contacts and "stacking-like" phenomena involving the planar adenine part of the cofactor (Chakrabarti & Samanta, 1995), as evidenced by the loss of electron density for the T33 side-chain in the mutant. On the other hand, reorientation of the peptide plane located between residues 33 and 34 enables the main-chain NH of residue 34 to form a hydrogen bond with one of the oxygens of the D32 carboxylate group (Figure 4). In such a case, geometrical considerations published in the literature (Sundaralingam & Sekharudu, 1989) tentatively indicate that the presence of a second hydrogen bond on an acceptor atom significantly alters the strength of the other hydrogen bond, because of steric hindrance and repulsive electrostatic effects between the two proton-donor groups in this bidentate configuration.

Comparison of the NADP⁺-binding features of D32G-S and B-S mutants

Our crystallographic results reveal that the adenosine part of the NADP⁺ cofactor is well stabilized in the D32G-Sp structure, in excellent agreement with a dissociation constant value (1.3 μ M) of the same order as the one measured for NAD⁺ in the holo wild-type enzyme (0.9 μ M). Furthermore, the NADP⁺ 2'-phosphate group appears to be directly involved in this tight binding because it strongly interacts with both the hydroxyl

group of S188 and one water molecule (W366) present in each of the four monomers (Figure 5 and Table 4). Such a behaviour is probably related to the fact that the phosphate group certainly adopts a dianionic state when bound to the D32G-S mutant: all the crystallization studies were carried out at pH 6.9 and NMR experiments have reliably demonstrated that the 2'-phosphate group pK_a is less than 5.7 in the presence of the D32G-S mutant (Eyschen *et al.*, 1996). It must be also emphasized that a further water molecule takes exactly the place of the removed carboxylate group of D32, and forms a critical bridge between the 2'-oxygen of the adenosine ribose and the main-chain NH of L33 (Figure 3). Due to the absence of well-organized interactions between the 2'-phosphate group and the protein environment (Figure 5) the NADP⁺ cofactor is not so well stabilized when bound to the S (G. B., S. R.-C. & A. J. Wonacott, unpublished results) and the B-S mutants, so giving rise to K_D values of 36 and 21 μ M, respectively.

As concerns more specifically the Sp structure (G. B., S. R.-C. & A. J. Wonacott, unpublished results), the ribose moiety of the adenosine only forms one hydrogen bond with the protein (O3'-A-NADP---OD1-Asp32) (Table 6), but nevertheless remains strongly embedded in the NADP⁻-binding cleft (Figure 5). Consequently, a close contact occurs between the D32 carboxylate group and the 2'-phosphate group of NADP⁺, which strongly alters the electrostatic state of the later function: as deduced from the NMR experiments previously reported (Eyschen *et al.*, 1996), the 2'-phosphate pK_a reaches 7.70 in this case and thus leads to a monoanionic state in the domain of pH used for the crystallization experiments. So it can be tentatively concluded that the 2'-phosphate hydrogen-bonding capability is substantially weakened with respect to that observed in the D32G-S mutant, thus hindering good interactions to occur with the protein.

In the case of the B-S mutant, the situation is a little bit more complex because of small structural differences between the four binding sites in the mutated tetramer: two of the four subunits (Q and R) are devoid of hydrogen-bonding interactions between the protein and the adenosine 2'-oxygen atom of NADP⁺, whereas the other two (O and P) contain a water molecule (W585) which acts as a bridge between the 2'-oxygen and one of the two oxygen atoms of the D32 carboxylate group (Figure 5 and Table 4). This water molecule and some others also present in the vicinity of the 2'-phosphate group, probably are allowed to enter the adenosine-binding cleft because of the substantial volume released by the introduction of G34 and G35 in the sequence. Simultaneously, the 2'-phosphate group appears to be more apart from the D32 carboxylate group than in the S mutant (this difference between distances is about 1 Å). Thus, in conjunction with the proton donor potential of the water environment, this phenomenon permits the 2'-phosphate pK_a (6.8) to remain close to the standard value corresponding to a free

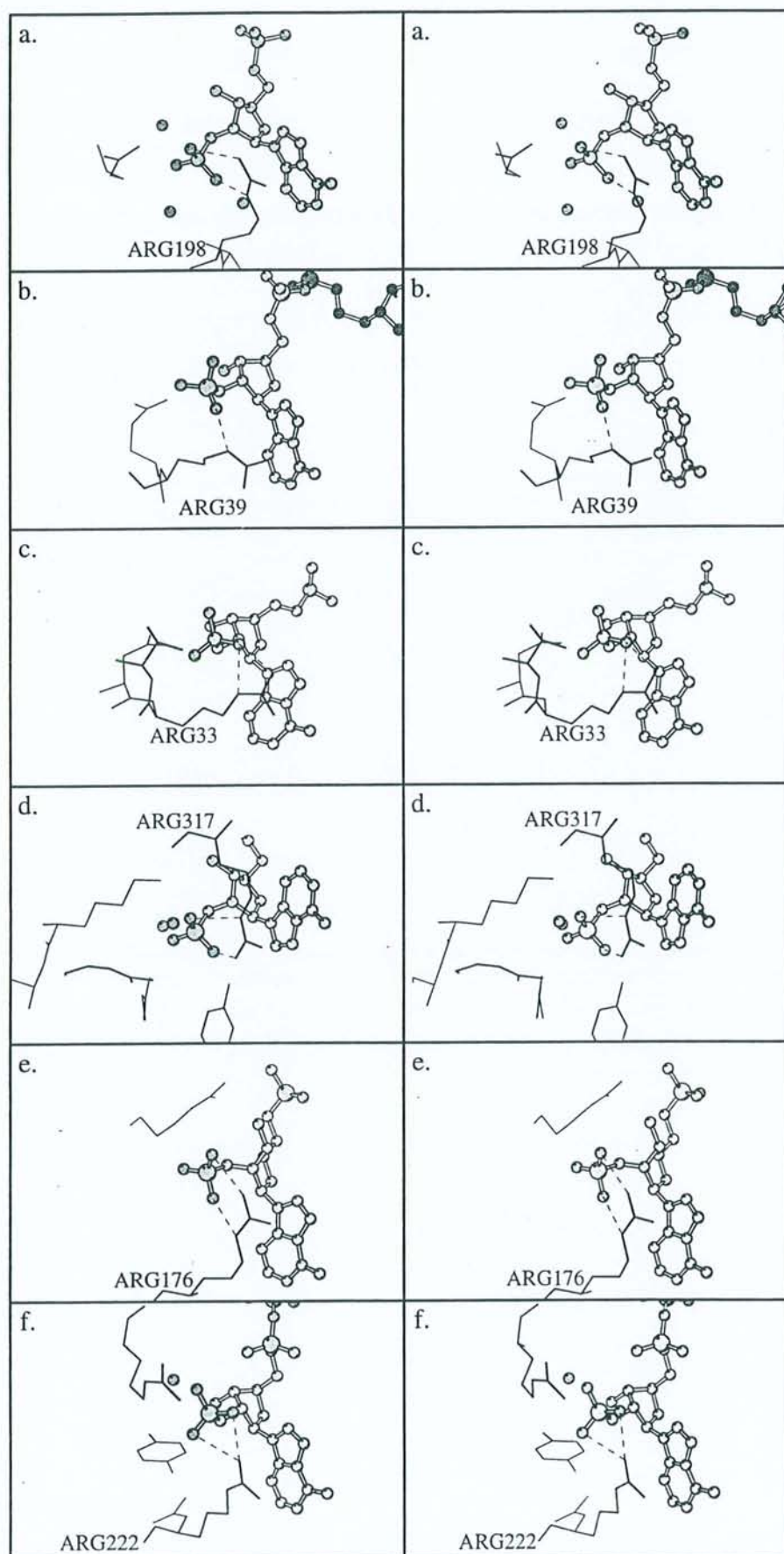


Figure 7. 2'-Phosphate/arginine interactions in crystallized NADP-bound forms of oxidoreductases belonging to the Rossmann-fold family. Dihydrodipicolinate reductase (a; Scapin *et al.*, 1995), glutathione reductase (b; Mittl & Schulz, 1994), 6-phosphogluconate dehydrogenase (c; Adams *et al.*, 1994), quinone oxidoreductase (d; Thorn *et al.*, 1995), thioredoxin reductase (e; Waksman *et al.*, 1994) and trypanothione reductase (f; Bailey *et al.*, 1994).



Figure 8. Adenosine moieties of superimposed NADP molecules extracted from crystallized complexes with oxidoreductases belonging to the Rossmann-fold family. In the D32G-Sp mutant (broken lines), the NADP 2'-phosphate adopts the same disposition as the one observed in five holo wild-type oxidoreductases (thin lines). A dissimilar orientation is only detected in the case of the quinone oxidoreductase from *E. coli* (thick lines) (Thorn *et al.*, 1995).

NADP⁺ molecule in aqueous solution (6.2) (Eyschen *et al.*, 1996). Such an environment nevertheless alters the electronegative strength of the 2'-phosphate with respect to that observed in the D32G-Sp complex and probably precludes a hydrogen bond from occurring with the side-chain hydroxyl group of S188.

Comparison with natural NADP-binding sites

Examination of the Brookhaven Protein Data Bank (Bernstein *et al.*, 1977) reveals that six crystal structures containing NADP/Rossmann fold complexes have been deposited: dihydrodipicolinate reductase from *E. coli* (Scapin *et al.*, 1995), glutathione reductase from *E. coli* (Mittl & Schulz, 1994), 6-phosphogluconate dehydrogenase from sheep (Adams *et al.*, 1994), quinone oxidoreductase from *E. coli* (Thorn *et al.*, 1995), thioredoxin reductase from *E. coli* (Waksman *et al.*, 1994) and trypanothione reductase from *Crithidia fasciculata* (Bailey *et al.*, 1994). Thus an interesting point to consider concerns the comparison of these enzymes with the non-natural NADP⁺-binding sites which have been generated in the double (S) (G. B., S. R.-C. & A. J. Wonacott, unpublished results), triple (D32G-S) and quintuple (B-S) mutants. In the six natural NADP-binding clefts, the most outstanding feature obviously consists of the systematic presence of at least one arginine residue in the close vicinity of the 2'-phosphate group of NADP. This arginine residue always acts as a proton donor towards NADP (Figure 7) by one or more of the five hydrogen atoms of its positively charged guanidinium group. Furthermore, a stacking effect is always observed between the adenine ring and this guanidinium fragment. Except for the quinone oxidoreductase, the arginine residue belongs to the loop located between strand B and helix C along the primary structure, thus corresponding to the

32-35 sequence in GAPDH. In the quinone oxidoreductase, the arginine residue is located in the $\alpha 5$ helix of the catalytic domain (Thorn *et al.*, 1995). Therefore, the position of this residue along the polypeptide chain seems not to be critical, provided that a spatial proximity with the 2'-phosphate group would be achieved in the three-dimensional structure. Besides the arginine residue, a noteworthy feature of three of the six natural NADP-binding sites consists of the presence of a hydroxylated side-chain close to the 2'-phosphate group: threonine in the 6-phosphogluconate dehydrogenase, tyrosine in the quinone oxidoreductase and trypanothione reductase (Figure 7c, d, e).

Among the mutated GAPDHs presented here, none of them shows an arginine residue interacting with the 2'-phosphate group. In fact, the 2'-phosphate group is recognized in the sole D32G-S mutant, because of a hydrogen bond with the hydroxyl group of S188. In conjunction with the release of electrostatic repulsion due to the D32 substitution in this mutant, this interaction allows the enzyme/cofactor affinity to be of the same order as the one observed for wild-type NAD(P)/oxidoreductase complexes.

The least square superimposition of the NADP molecules from wild-type complexes and NADP⁺ bound to the D32G-S mutant indicates that the latter does not conformationally differ from the vast majority of the former (Figure 8). In fact, the only different NADP conformation is observed in the quinone oxidoreductase structure (Thorn *et al.*, 1995) where the 2'-phosphate group points in the direction of the protein core (Figure 7d). In this case, the 2'-phosphate group is directed towards non-bulky residues present in the loop $\beta B \rightarrow \alpha C$ equivalent to that of GAPDH in positions 32 to 35. Concomitantly, it is interesting to note the unexpected *syn* conformation of the adenine ring in the binding cleft. In all of the other natural

NADP/enzyme complexes, the presence of an arginine in the loop $\beta\beta \rightarrow \alpha C$ orientates the 2'-phosphate group towards the outside of the NADP-binding sites, as in our mutant structures (Figure 8).

Crystal structures of NADP/enzyme complexes have also been solved in the case of proteins not belonging to the Rossmann-fold family. Among them, dihydrofolate reductase appears to be the most studied with five crystalline NADP-bound holo forms from *Lactobacillus casei* (Bolin *et al.*, 1982; Filman *et al.*, 1982), *E. coli* (Bystroff *et al.*, 1990), chicken (McTigue *et al.*, 1992), *Pneumocystis carinii* (Champness *et al.*, 1994) and man (Cody *et al.*, 1992). Other available data concern isocitrate dehydrogenase from *E. coli* (Hurley *et al.*, 1991; Stoddard *et al.*, 1993), aldose reductase from man (Wilson *et al.*, 1992) and catalases from beef liver (Fita & Rossmann, 1985) and *Proteus mirabilis* (Gouet *et al.*, 1995). In both catalases, the NADP 2'-phosphate group only interacts with one positively charged arginine residue. In each of the other seven models, the 2'-phosphate moiety is deeply embedded in a network of hydrogen bonds involving cationic residues (arginine, lysine and protonated histidine) as well as hydroxylated ones: serine in dihydrofolate reductase from man, chicken, *E. coli* and in aldose reductase from man; threonine in dihydrofolate reductase from *L. casei* and *P. carinii*; tyrosine in isocitrate dehydrogenase from *E. coli*. Moreover, as explained above, experiments based upon site-directed mutagenesis have established that the selective recognition of NADP by *L. casei* dihydrofolate reductase (Thomas *et al.*, 1994) and *Thermus thermophilus* isocitrate dehydrogenase (Yaoi *et al.*, 1994) substantially depends on those hydroxylated side-chains. As judged by these results, a 2'-phosphate/hydroxyl hydrogen bond certainly constitutes an efficient way to reinforce the stabilization of the NADP adenosine moiety in a binding cleft, but such an interaction more often than not appears to be used in combination with salt bridges containing other hydrogen bonds. Therefore, the D32G-Sp structure turns out to be a convincing crystallographic demonstration that a NADP-preferring binding site may result from the sole presence of a hydroxylated amino acid in the close neighbourhood of the cofactor 2'-phosphate.

Remarks about the chloroplastic GAPDHs

The chloroplastic GAPDHs which should belong to the Rossmann-fold family (Martin & Cerff, 1986; Shih *et al.*, 1986) recognize both NAD and NADP with a preference for NADP. These enzymes possess an aspartic acid residue in position 32 located at the C terminus of the second strand of the first $\beta\alpha\beta\alpha\beta$ motif. This position, when occupied by an acidic residue, is conventionally considered as indicative of a NAD-specific enzyme (Wierenga *et al.*, 1986; Scrutton *et al.*, 1990; Bork & Grunwald, 1990). Yet, in the dihydronicotinamide reductase mentioned above, the equivalent position is occupied by a glutamate residue chemically close to the as-

partic one. In this case, the enzyme exhibits a dual-cofactor selectivity with a weak preference for NAD⁺ (Scapin *et al.*, 1995).

Therefore, the definition of this aspartic acid residue as an antideterminant factor towards NADP is not as evident as usually postulated, and it clearly ensues that its effect on NAD/NADP discrimination must be largely modulated by other amino acids located in its close vicinity. In our experiments, the quintuple B-S mutant was designed with the hope of obtaining a "chloroplastic-like" behaviour as regards NAD/NADP selectivity. Dissociation constants, kinetic parameters (Clermont *et al.*, 1993) and crystallographic data indicate that the mutations introduced are not sufficient or well enough adapted to build a chloroplastic-like enzyme. Yet, and in the light of the primary structure alignments (Corbier *et al.*, 1990a), it can be asserted that all cationic residues potentially able to lie in the 2'-phosphate neighbourhood in chloroplastic enzymes are also present in the *B. stearothermophilus* enzyme but do not adopt the required orientations for interacting with the 2'-phosphate. Thus, one must tentatively conclude that second-order or long-range effects, not easily detectable or predictable, are of crucial importance for designing an efficient chloroplastic-like GAPDH mutant.

Concerning this matter, an interesting suggestion to consider involves the S-loop region which strongly participates in inter-subunit interfaces, in particular at the level of the cofactor-binding cleft. R183 and G180 are two residues which are conserved in all presently known primary structures of chloroplastic GAPDHs. In the *B. stearothermophilus* enzyme and in our mutants, R183 is also present but G180 is replaced by an asparagine residue. As explained earlier, no interaction occurs between R183 (monomer R) and the NADP⁻ 2'-phosphate (monomer O) in the mutants, and its guanidinium group only hydrogen bonds with the main-chain carbonyl of S188. As for the N180 residue, which is not systematically present in cytosolic enzymes, its side-chain amide group (ND2) forms a hydrogen bond with an oxygen atom (OP2A) of the NAD(P)⁺ pyrophosphate moiety in the particular case of the *B. stearothermophilus* enzyme (Table 6). So removing this interaction by the introduction of a glycine in position 180 would perhaps significantly enhance the flexibility of the S-loop and thus enable a reorientation and a translocation of the R183 side-chain (monomer R) in the direction of the extra 2'-phosphate of NADP⁺ (monomer O). Hopefully the problem will be solved when the crystal structure of the chloroplastic GAPDH in holo form will be available.

Materials and Methods

Preparation and crystallization

Production and purification of the mutants were performed as described by Clermont *et al.* (1993). Each protein, supplied as an ammonium sulphate precipitate,

Table 7. Crystallographic data for the four mutants

	<i>a</i> (Å)	<i>b</i> (Å)	<i>c</i> (Å)	β (°)	Space group	Resolution limit (Å)	Completeness % ^a	Unique reflections <i>F</i> > 2σ(<i>F</i>)	Unique reflections all data	Redundancy	<i>R</i> _{sym} %
D32G-Sn	82.95	124.68	82.97	108.63	<i>P</i> 2 ₁	2.5	81	43095	45468	2.7	6.8
D32G-Sp	82.37	124.27	82.45	108.93	<i>P</i> 2 ₁	2.2	79	65530	70186	2.4	9.7
B-Sn	82.87	124.80	82.92	108.72	<i>P</i> 2 ₁	2.45	90	53718	56855	3.1	7.8
B-Sp	134.14	123.96	96.73	90.0	<i>P</i> 2 ₁ 2 ₁ 2	2.5	89	48832	52001	2.5	5.4

^a Completeness of the set of reflections used in the refinement.

was dialysed three times in a 3 ml Centricon against 30 mM NAD(P)⁺, 200 mM phosphate buffer (pH 7.2), 1 mM EDTA and 1 mM DTT to replace any bound or degraded coenzyme with fresh NAD(P)⁺. Excess NAD(P)⁺ was removed by further cycles of dialysis in the Centricon with the above solution, but leaving out NAD(P)⁺. The final concentration of the protein solution was adjusted to 7 mg/ml in 2 M ammonium sulphate, 30 mM NAD(P)⁺, 50 mM phosphate buffer (pH 6.9), 1 mM EDTA and 1 mM DTT. Hanging drops of 10 µl were set up against a reservoir of 2.5 M ammonium sulphate at pH 6.9. For each mutant, crystals grew in less than two weeks until a suitable size for X-ray analysis was attained. The D32G-Sp, D32G-Sn, B-Sn crystals grew as monoclinic prisms in space group *P*2₁ with one tetramer per asymmetric unit. These crystals are isomorphous to that of the wild-type holoenzyme (Skarzynski *et al.*, 1987). The B-Sp crystals belong to space group *P*2₁2₁2, with one tetramer per asymmetric unit. In this case, the unit cell dimensions are very similar to those of native GAPDH grown with one molecule of NAD⁺ per tetramer ("HOLO I" crystal form; Leslie & Wonacott, 1983) and to those of the S mutant complexed with NADP⁺ (G. B., S. R.-C. & A. J. Wonacott, unpublished results). Details of cell parameters are given in Table 7.

Data collection

The data sets of the D32G-Sp and B-Sn complexes were collected at room temperature at the Grenoble LCM/IBS laboratory on a 30 cm Mar-Research area detector with a rotating anode generator (RU200) using CuK_α radiation, to a resolution of 2.2 and 2.45 Å, respectively. The data were processed with the program XDS (Kabsch, 1988, 1993) and were merged with the ROTAVATA procedure from the program package CCP4 (Collaborative Computing Project Number 4, 1994). The D32G-Sn and B-Sp crystals were measured at room temperature on our home area detector (Dip2000) with a Kappa goniometer using CuK_α radiation from a rotating anode generator (Nonius FR591 Model) equipped with a graphite monochromator. Both 2.5 Å resolution data sets were processed and scaled with the Denzo software and Scalepack program, respectively (Otwinowski, 1993). The principal crystallographic data are summarized in Table 7.

Model building and refinement

The crystal structures of GAPDH mutants complexed with either NAD⁺ or NADP⁺ were solved by using the structural coordinates of the *B. stearothermophilus* holo wild-type GAPDH refined at 1.8 Å resolution (access number 1gd1, Protein Data Bank; Bernstein *et al.*, 1977).

At the starting point of the refinement procedure, the mutated amino acid residues were replaced by glycine residues, and the cofactors, the sulphate ions and water molecules were not included in the models. The refinement was carried out with the program X-PLOR version 3.1 (Brünger *et al.*, 1987; Brünger, 1992a,b) using the potential parameters of Engh & Huber (1991). The models were examined and manually rebuilt with the graphic program TURBO-FRODO (Roussel & Cambillau, 1989). Refinement included data (*F* > 2σ) in the range 8 to 2.5 Å, 8 to 2.2 Å, 8 to 2.45 Å and 8 to 2.5 Å for the structures of D32G-Sn, D32G-Sp, B-Sn and B-Sp complexes, respectively. The refinements were performed without the use of non-crystallographic symmetry restraints. Rigid body refinement of each structure converged very quickly and was followed by a simulated annealing procedure (Brünger *et al.*, 1990): during the molecular dynamics simulation each system was initially heated to 2000 K and slowly cooled to 300 K in 50 K steps with 0.5 fs time step. Next, 120 conventional positional-refinement steps were applied to complete this refinement cycle. At this stage, the calculated electron density maps (*F_o* - *F_c*) gave unambiguously the positions of the cofactors, which were introduced in each model. Simulated annealing and positional refinement were then repeated and followed by a temperature-factor refinement. The water molecules and sulphate ions were then added to build the solvent network around each mutant: each time a peak on the (*F_o* - *F_c*) maps was contoured at 3 σ, it was tentatively identified as a water molecule, provided that favorable interactions would be allowed between this site and the protein. Finally, this molecule was conserved in the model if it was contoured at 1 σ on the (2*F_o* - *F_c*) map from the following refinement. As expected, most of the water molecules present in the first hydration shell of each mutant appeared to lie in similar positions to those observed in the wild-type structure. At this stage the (*F_o* - *F_c*) maps confirmed the respective mutations of residues Leu187, Pro188 and Asp32 into an alanine, a serine and a glycine in the D32G-S mutant. In the B-S mutant, the Leu187Ala, Pro188Ser, Thr34Gly and Asp35Gly mutations were also confirmed in the residual map, whereas the Leu33Thr mutation could not be firmly demonstrated because of a very poor electron density for the side-chain of this residue. Then, the models were adjusted and refined until the final convergence was reached. Residuals are listed in Table 3. Illustrations were prepared with the MOLSCRIPT software (Kraulis, 1991).

Data deposition

The atomic coordinates have been deposited in the Protein Data Bank, Chemistry Department, Brookhaven National Laboratory, Upton, NY 11973, with the follow-

ing accession codes: 1dbv (D32G-Sn), 2dbv (D32G-Sp), 3dbv (B-Sn) and 4dbv (B-Sp).

Acknowledgements

We are grateful to Drs Brantant, Rahuel-Clermont & Wonacott for providing us the atomic coordinates of the Sp mutant. This work was financially supported by the Centre National de la Recherche Scientifique, the University Henri Poincaré Nancy I and the Institut Fédératif de Recherche Protéines. This work was technically supported by the Service Commun de Diffraction X sur Monocristaux and the Service Commun de Biophysicochimie des Interactions Moléculaires of the University Henri Poincaré Nancy I.

References

- Adams, M. J., Ellis, G. H., Gover, S., Naylor, C. E. & Phillips, C. (1994). Crystallographic study of coenzyme, coenzyme analogue and substrate binding in 6-phosphogluconate dehydrogenase: implications for NADP specificity and the enzyme mechanism. *Structure*, **2**, 651–668.
- Bailey, S., Fairlamb, A. H. & Hunter, W. N. (1994). Structure of trypanothione reductase from *Crithidia fasciculata* at 2.6 Å resolution; enzyme-NADP interactions at 2.8 Å resolution. *Acta Crystallog. sect. D*, **50**, 139–154.
- Baker, P. J., Britton, K. L., Rice, D. W., Rob, A. & Stillman, T. J. (1992). Structural consequences of sequence patterns in the fingerprint region of the nucleotide binding fold. Implications for nucleotide specificity. *J. Mol. Biol.* **228**, 662–671.
- Bernard, N., Johnsen, K., Holbrook, J. J. & Delcour, J. (1995). D175 discriminates between NADH and NADPH in the coenzyme binding site of *Lactobacillus delbrueckii* subsp. *bulgaricus* D-lactate dehydrogenase. *Biochem. Biophys. Res. Commun.* **208**, 895–900.
- Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr, Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* **112**, 535–542.
- Biesecker, G., Harris, J. I., Thierry, J. C., Walker, J. E. & Wonacott, A. J. (1977). Sequence and structure of D-glyceraldehyde 3-phosphate dehydrogenase from *Bacillus stearothermophilus*. *Nature*, **266**, 328–333.
- Bocanegra, J. A., Scrutton, N. S. & Perham, R. N. (1993). Creation of an NADP-dependent pyruvate dehydrogenase multienzyme complex by protein engineering. *Biochemistry*, **32**, 2737–2740.
- Bolin, J. T., Filman, D. J., Matthews, D. A., Hamlin, R. C. & Kraut, J. (1982). Crystal structures of *Escherichia coli* and *Lactobacillus casei* dihydrofolate reductase refined at 1.7 Å resolution. I. General features and binding of methotrexate. *J. Biol. Chem.* **257**, 13650–13662.
- Bork, P. & Grunwald, C. (1990). Recognition of different nucleotide-binding sites in primary structures using a property-pattern approach. *Eur. J. Biochem.* **191**, 347–358.
- Brünger, A. T. (1992a). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, **355**, 472–475.
- Brünger, A. T. (1992b). *X-PLOR Manual, version 3.1*, Yale University, Newhaven.
- Brünger, A. T., Kuriyan, J. & Karplus, M. (1987). Crystallographic R-factor refinement by molecular dynamics. *Science*, **229**, 458–460.
- Brünger, A. T., Krukowski, A. & Erickson, J. W. (1990). Slow-cooling protocols for crystallographic refinement by simulated annealing. *Acta Crystallog. sect. A*, **46**, 585–593.
- Buehner, M., Ford, G. C., Moras, D., Olsen, K. W. & Rossmann, M. G. (1974). Three-dimensional structure of D-glyceraldehyde-3-phosphate dehydrogenase. *J. Mol. Biol.* **90**, 25–49.
- Bystroff, C., Oatley, S. J. & Kraut, J. (1990). Crystal structures of *Escherichia coli* dihydrofolate reductase: the NADP⁺ holoenzyme and the folate-NADP⁺ ternary complex. Substrate binding and a model for the transition state. *Biochemistry*, **29**, 3263–3277.
- Cerff, R. (1978). Glyceraldehyde-3-phosphate dehydrogenase (NADP) from *Sinapis alba* L. NAD(P)-induced conformation changes of the enzyme. *Eur. J. Biochem.* **82**, 45–53.
- Cerff, R. (1979). Quaternary structure of higher plant glyceraldehyde-3-phosphate dehydrogenases. *Eur. J. Biochem.* **94**, 243–247.
- Chakrabarti, P. & Samanta, U. (1995). CH/π interaction in the packing of the adenine ring in protein structures. *J. Mol. Biol.* **251**, 9–14.
- Champness, J. N., Achari, A., Ballantine, S. P., Bryant, P. K., Delves, C. J. & Stammers, D. K. (1994). The structure of *Pneumocystis carinii* dihydrofolate reductase to 1.9 Å resolution. *Structure*, **2**, 915–924.
- Chen, R., Greer, A. & Dean, A. M. (1995). A highly active decarboxylating dehydrogenase with rationally inverted coenzyme specificity. *Proc. Natl Acad. Sci. USA*, **92**, 11666–11670.
- Chen, Z., Lee, W. R. & Chang, S. H. (1991). Role of aspartic acid 38 in the cofactor specificity of *Drosophila* alcohol dehydrogenase. *Eur. J. Biochem.* **202**, 263–267.
- Chen, Z., Tsigelny, I., Lee, W. R., Baker, M. E. & Chang, S. H. (1994). Adding a positive charge at residue 46 of *Drosophila* alcohol dehydrogenase increases cofactor specificity for NADP⁺. *FEBS Letters*, **356**, 81–85.
- Clermont, S., Corbier, C., Mely, Y., Gérard, D., Wonacott, A. & Brantant, 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**, 10178–10184.
- Cody, V., Luft, J. R., Ciszak, E., Kalman, T. I. & Freisheim, J. H. (1992). Crystal structure determination at 2.3 Å of recombinant human dihydrofolate reductase ternary complex with NADPH and methotrexate-γ-tetrazole. *Anti-cancer Drug Des.* **7**, 483–491.
- Collaborative Computing Project Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.
- Corbier, C., Clermont, S., Billard, P., Skarzynski, T., Brantant, C., Wonacott, A. & Brantant, G. (1990a). Probing the coenzyme specificity of glyceraldehyde-3-phosphate dehydrogenases by site-directed mutagenesis. *Biochemistry*, **29**, 7101–7106.
- Corbier, C., Mougin, A., Mely, Y., Adolph, H. W., Zeppezauer, M., Gérard, D., Wonacott, A. &

- Branlant, G. (1990b). The nicotinamide subsite of glyceraldehyde-3-phosphate dehydrogenase studied by site-directed mutagenesis. *Biochimie*, **72**, 545–554.
- Duée, E., Olivier-Deyris, L., Fanchon, E., Corbier, C., Branlant, G. & Dideberg, O. (1996). Comparison of the structures of wild-type and a N313T mutant of *Escherichia coli* glyceraldehyde 3-phosphate dehydrogenase: implication for NAD binding and cooperativity. *J. Mol. Biol.* **257**, 814–838.
- El-Kabbani, O., Green, N. C., Lin, G., Carson, M., Narayana, S. V. L., Moore, K. M., Flynn, T. G. & DeLucas, L. J. (1994). Structures of human and porcine aldehyde reductase: an enzyme implicated in diabetic complications. *Acta Crystallog. sect. D*, **50**, 859–868.
- El-Kabbani, O., Judge, K., Ginell, S. L., Myles, D. A. A., DeLucas, L. J. & Flynn, T. G. (1995). Structure of porcine aldehyde reductase holoenzyme. *Nature Struct. Biol.* **2**, 687–692.
- Engh, R. A. & Huber, R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallog. sect. A*, **47**, 392–400.
- Eyschen, J., Vitoux, B., Rahuel-Clermont, S., Marraud, M., Branlant, G. & Cung, M. T. (1996). Phosphorous-31 nuclear magnetic resonance studies on coenzyme binding and specificity in glyceraldehyde-3-phosphate dehydrogenase. *Biochemistry*, **35**, 6064–6072.
- Fan, F., Lorenzen, J. A. & Plapp, B. V. (1991). An aspartate residue in yeast alcohol dehydrogenase I determines the specificity for coenzyme. *Biochemistry*, **30**, 6397–6401.
- Ferri, G., Comerio, G., Iadarola, P., Zapponi, M. C. & Speranza, M. L. (1978). Subunit structure and activity of glyceraldehyde-3-phosphate dehydrogenase from spinach chloroplasts. *Biochim. Biophys. Acta*, **522**, 19–31.
- Ferri, G., Stoppini, M., Meloni, M. L., Zapponi, M. C. & Iadarola, P. (1990). Chloroplast glyceraldehyde-3-phosphate dehydrogenase (NADP): amino acid sequence of the subunits from isoenzyme I and structural relationship with isoenzyme II. *Biochim. Biophys. Acta*, **1041**, 36–42.
- Filman, D. J., Bolin, J. T., Matthews, D. A. & Kraut, J. (1982). Crystal structures of *Escherichia coli* and *Lactobacillus casei* dihydrofolate reductase refined at 1.7 Å resolution. II. Environment of bound NADPH and implications for catalysis. *J. Biol. Chem.* **257**, 13663–13672.
- Fita, I. & Rossmann, M. G. (1985). The NADPH binding site on beef liver catalase. *Proc. Natl Acad. Sci. USA*, **82**, 1604–1608.
- Gouet, P., Jouve, H.-M. & Dideberg, O. (1995). Crystal structure of *Proteus mirabilis* PR catalase with and without bound NADPH. *J. Mol. Biol.* **249**, 933–954.
- Grimshaw, C. E., Matthews, D. A., Varughese, K. I., Skinner, M., Xuong, N. H., Bray, T., Hoch, J. & Whiteley, J. M. (1992). Characterization of nucleotide binding properties of a mutant dihydropteridine reductase containing an aspartate 37-isoleucine replacement. *J. Biol. Chem.* **267**, 15334–15339.
- Haeflner-Gormley, L., Chen, Z., Zalkin, H. & Colman, R. F. (1992). Importance of lysine-286 at the NADP site of glutamate dehydrogenase from *Salmonella typhimurium*. *Biochemistry*, **31**, 7807–7814.
- Hanukoglu, I. & Gutfinger, T. (1989). cDNA sequence of adrenodoxin reductase. Identification of NADP-binding sites in oxidoreductases. *Eur. J. Biochem.* **180**, 479–484.
- Harper, E. T. & Rose, G. D. (1993). Helix stop signals in proteins and peptides: the capping box. *Biochemistry*, **32**, 7605–7609.
- Harris, J. I. & Waters, M. (1976). Glyceraldehyde-3-phosphate dehydrogenase. In *The Enzymes* (Boyer, P. D., ed.), vol. 13, pp. 1–49, Academic Press, New York.
- Hoog, S. S., Pawlowski, J. E., Alzari, P. M., Penning, T. M. & Lewis, M. (1994). Three-dimensional structure of rat liver 3 α -hydroxysteroid/dihydrodiol dehydrogenase: a member of the aldo-keto reductase superfamily. *Proc. Natl Acad. Sci. USA*, **91**, 2517–2521.
- Huang, S., Appleman, J. R., Tan, X., Thompson, P. D., Blakley, R. L., Sheridan, R. P., Venkataraghavan, R. & Freisheim, J. H. (1990). Role of lysine-54 in determining cofactor specificity and binding in human dihydrofolate reductase. *Biochemistry*, **29**, 8063–8069.
- Hurley, J. H. & Dean, A. M. (1994). Structure of 3-isopropylmalate dehydrogenase in complex with NAD⁺: ligand-induced loop closing and mechanism for cofactor specificity. *Structure*, **2**, 1007–1016.
- Hurley, J. H., Dean, A. M., Koshland, D. E., Jr & Stroud, R. M. (1991). Catalytic mechanism of NADP⁺-dependent isocitrate dehydrogenase: implications from the structures of magnesium-isocitrate and NADP⁺ complexes. *Biochemistry*, **30**, 8671–8678.
- Hurley, J. H., Chen, R. & Dean, A. M. (1996). Determinants of cofactor specificity in isocitrate dehydrogenase: structure of an engineered NADP⁺ $\rightarrow$ NAD⁺ specificity-reversal mutant. *Biochemistry*, **35**, 5670–5678.
- Kabsch, W. (1988). Evaluation of single-crystal X-ray diffraction data from a position-sensitive detector. *J. Appl. Crystallog.* **21**, 916–924.
- Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallog.* **26**, 795–800.
- Kim, H., Feil, I. K., Verlinde, C. L. M. J., Petra, P. H. & Hol, W. G. J. (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**, 14975–14986.
- Korndörfer, I., Steipe, B., Huber, R., Tomschy, A. & 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**, 511–521.
- Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946–950.
- Lauvergeat, V., Kennedy, K., Feuillet, C., McKie, J. H., Gorrichon, L., Baltas, M., Boudet, A. M., Grima-Pettenati, J. & Douglas, K. T. (1995). Site-directed mutagenesis of a serine residue in cinnamyl alcohol dehydrogenase, a plant NADPH-dependent dehydrogenase, affects the specificity for the coenzyme. *Biochemistry*, **34**, 12426–12434.
- Lesk, A. M. (1995). NAD-binding domains of dehydrogenases. *Curr. Opin. Struct. Biol.* **5**, 775–783.
- Leslie, A. G. W. & Wonacott, A. J. (1983). Coenzyme binding in crystals of glyceraldehyde-3-phosphate dehydrogenase. *J. Mol. Biol.* **165**, 375–391.

- Levy, H. R., Vought, V. E., Yin, X. & Adams, M. J. (1996). Identification of an arginine residue in the dual coenzyme-specific glucose-6-phosphate dehydrogenase from *Leuconostoc mesenteroides* that plays a key role in binding NADP⁺ but not NAD⁺. *Arch. Biochem. Biophys.* **326**, 145–151.
- Lin, Z.-J., Li, J., Zhang, F.-M., Song, S.-Y., Yang, J., Liang, S.-J. & Tsou, C.-L. (1993). Structure of D-glyceraldehyde-3-phosphate dehydrogenase from *Palinurus versicolor* carrying the fluorescent NAD derivatives at 2.7 Å resolution. *Arch. Biochem. Biophys.* **302**, 161–166.
- Martin, W. & Cerff, R. (1986). Prokaryotic features of a nucleus-encoded enzyme. cDNA sequences for chloroplast and cytosolic glyceraldehyde-3-phosphate dehydrogenases from mustard (*Sinapis alba*). *Eur. J. Biochem.* **159**, 323–331.
- Mathews, F. S. (1991). New flavoenzymes. *Curr. Opin. Struct. Biol.* **1**, 954–967.
- McTigue, M. A., Davies, J. F., II, Kaufman, B. T. & Kraut, J. (1992). Crystal structure of chicken liver dihydrofolate reductase complexed with NADP⁺ and bipterin. *Biochemistry*, **31**, 7264–7273.
- Mercer, W. D., Winn, S. I. & Watson, H. C. (1976). Twinning in crystals of human skeletal muscle D-glyceraldehyde-3-phosphate dehydrogenase. *J. Mol. Biol.* **104**, 277–283.
- Metzger, M. H. & Hollenberg, C. P. (1995). Amino acid substitutions in the yeast *Pichia stipitis* xylitol dehydrogenase coenzyme-binding domain affect the coenzyme specificity. *Eur. J. Biochem.* **228**, 50–54.
- Mittl, P. R. E. & Schulz, G. E. (1994). Structure of glutathione reductase from *Escherichia coli* at 1.86 Å resolution: comparison with the enzyme from human erythrocytes. *Protein Sci.* **3**, 799–809.
- Mittl, P. R. E., Berry, A., Scrutton, N. S., Perham, R. N. & Schulz, G. E. (1994). Anatomy of an engineered NAD-binding site. *Protein Sci.* **3**, 1504–1514.
- Miyazaki, K. & Oshima, T. (1994). Co-enzyme specificity of 3-isopropylmalate dehydrogenase from *Thermus thermophilus* HB8. *Protein Eng.* **7**, 401–403.
- Moras, D., Olsen, K. W., Sabesan, M. N., Buehner, M., Ford, G. C. & Rossmann, M. G. (1975). Studies of asymmetry in the three-dimensional structure of lobster D-glyceraldehyde-3-phosphate dehydrogenase. *J. Biol. Chem.* **250**, 9137–9162.
- Mougin, 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**, 45–48.
- Nishiyama, M., Birktoft, J. J. & Beppu, T. (1993). Alteration of coenzyme specificity of malate dehydrogenase from *Thermus flavus* by site-directed mutagenesis. *J. Biol. Chem.* **268**, 4656–4660.
- Otwinowski, Z. (1993). Oscillation data reduction program. In *Proceedings of the CCP4 Study Weekend: Data Collection and Processing* (Sawyer, L., Isaacs, N. & Bailey, S., eds), pp. 56–62, SERC Daresbury Laboratory, Warrington.
- Pai, E. F. (1991). Variations on a theme: the family of FAD-dependent NAD(P)H-(disulphide)-oxidoreductases. *Curr. Opin. Struct. Biol.* **1**, 796–803.
- Racker, E. & Krimsky, I. (1952). The mechanism of oxidation of aldehydes by glyceraldehyde-3-phosphate dehydrogenase. *J. Biol. Chem.* **198**, 731–743.
- Rahuel-Clermont, S. (1994). Etude des facteurs moléculaires responsables de la spécificité de cofacteur des déshydrogénases à NAD(P). Thèse d'Université, Université Henri Poincaré, Nancy.
- Rondeau, J.-M., Tête-Favier, F., Podjarny, A., Reymann, J.-M., Barth, P., Biellmann, J.-F. & Moras, D. (1992). Novel NADPH-binding domain revealed by the crystal structure of aldose reductase. *Nature*, **355**, 469–472.
- Rossmann, M. G., Adams, M. J., Buehner, M., Ford, G. C., Hackert, M. L., Liljas, A., Rao, S. T., Banaszak, L. J., Hill, E., Tsernoglou, D. & Webb, L. (1973). Molecular symmetry axes and subunit interfaces in certain dehydrogenases. *J. Mol. Biol.* **76**, 533–537.
- Rossmann, M. G., Moras, D. & Olsen, K. W. (1974). Chemical and biological evolution of a nucleotide-binding protein. *Nature*, **250**, 194–199.
- Rossmann, M. G., Liljas, A., Brändén, C.-I. & Banaszak, L. J. (1975). Evolutionary and structural relationships among dehydrogenases. In *The Enzymes* (Boyer, P. D., ed.), vol. 11, pp. 61–102, Academic Press, New York.
- Roussel, A. & Cambillau, C. (1989). TURBO-FRODO. In *Silicon Graphics Geometry Partner Directory* (Silicon Graphics, ed.), pp. 77–78, Silicon Graphics, Mountain View.
- Scapin, G., Blanchard, J. S. & Sacchettini, J. C. (1995). Three-dimensional structure of *Escherichia coli* dihydrodipicolinate reductase. *Biochemistry*, **34**, 3502–3512.
- Schulz, G. E. (1992). Binding of nucleotides by proteins. *Curr. Opin. Struct. Biol.* **2**, 61–67.
- Scrutton, N. S., Berry, A. & Perham, R. N. (1990). Redesign of the coenzyme specificity of a dehydrogenase by protein engineering. *Nature*, **343**, 38–43.
- Seale, J. W., Srinivasan, R. & Rose, G. D. (1994). Sequence determinants of the capping box, a stabilizing motif at the N-termini of α -helices. *Protein Sci.* **3**, 1741–1745.
- Sem, D. S. & Kasper, C. B. (1993a). Enzyme-substrate binding interactions of NADPH-cytochrome P-450 oxidoreductase characterized with pH and alternate substrate/inhibitor studies. *Biochemistry*, **32**, 11539–11547.
- Sem, D. S. & Kasper, C. B. (1993b). Interaction with arginine 597 of NADPH-cytochrome P-450 oxidoreductase is a primary source of the uniform binding energy used to discriminate between NADPH and NADH. *Biochemistry*, **32**, 11548–11558.
- Shih, M.-C., Lazar, G. & Goodman, H. M. (1986). Evidence in favor of the symbiotic origin of chloroplasts: primary structure and evolution of tobacco glyceraldehyde-3-phosphate dehydrogenases. *Cell*, **47**, 73–80.
- Skarzynski, T. & Wonacott, A. J. (1988). Coenzyme-induced conformational changes in glyceraldehyde-3-phosphate dehydrogenase from *Bacillus stearothermophilus*. *J. Mol. Biol.* **203**, 1097–1118.
- Skarzynski, T., Moody, P. C. E. & Wonacott, A. J. (1987). Structure of holo-glyceraldehyde-3-phosphate dehydrogenase from *Bacillus stearothermophilus* at 1.8 Å resolution. *J. Mol. Biol.* **193**, 171–187.
- Stoddard, B. L., Dean, A. & Koshland, D. E., Jr (1993). Structure of isocitrate dehydrogenase with isocitrate, nicotinamide adenine dinucleotide phosphate, and calcium at 2.5 Å resolution: a pseudo-Michaelis ternary complex. *Biochemistry*, **32**, 9310–9316.
- Stryer, L. (1995). *Biochemistry*, 4th edit., Freeman & Co, New York.

- Sundaralingam, M. & Sekharudu, Y. C. (1989). Water-inserted α -helical segments implicate reverse turns as folding intermediates. *Science*, **244**, 1333–1337.
- Tanner, J. J., Hecht, R. M. & Krause, K. L. (1996). Determinants of enzyme thermostability observed in the molecular structure of *Thermus aquaticus* D-glyceraldehyde-3-phosphate dehydrogenase at 2.5 Å resolution. *Biochemistry*, **35**, 2597–2609.
- Thomas, J. A., Arnold, J. R. P., Basran, J., Andrews, J., Roberts, G. C. K., Birdsall, B. & Feeney, J. (1994). Effects of substitution of Thr63 by alanine on the structure and function of *Lactobacillus casei* dihydrofolate reductase. *Protein Eng.* **7**, 783–792.
- Thorn, J. M., Barton, J. D., Dixon, N. E., Ollis, D. L. & Edwards, K. J. (1995). Crystal structure of *Escherichia coli* QOR quinone oxidoreductase complexed with NADPH. *J. Mol. Biol.* **249**, 785–799.
- Vellieux, F. M. D., Hajdu, J. & Hol, W. G. J. (1995). Refined 3.2 Å structure of glycosomal holo glyceraldehyde phosphate dehydrogenase from *Trypanosoma brucei brucei*. *Acta Crystallog. sect. D*, **51**, 575–589.
- Waksman, G., Krishna, T. S. R., Williams, C. H., Jr & Kuriyan, J. (1994). Crystal structure of *Escherichia coli* thioredoxin reductase refined at 2 Å resolution. *J. Mol. Biol.* **236**, 800–816.
- Watson, H. C., Duée, E. & Mercer, W. D. (1972). Low resolution structure of Glyceraldehyde-3-phosphate dehydrogenase. *Nature New Biol.* **240**, 130–139.
- Wierenga, R. K., Terpstra, P. & Hol, W. G. J. (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**, 101–107.
- Wilson, D. K., Bohren, K. M., Gabbay, K. H. & Quijcho, F. A. (1992). An unlikely sugar substrate site in the 1.65 Å structure of the human aldose reductase holoenzyme implicated in diabetic complications. *Science*, **257**, 81–84.
- Yaoi, T., Miyazaki, K. & Oshima, T. (1994). Role of Arg²³¹ and Tyr²⁸⁴ of *Thermus thermophilus* isocitrate dehydrogenase in the coenzyme specificity. *FEBS Letters*, **355**, 171–172.

Edited by R. Huber

(Received 11 December 1996; received in revised form 17 February 1997; accepted 20 February 1997)

DIII COMPARAISON STRUCTURALE DE 25 ENZYMES DE TYPE 'ROSSMANN FOLD'

DIIIa Introduction

Les études comparatives réalisées par Wierenga *et al.* (1986), Bork & Grunwald (1990) et Baker *et al.* (1992b) sur les enzymes présentant le repliement caractéristique décrit par Rossmann *et al.* (1975) ont mis en évidence une empreinte caractéristique constituée de six résidus consécutifs de la boucle β A- α B et de l'extrémité N-terminale de l'hélice α B. La séquence, nommée de P1 à P6 (Baker *et al.*, 1992b), est de la forme GXGXXG/A mise à part celle de la malate déshydrogénase qui est de la forme GXXGXXG (où X désigne un résidu quelconque). A partir d'un ensemble de sept structures déterminées à haute résolution, Baker *et al.* (1992b) ont mis en évidence deux familles structurales, suivant qu'une glycine ou une alanine se situe en position P6. Scrutton *et al.* (1990) et Mittl *et al.* (1994) ont montré que la nature de ce résidu en P6 joue un rôle essentiel (effet de deuxième couche) dans la spécificité de cofacteur dans la glutathione réductase (cf. § AIIb en particulier la figure A6). Depuis 1992 de nombreuses structures supplémentaires d'enzymes constituées d'un repliement de type 'Rossmann fold' et présentant des variations au niveau de la signature ont été déterminées à haute résolution.

Dans notre étude sur la spécificité de cofacteur NAD/NADP dans les GAPDH, les résultats enzymatiques et structuraux sur le quintuple mutant B-S ont montré que les cinq mutations modélisées à partir de la séquence de la GAPDH chloroplastique n'étaient pas suffisantes pour créer une dualité de cofacteur à l'avantage du NADP dans la GAPDH de *Bacillus stearothermophilus*.

Cette dernière partie présente une comparaison de 25 structures disponibles dans la PDB (Bernstein *et al.*, 1977) complexées soit avec NAD, soit avec NADP (tableau D1) pour tenter de mettre en évidence les résidus de la GAPDH *B. stearothermophilus* responsables d'effets de deuxième couche ou d'effets à longue distance dans la spécificité de cofacteur.

Tableau D1 : Les 25 complexes enzyme/cofacteur sélectionnés dans l'étude comparative

entrée PDB	Enzyme	organisme	cofac.	Rés. (Å)	Rf (%)	Référence
2OHX	alcool déshydrogénase (LADH)	<i>Equus caballus</i> (foie)	NAD	1.8	17.3	Al-Karadaghi <i>et al.</i> (1994)
9LDT	lactate déshydrogénase (LDH)	<i>sus scrofa</i>	NAD	2.0	23.3	Dunn <i>et al.</i> (1991)
1BMD	malate déshydrogénase (MDH)	<i>Thermus flavus</i>	NAD	1.9	15.4	Kelly <i>et al.</i> (1993)
1HYH	L-2-hydroxiisocaproate déshydrogénase (LHicDH)	<i>Lactobacillus</i> <i>confusus</i>	NAD	2.2	21.	Niefind <i>et al.</i> (1995)
2NAD	formate déshydrogénase (FDH)	<i>Pseudomonas sp</i> <i>101</i>	NAD	2.0	11.4	Lamzin <i>et al.</i> (1994)
2DLD	D-lactate déshydrogénase (DLDH)	<i>Lactobacillus</i> <i>pentosus</i>	NAD	2.7	19.5	Bernard <i>et al.</i> (rés. non publiés)
1PSD	D-phosphoglycérate déshydrogénase (PGDH)	<i>Escherichia coli</i>	NAD	2.75	17.7	Schuller <i>et al.</i> (1995)
1DHR	dihydropyridine réductase (DHPiR)	<i>Rattus norvegicus</i> (foie)	NAD	2.3	^a	Varughese <i>et al.</i> (1992)
2HSD	3 α ,20 β -hydroxystéroïde déshydrogénase (3 α 20 β HSD)	<i>Streptomyces</i> <i>hydrogenans</i>	NAD	2.6	18.8	Ghosh <i>et al.</i> (1994)
1FMC	7 α -hydroxystéroïde déshydrogénase (7 α HSD)	<i>Escherichia coli</i>	NAD	1.8	20.7	Tanaka <i>et al.</i> (1996)
1NAI	UDP-galactose-4- épimérase (GALE)	<i>Escherichia coli</i>	NAD	2.0	19.3	Thoden <i>et al.</i> (1996)
1ENP	enoyl acyl carrier protéine réductase (ENR)	<i>Brassica napus</i>	NAD	1.9	16.3	Rafferty <i>et al.</i> (1995)
1GD1	glycéraldéhyde-3-phosphate déshydrogénase (GAPDH)	<i>Bacillus</i> <i>stearothermophilus</i>	NAD	1.8	17.7	Skarzynski <i>et al.</i> (1987)
1ENY	protéine InhA (InhA)	<i>Mycobacterium</i> <i>tuberculosis</i>	NAD	2.2	19.6	Dessen <i>et al.</i> (1995)
2NPX	NADH peroxidase (NPX)	<i>Streptococcus</i> <i>faecalis</i>	NAD	2.4	15.9	Stehle <i>et al.</i> (1993)
1LVL	lipoamide déshydrogénase (LipDH)	<i>Pseudomonas</i> <i>putida</i>	NAD	2.45	21.5	Mattevi <i>et al.</i> (1992)
1DRU	dihydrodipicolinate réductase (DHPiR)	<i>Escherichia coli</i>	NAD	2.2	19.2	Reddy <i>et al.</i> (1996)
1DIH	dihydrodipicolinate réductase (DHPiR)	<i>Escherichia coli</i>	NADP	2.2	18.7	Scapin <i>et al.</i> (1995)
1QOR	quinone oxydoréductase (QOR)	<i>Escherichia coli</i>	NADP	2.2	14.	Thorn <i>et al.</i> (1995)
1FDT	17 β -hydroxystéroïde déshydrogénase (17 β HSD)	<i>Homo sapiens</i>	NADP	2.2	19.3	Ghosh <i>et al.</i> (1995)
1CYD	carbonyle réductase (MLCR)	<i>Mus musculus</i>	NADP	1.8	15.4	Tanaka <i>et al.</i> (1996)
1PGO	6-phosphogluconate déshydrogénase (6PGDH)	<i>Ovis aries</i> (foie)	NADP	2.5	17.	Adams <i>et al.</i> (1994)
1GET	glutathione réductase (GR)	<i>Escherichia coli</i>	NADP	2.	20.8	Mittl & Schulz (1994)
1TYP	trypanothionéréductase (TR)	<i>Crithidia fasciculata</i>	NADP	2.8	14.8	Bailey <i>et al.</i> (1994)
1TDF	thiorédoxine réductase (ThR)	<i>Escherichia coli</i>	NADP	2.3	17.9	Waksman <i>et al.</i> (1994)

^a valeur non indiquée dans le fichier PDB.

DIIIb Mise en évidence de cinq familles

Nous avons posé le problème de façon complémentaire par rapport à la partie 'comparison with natural NADP-binding sites' développée dans la publication : au lieu de superposer les dinucléotides pour comparer les voisinages directs des 2'-phosphates du NADP dans les différents sites, ce sont les 'Rossmann fold' qui ont été superposés pour comparer les différentes positions adoptées par l'adénosine 5'-diphosphate (2'-phosphate) dans les 25 structures. La protéine de référence utilisée est la GAPDH de *B. stearothermophilus* (Skarzynski *et al.*, 1987). Pour superposer localement deux 'Rossmann fold', seuls les carbones α de 12 résidus du brin βA , de la boucle βA - αB et de l'hélice αB ont été utilisés (du troisième résidu avant la position P1 au troisième résidu après la position P6, soit dans la GAPDH du résidu 4 à 16) (Swindells, 1993). La figure D1 présente cinq conformations différentes de la séquence caractéristique, mises en évidence par la superposition des 25 structures.

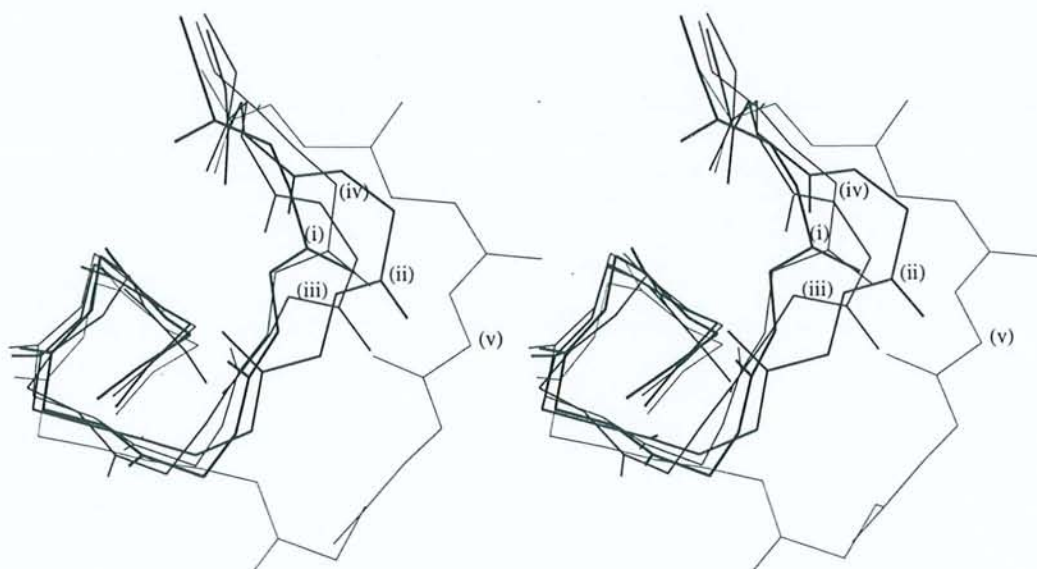


Figure D1 : Superposition des signatures présentant des conformations différentes de cinq 'Rossmann fold'

(i) signature G/AXG/AXXG (LADH représentée) ; (ii) signature GXXXGXG (3a20bHSD représentée) ; (iii) signature G/AXXGXG (QOR représentée) ; (iv) signature GXGXXA/S (TR représentée) ; (v) signature GXXXXXXXXG/A (ENR représentée)

Tableau D2 : Alignement des séquences caractéristiques des 24 enzymes sélectionnées dans l'étude comparative suivant la nomenclature de Baker *et al.* (1992b).

Enzyme	famille	N° du premier résidu	stucture primaire									
			P1					P6				
LDH	i	28	G	V	-	-	-	G	A	V	G	
LADH	i	199	G	L	-	-	-	G	G	V	G	
FDH	i	198	A	A	-	-	-	G	R	I	G	
LHicDH	i	28	G	L	-	-	-	G	N	V	G	
GAPDH	i	7	G	F	-	-	-	G	R	I	G	
DLDH	i	153	G	T	-	-	-	G	H	I	G	
PGDH	i	158	G	Y	-	-	-	G	H	I	G	
NPX	i	156	G	S	-	-	-	G	Y	I	G	
LipDH	i	178	G	G	-	-	-	G	Y	I	G	
6PGDH	i	9	G	L	-	-	-	A	V	M	G	
3 α 20 β HSD	ii	13	G	G	A	-	-	R	G	L	G	
7 α HSD	ii	18	G	A	G	-	-	A	G	I	G	
MLCR	ii	14	G	A	G	-	-	K	G	I	G	
17 β HSD	ii	8	G	C	S	-	-	S	G	I	G	
MDH	iii	10	G	A	A	-	-	G	Q	I	G	
GALE	iii	7	G	G	S	-	-	G	Y	I	G	
DHPiR	iii	13	G	G	R	-	-	G	A	L	G	
QOR	iii	148	A	A	A	-	-	G	G	V	G	
DHPiR	iii	12	G	A	G	-	-	G	R	M	G	
GR	iv	174	G	A	-	-	-	G	Y	I	A	
TR	iv	195	G	G	-	-	-	G	Y	I	S	
Thr	iv	152	G	G	-	-	-	G	N	T	A	
ENR	v	25	G	I	A	D	D	N	G	T	G	
InhA	v	14	G	I	I	T	D	S	S	I	A	

L'appartenance d'une enzyme à une famille donnée semble être en relation directe avec la position des glycines ou des alanines dans la séquence caractéristique :

i) La première famille possède la séquence caractéristique G/AXG/AXXG mise en évidence par Wierenga *et al.* (1986). Elle comprend 9 enzymes NAD-dépendantes (GADPH, LADH, LDH, NPX, DLDH, PGDH, LHicDH, FDH, LipDH) et une enzyme NADP dépendante (6PGDH) (tableau D1).

ii) La deuxième famille ne contient que des enzymes de la famille des oxydoréductases à chaîne courte. Deux sont NAD-dépendantes (3 α 20 β HSD, 7 α HSD) et deux

NADP-dépendantes (17 β HSD, MLCR). Sa séquence caractéristique GXXXGXG contient un résidu de plus que celle de la première famille (tableau D2).

iii) La signature de la troisième famille est de la forme G/AXXGXXG. Trois enzymes sont NAD-dépendantes (MDH, DHPtR, GALE), une NADP-dépendante (QOR) et enfin une à dualité de cofacteur, cocrystallisée avec NAD et NADP (DHPiR). Il est intéressant de remarquer la présence d'une oxydoréductases et d'une épimérase à chaîne courte, en particulier la DHPtR. Celle-ci possède entre les résidus 14 et 20 la séquence GGRGALG. Tanaka *et al.* (1996) ont identifié dans cette structure primaire la signature de la deuxième famille GXXXAXG. Cependant la superposition des empreintes caractéristiques de la MDH, de la DHPtR et de la MLCR montre bien que la conformation adoptée par la séquence peptidique de la DHPtR est plus proche de celle de la MDH que de celle de la MLCR (figure D2). La signature de la DHPtR se présente donc sous la forme GXXGXXG.

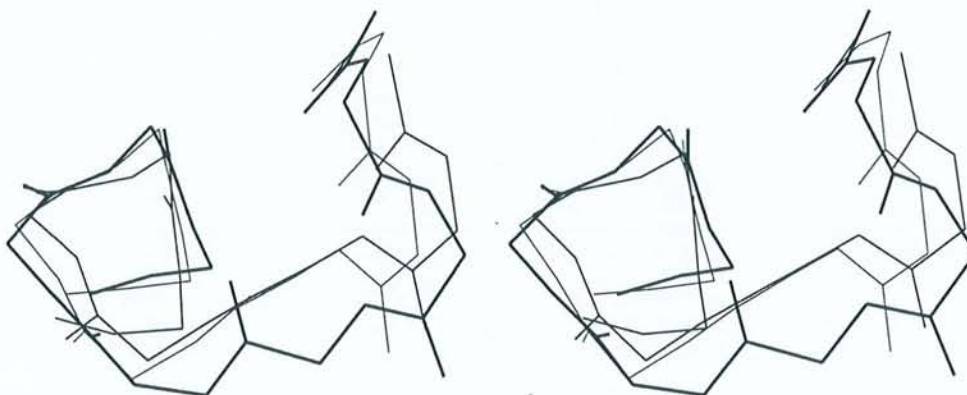


Figure D2 : Mise en évidence de l'appartenance de la DHPtR à la famille iii.

Superposition des signatures de la MLCR (traits épais), MDH (traits moyens), DHPtR (traits fins).

iv) Les trois flavoenzymes NADP-dépendantes (GR, TR, ThR) forment la quatrième famille (signature GXGXXA/S), avec comme seule différence avec la première famille la présence d'une alanine en position P6 à la place d'une glycine. Baker *et al.* (1992a,b) ont montré que la signature GXGXXA/S de la glutamate déshydrogénase de *Clostridium symbiosum* NAD-dépendante possède les mêmes caractéristiques structurales que celle de la

GR. La présence d'une alanine en P6 suffit à changer la conformation de la chaîne polypeptidique par rapport à celle de la signature caractéristique GXGXXG. En effet, le carbonyle du résidu en P1 ne peut plus pointer vers le carbone α du résidu en P6 comme dans la première famille à cause du méthyle latéral de l'alanine (figure D3). La liaison peptidique entre les résidus situés en P1 et P2 subit alors une rotation d'environ 90° . Dans la GAPDH (famille i) et la GR (famille iv) les angles dièdres ψ du résidu P1 valent respectivement $+95^\circ$ et -173° .

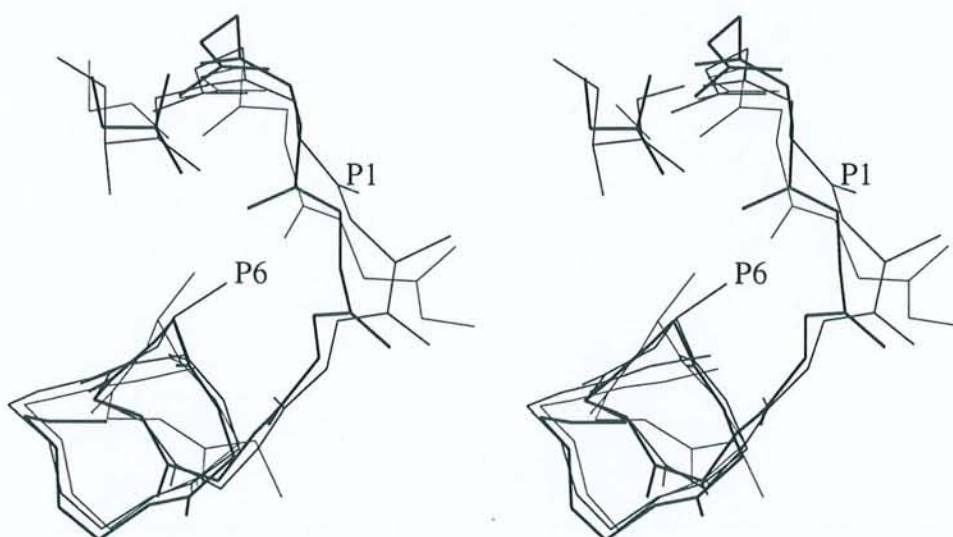


Figure D3 : Orientation du carbonyle du résidu localisé en P1 dans les familles i (traits gras), iv (traits moyens) et v (traits fins)

Les enzymes représentées pour les familles i, iv, et v sont respectivement la GAPDH, la GR et l'InhA.

v) La boucle β A- α B des deux énoyle réductases est plus longue de trois résidus par rapport à celle de la première famille. Il n'est pas possible de mettre évidence la glycine caractéristique localisée entre les deux autres. La signature se présente sous la forme GXXXXXXXXG/A. La nature du résidu en dernière position de la séquence caractéristique ne conduit pas comme dans les enzymes présentant la signature GXGXXG/A à deux familles distinctes. En effet, dans les deux structures, les carbonyles des résidus en P1 pointent dans une direction équivalente, différente de celles observées dans les familles i et iv : ils se dirigent

vers l'intérieur de la boucle $\beta A-\alpha B$ conduisant à un angle dièdre ψ du résidu en P1 d'environ 30° (figure D3). C'est l'insertion des trois résidus dans la séquence caractéristique qui rend cette conformation possible.

Comme nous l'avons précisé dans la partie A1b2, les structures secondaires du 'Rossmann fold' les mieux conservées sont les brins βA , βB , βC et βD ainsi que l'hélice αB . Celle-ci constitue principalement le site de fixation du pyrophosphate et de la partie adénosine du cofacteur. Il n'est donc pas étonnant de remarquer que ces deux parties du dinucléotide sont les mieux superposées dans les 25 structures comparées. Cependant, ces superpositions présentent des déviations notables exposées famille par famille dans le paragraphe suivant.

DIIIc Comportement du pyrophosphate dans les cinq familles

DIIIc1 La famille i (G/AXG/AXXG)

Exceptés les cas de la LipDH et de la DLDH, les parties adénosine 5'-diphosphate des enzymes NAD-dépendantes de cette famille possèdent une position commune par rapport à la chaîne polypeptidique de la signature G/AXG/AXXG, nommée position référence (valable pour les cinq familles présentées dans cette étude). Trois interactions communes peuvent être relevées : une liaison hydrogène entre le groupe NH de l'acide aminé précédant le dernier résidu de la signature et un oxygène du 5'-phosphate de la partie NMN ainsi que deux liaisons hydrogène entre le carboxyle du résidu acide en P7 et les hydroxyles du ribose de l'adénosine (voir dans le cas de la GAPDH sur les figures A14b, c). Les pyrophosphates des cofacteurs complexés dans la DLDH et dans la LipDH ne se trouvent pas dans la position référence. Cette différence semble être liée dans les deux cas au quatrième résidu de la signature (Tyr181 dans la LipDH et His159 dans la DLDH). En effet dans les deux structures les positions des chaînes latérales empêchent les pyrophosphates de se trouver dans la position référence par contrainte stérique. Il est intéressant de noter que la quatrième position des séquences caractéristiques de la NPX et de la PGDH sont occupées respectivement par une tyrosine (comme dans la LipDH) et une histidine (comme dans la DLDH) et que les chaînes latérales de ces résidus ne pointent pas vers les pyrophosphates qui se trouvent alors dans la position référence.

La signature caractéristique de la seule enzyme NADP-dépendante (6PGDH) de cette famille possède une alanine en troisième position alors que toutes les autres NAD-dépendantes présentent une glycine. Le méthyle de cette alanine, en pointant dans la direction du pyrophosphate du NADP, empêche celui-ci d'adopter la position référence (figure D4).

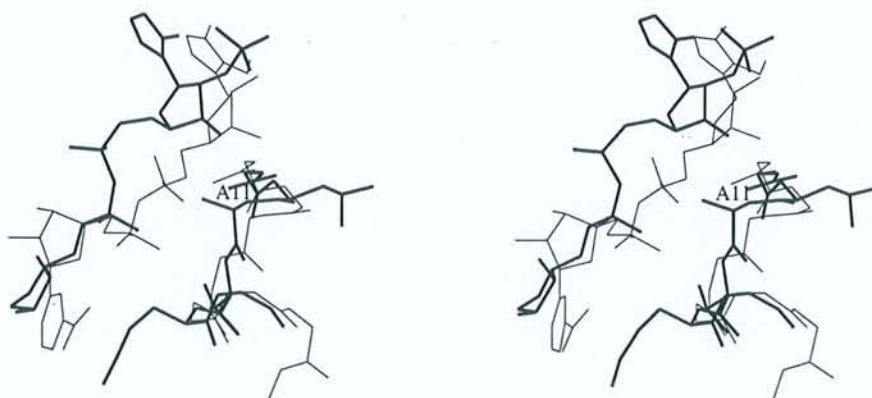


Figure D4 : Superposition du site de fixation du pyrophosphate de la GAPDH (traits fins) avec celui de la 6PGDH (traits épais).

L'alanine 11 de la 6PGDH empêche le pyrophosphate du NADP de se positionner comme dans la GAPDH.

DIIIc2 La famille ii (GXXXGXXG)

Sur les quatre oxydoréductases à courte chaîne, la partie adénosine 5'-diphosphate 2'-phosphate du cofacteur NADP ne se trouve pas dans la position référence uniquement dans la 17 β HSD. Ce changement de position semble être dû à la nature du troisième résidu de la séquence caractéristique de cette famille (tableau D2). Dans la 17 β HSD cette troisième position est occupée par une sérine alors que l'on trouve soit une glycine soit une alanine dans les trois autres. La chaîne latérale de cette sérine empêche le ribose du NADP de se trouver dans la position référence.

La remarque la plus intéressante de ce paragraphe concerne la position du cofacteur dans la structure de MLCR. Sur les sept enzymes NADP-dépendantes comparées dans cette étude, c'est le seul cas où la partie adénosine 5'-diphosphate 2'-phosphate du NADP se trouve dans la position référence (figure D5). Tanaka *et al.* (1996) ont montré que le site de fixation du 2'-

phosphate dans la MLCR est favorable à la stabilisation des charges négatives portées par ce groupement, d'une part par la présence de deux résidus basiques (Lys17 situé dans la signature et Arg39 situé dans la boucle β B- α C) et d'autre part par l'absence de résidu acide en position P7 (Thr38). Si l'on compare avec le site de fixation du cofacteur de la 7 α HSD NAD-dépendante, il est évident que le 2'-phosphate ne pourrait pas se loger à la même place dans cette dernière enzyme à cause de l'acide aspartique 42 (position P7) (figure D5).

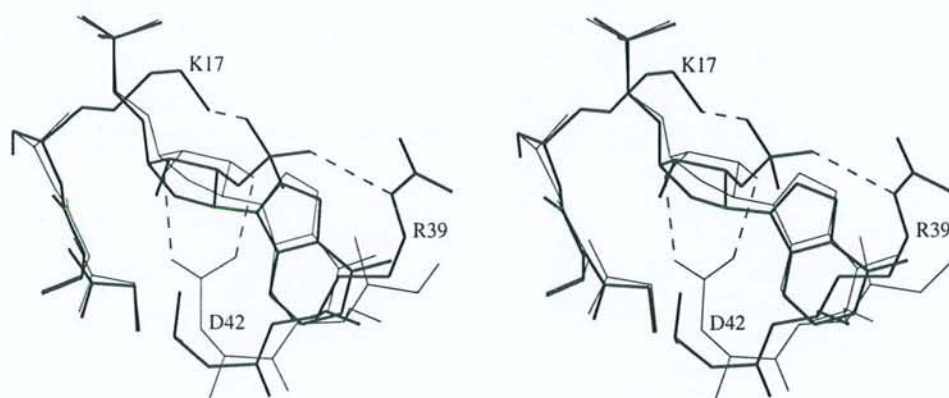


Figure D5 : Superposition du site de fixation de la partie adénosine du cofacteur de la 7 α HSD (traits fins) avec celui de la MLCR (traits épais).

DIIIc3 La famille iii (GXXGXXG)

Toutes les structures des holoenzymes de cette famille révèlent l'interaction entre le groupe NH de l'acide aminé précédant le dernier résidu de la signature et un oxygène du 5'-phosphate de la partie NMN du coenzyme. D'autre part, tous les cofacteurs complexés dans les enzymes NAD-dépendantes se trouvent dans la position référence.

Comme nous l'avons montré dans le paragraphe 'comparison with natural NADP-binding sites' de la publication, le 2'-phosphate du NADP lié dans la QOR pointe vers la boucle β B- α C formant une interaction avec la chaîne principale (figure 7 de la publication). L'enfouissement du ribose de l'adénosine est maximal dans cette protéine (99% ; ce calcul a été réalisé en suivant la méthode décrite par Moodie *et al.* (1996)). Pour adopter cette disposition, les angles dièdres du pyrophosphate de la partie adénosine 5'-diphosphate 2'-phosphate sont

différents de ceux observés dans les structures des enzymes NAD-dépendantes de cette famille. Le 5'-phosphate de la partie AMP 2'-phosphate se trouve alors écarté de la position référence et son ribose est déplacé d'environ 1 Å vers l'extérieur du site. L'enfouissement est complété par un résidu basique n'appartenant pas au 'Rossmann fold'.

La DHPiR est la première enzyme à dualité de cofacteur cocrystallisée avec NAD et NADP (Scapin *et al.*, 1995 ; Reddy *et al.*, 1996). Dans les deux cas, l'ensemble du dinucléotide se trouve déplacé d'environ 1 Å vers le site actif par rapport à la position référence. Les effets de première couche sur la dualité de cofacteur ont été décrits dans le paragraphe AIIId. Ils sont principalement liés à la présence de deux résidus consécutifs d'acidités différentes (Glu38 et Arg39). Il est intéressant de noter le faible enfouissement du ribose de l'adénosine du NADP complexé (63%). La dernière remarque concerne l'arginine 16, cinquième résidu de la signature. Ce résidu n'adopte pas la même conformation dans les deux structures de la DHPiR complexée soit avec NAD soit avec avec NADP (figure D6). Le pyrophosphate semble moins bien stabilisé par ce résidu basique dans le complexe NAD que dans le complexe NADP. En effet, le facteur d'agitation thermique de ce résidu basique est plus élevée dans la structure avec NAD.

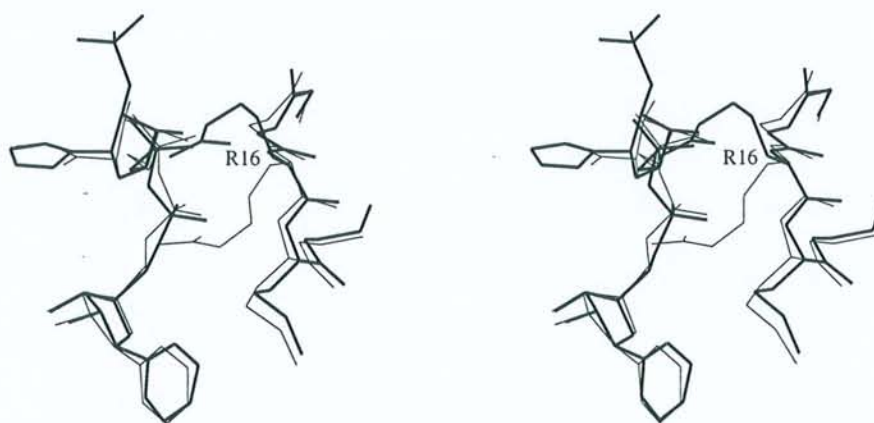


Figure D6 : Superposition du site de fixation du pyrophosphate du complexe DHPiR/NAD (traits fins) avec celui du complexe DHPiR/NADP (traits épais).

L'arginine 16 n'adopte pas la même conformation dans les deux complexes.

DIIIc4 La famille iv (GXGXXA/S)

Les trois molécules de NADP dans les trois structures de flavoenzyme ne sont pas superposables et aucune d'elles n'adopte la position référence. La déviation observée dans la GR semble être liée à la conformation de la signature de cette famille qui est différente de celle observée dans la famille i (cf. AIIb ; DIIIb)(Mittl *et al.*, 1994). Le cas de la TR est similaire à celui de la LipDH de la famille i : le pyrophosphate ne peut se placer dans la position référence à cause de la chaîne latérale du résidu Tyr198. C'est l'arginine 181 située dans boucle β B- α C qui semble responsable du positionnement du pyrophosphate du NADP dans la ThR. Cette arginine est importante dans la fixation du cofacteur puisqu'elle participe à la stabilisation du groupement 2'-phosphate du NADP.

DIIIc5 La famille v (GXXXXXXXXG/A)

Les parties adénosine 5'-diphosphate des molécules de NAD complexées dans les structures de ENR et InhA se trouvent dans la position référence. Les deux pyrophosphates interagissent avec les groupes NH des chaînes principales des résidus situés sur l'avant-dernière position de la signature. Cependant aucune de ces deux enzymes ne présente de résidu acide en P7. Dans les deux cas, les hydroxyles des riboses des adénosines sont stabilisés par des molécules d'eau jouant le rôle de pont avec les protéines.

DIIId Réflexion sur la dualité de cofacteur dans la GAPDH chloroplastique

Avant de proposer des hypothèses sur le mode de fixation du NADP dans la GAPDH chloroplastique, il est important de garder à l'esprit les observations suivantes :

- les GAPDH glycolytiques et chloroplastiques présentent entre elles une forte identité de séquence et possèdent probablement des structures tridimensionnelles très similaires;
- la partie adénosine 5'-diphosphate du NAD se trouve dans la position référence décrite ci-dessus dans les 9 structures de GAPDH glycolytiques déterminées à ce jour.

Après superposition des carbones α de 16 résidus (du brin βA , de la boucle βA - αB et de l'hélice αB) de 25 structures cristallines, l'étude comparative de la position de la partie adénosine 5'-diphosphate (2'-phosphate) du cofacteur dans les structures a montré que dans 15 cas celles-ci adoptent une position commune dans les enzymes. Dans les autres cas, il semble que ce soit principalement la chaîne latérale d'un résidu de la signature qui empêche le cofacteur d'adopter la position référence par contrainte stérique. En outre, cette étude n'a pas permis de mettre en évidence une relation directe entre l'appartenance d'une enzyme à l'une des cinq familles structurales et sa spécificité de cofacteur. Il est intéressant de noter que dans une seule structure d'enzyme NADP-dépendante (MLCR) la partie adénosine 5'-diphosphate 2'-phosphate adopte la position référence. Aucune des sept enzymes NADP-dépendantes ne possède de résidu acide en P7 qui déstabiliserait probablement le 2'-phosphate s'il était présent. Par exemple, la distance entre le 2'-phosphate et l'atome C γ de l'asparagine 32 de la 6PGDH (position P7) vaut 4.13 Å. Cette distance serait trop courte si l'asparagine était remplacée par un acide aspartique. Dans le cas de la DHPiR (enzyme à dualité de cofacteur) la distance entre le carboxylate de l'acide glutamique 38 (position P7) et le 2'-phosphate (5.67 Å) est suffisante pour atténuer les effets de répulsion électrostatique.

Lorsque la DHPiR est complexée avec NAD, la structure montre que le ribose de l'adénosine et la boucle βB - αC sont rapprochés, permettant la formation des deux interactions caractéristiques entre le carboxylate de l'acide glutamique 38 et les hydroxyles du ribose de l'adénosine. De plus, la chaîne latérale du résidu Arg16 (cinquième position dans la signature GXXGXXG) ne possédant pas la même conformation dans les deux complexes NAD/NADP, joue peut-être un rôle dans la reconnaissance du cofacteur (figure D6). Dans toutes les GAPDH la position équivalente est également occupée par une arginine (quatrième position de la signature GXGXXG, soit dans la GAPDH de *B. stearothermophilus* le résidu Arg10). Cependant, l'analyse des neuf structures de GAPDH a montré que ce résidu est impliqué dans la stabilisation des structures secondaires de la protéine (cf. § AIVb6). De plus les trois résidus impliqués dans ce cluster (Arg10, Asp47, Glu314) sont conservées dans les deux types de

GAPDH. L'implication de cette arginine dans la reconnaissance du cofacteur dans la GAPDH chloroplastique est donc peu probable.

Les mutations L33T, T34G et D35G dans le mutant B-S constituent un remodelage de la boucle β B- α C (côté droit de l'adénine, voir figure A4). Les structures de ce mutant complexé avec NAD ou NADP ont mis en évidence un changement conformationnel au niveau de la boucle β B- α C par rapport à l'holoenzyme de type sauvage (figure 3. de la publication). Cette modification structurale n'induit néanmoins pas de déplacement au niveau du résidu Asp32 (position P7) ce qui a pour effet de déstabiliser la fixation 2'-phosphate. En supposant que la conformation de la boucle β B- α C adoptée par la GAPDH chloroplastique soit celle observée dans les deux structures du mutant B-S, il est raisonnable de penser que la partie adénosine 5'-diphosphate 2'-phosphate du NADP ne se trouve pas dans la position référence dans la structure de la GAPDH chloroplastique. L'adénine est stabilisée dans les GAPDH glycolytiques par deux résidus hydrophobes situés dans les boucles β B- α C (Leu33 dans la GAPDH de *B. stearothermophilus*) et β D- α E (Thr96). Le résidu Thr96 est conservé dans toutes les GAPDH, il est donc probable que l'interaction adénine/thréonine subsiste dans les GAPDH chloroplastiques. Une position différente de la boucle β D- α E dans la GAPDH chloroplastique induirait alors une position différente de l'adénine libérant ainsi de l'espace du côté droit du cycle adénine (figure A4) et permettant à l'arginine 77 conservée dans les GAPDH chloroplastiques de stabiliser d'une part le 2'-phosphate et d'autre part le cycle adénine. Néanmoins, la comparaison des séquences des GAPDH chloroplastiques et glycolytiques ne permet pas de révéler de différences évidentes dans la boucle β D- α E et formuler d'hypothèse sur l'éventuel changement conformationnel de cette boucle entre les deux types d'enzymes.

CONCLUSION

Notre objectif était d'apporter de nouvelles informations structurales sur les facteurs moléculaires responsables de la spécificité de cofacteur dans les oxydoréductases. Dans cette optique, nous avons entrepris l'étude cristallographique de deux mutants de la GAPDH de *Bacillus stearothermophilus* :

- le triple mutant D32G, L187A et P188S (D32G-S) qui présente une inversion de sélectivité de NAD vers NADP par rapport à l'enzyme de type sauvage ;

- le mutant L33T, T34G, D35G, L187A et P188S (B-S) construit pour mimer le site du coenzyme des GAPDH à dualité de cofacteur, en l'occurrence les GAPDH chloroplastiques.

Les études cristallographiques déjà réalisées sur la GAPDH de *B. stearothermophilus* (formes holo et apo, double mutant S) nous ont permis de réaliser de nombreuses comparaisons destinées à établir les relations structure-affinité de chaque mutant pour le NAD ou le NADP.

- A l'état cristallin, les quatre complexes étudiés et le complexe **Sp** (double mutant L187A, P188S complexé avec NADP) présentent des structures plus proches de la forme holo que de la forme apo de l'enzyme de type sauvage. De plus, les structures des complexes D32G-**Sn** (mutant D32G-S complexé avec NAD), B-**Sp** (mutant B-S complexé avec NADP) et **Sp** se trouvent dans une conformation plus ouverte que les structures des complexes D32G-**Sp** (mutant D32G-S complexé avec NADP) et B-**Sn** (mutant B-S complexé avec NAD). Ces deux dernières semblent alors présenter un caractère plus 'holo' que les trois autres. Ces observations concordent avec les bandes de Racker de ces cinq complexes.

- Les résultats structuraux sur le double, le triple et le quintuple mutant complexés avec NADP apparaissent en parfait accord avec ceux obtenus en enzymologie et par RMN du phosphore 31. Dans le cas du complexe D32G-**Sn**, l'élimination du résidu Asp 32 favorise l'état dianionique du 2'-phosphate qui se stabilise en formant de nombreuses interactions en particulier avec la sérine 188. Ce processus rend compte de la bonne affinité du triple mutant pour NADP. Dans les deux autres mutants, l'acide aspartique 32 se comporte en tant qu'antidéterminant de type électrostatique pour la fixation du NADP. En effet, le

2'-phosphate pointe vers le solvant et ne forme aucune interaction avec la protéine dans les structures des complexes **Sp** et **B-Sp**.

- La comparaison des structures **D32G-Sn** et **B-Sn** avec celle de l'holoenzyme de type sauvage a permis d'expliquer la baisse d'affinité de ces mutants pour NAD^+ . La plus forte baisse d'affinité observée dans le triple mutant **D32G-S** s'explique principalement par le manque d'interaction directe entre la protéine et le ribose de l'adénosine du NAD. L'augmentation notable de la constante de dissociation du NAD dans le quintuple mutant **B-S** semble être liée à deux effets : le carboxylate de l'acide aspartique 32, en formant une interaction supplémentaire dans la structure du complexe **B-Sn**, atténue probablement ses contacts avec les hydroxyles du ribose de l'adénosine ; le cycle adénine pris en tenaille par les résidus Thr96 et Leu33 dans l'holoenzyme de type sauvage apparaît moins bien stabilisé par la Thr33 dans le mutant **B-S**.

Une comparaison des structures tridimensionnelles des oxydoréductases complexées avec NADP connues à ce jour a montré que la reconnaissance du cofacteur était principalement liée à la présence d'une arginine. Celle-ci est impliquée dans la stabilisation du groupement 2'-phosphate ainsi que dans la stabilisation du cycle adénine par contact π - π . Cependant, des études récentes alliant mutagenèse dirigée et enzymologie, ont permis de mettre en évidence l'importance d'un résidu à chaîne hydroxylée dans la fixation du cofacteur NADP. Il n'est donc pas étonnant de trouver dans la structure du complexe **D32G-Sp** la sérine 188 comme seul résidu en interaction avec le 2'-phosphate. Néanmoins, même si ce mutant présente une bonne affinité pour NADP son activité est diminuée d'un facteur 5 par rapport à celle observée dans l'holoenzyme de type sauvage active avec NAD. Cette baisse d'activité est probablement liée à des effets à longue distance impliquant des résidus difficiles à mettre en évidence à l'intérieur de complexes simplement binaires (enzyme-cofacteur).

Les structures du mutant **B-S** complexé soit avec NAD soit avec NADP ont révélé un changement conformationnel de la boucle $\beta\text{B}-\alpha\text{C}$ qui a pour principal effet d'impliquer le résidu

Asp32 dans une interaction supplémentaire avec la chaîne principale de cette boucle. Cette interaction ne suffit cependant pas à diminuer la répulsion électrostatique entre la chaîne latérale de ce résidu acide et le 2'-phosphate dans le complexe B-Sp. Les cinq mutations L33T, T34G, D35G, L187A et P188S semblent donc insuffisantes pour mimer la fixation du NADP dans la GAPDH chloroplastique. Un alignement de structures primaires limité aux GAPDH, suggère à ce sujet l'intérêt potentiel d'investigations supplémentaires au niveau des positions 180 et 183 situées dans la S-loop.

Une étude comparative complémentaire du site de fixation de la partie adénosine 5'-diphosphate (2'-phosphate) de 25 enzymes de type 'Rossmann fold' cocrystallisées soit avec NAD soit avec NADP, a permis d'apporter de nouvelles informations. Dans la majorité des structures d'enzymes NAD dépendantes, les parties adénosine 5'-diphosphate sont superposables. Par contre, aucune position commune pour la partie adénosine 5'-diphosphate 2'-phosphate n'a pu être mise en évidence dans les structures des enzymes complexées avec NADP. Ces observations nous ont conduit à supposer que la molécule de NADP dans la GAPDH chloroplastique ne se trouve pas dans la position commune adoptée en particulier par le NAD dans la GAPDH de *B. stearothermophilus*. En l'absence de déplacement important de la boucle β B- α C, tel que le suggère la structure du complexe B-Sp, une position différente de la boucle β D- α E dans la GAPDH chloroplastique pourrait éventuellement écarter l'adénosine de la position observée dans le mutant et ainsi atténuer la répulsion électrostatique avec Asp 32.

ANNEXE 1

Bactéries		1	10	20	30	40	50	60
<i>Alcaligenes eutrophus</i> (chromosomique)		MTI	KVA	NGF	G	H	G	H
<i>Alcaligenes eutrophus</i> (plasmidique)		MTI	KVA	NGF	G	H	G	H
<i>Anabaena variabilis</i> (gène 1)	M L	KTI	KVG	NGF	G	H	G	H
<i>Anabaena variabilis</i> (gène 2)		MI	RVA	NGF	G	H	G	H
<i>Anabaena variabilis</i> (gène 3)		MKI	RVG	NGF	G	H	G	H
<i>Bacillus megaterium</i>		AV	KVG	NGF	G	H	G	H
<i>Bacillus stearothermophilus</i>		AV	KVG	NGF	G	H	G	H
<i>Bacillus subtilis</i>		AV	KVG	NGF	G	H	G	H
<i>Buchnera aphidicola</i>		MTI	KI	G	NGF	G	H	G
<i>Citrobacter freundii</i>	#							
<i>Corynebacterium glutamicum</i>		MTI	RVG	NGF	G	H	G	H
<i>Enterobacter aerogenes</i>	#							
<i>Escherichia blattae</i>	#							
<i>Escherichia coli</i> (gène A)		T	KVG	NGF	G	H	G	H
<i>Escherichia coli</i> (gène B)		T	RVVA	NGF	G	H	G	H
<i>Escherichia coli</i> (gène C)		MS	KVG	NGF	G	H	G	H
<i>Escherichia fergusonii</i>	#							
<i>Escherichia hemannii</i>	#							
<i>Escherichia vulneris</i>	#							
<i>Klebsiella pneumoniae</i>	#							
<i>Lactococcus lactis</i>		MV	V	KVG	NGF	G	H	G
<i>Pseudomonas aeruginosa</i>	#	MTI	RLA	NGF	G	H	G	H
<i>Rhodobacter sphaeroides</i>	#	MTI	RVVA	NGF	G	H	G	H
<i>Salmonella typhimurium</i>	#							
<i>Serratia marcescens</i>	#							
<i>Serratia odorifera</i>	#							
<i>Streptococcus sp.</i>		MV	V	KVG	NGF	G	H	G
<i>Thermotoga maritima</i>		A	RVVA	NGF	G	H	G	H
<i>Thermus aquaticus</i>		K	V	G	NGF	G	H	G
<i>Xanthobacter flavus</i>		MS	V	KVA	NGF	G	H	G
<i>Zygomonas mobilis</i>		A	V	KVA	NGF	G	H	G
Champignons								
<i>Aspergillus nidulans</i>		A	P	K	V	G	NGF	G
<i>Agaricus bisporus</i> (gène 1)		MV	N	V	G	NGF	G	H
<i>Agaricus bisporus</i> (gène 2)		MV	N	V	G	NGF	G	H
<i>Cephalosporium acremonium</i>		MA	I	K	V	G	NGF	G
<i>Claviceps purpurea</i>		MA	V	K	V	G	NGF	G
<i>Cochliobolus heterostrophus</i>		MV	V	K	V	G	NGF	G
<i>Colletotrichum gloeosporioides</i>	MA	P	I	K	V	G	NGF	G
<i>Cryphonectria parasitica</i>		MV	V	K	V	G	NGF	G
<i>Curvularia lunata</i>		MV	V	K	V	G	NGF	G
<i>Kluyveromyces fragilis</i>		V	K	V	G	NGF	G	H
<i>Kluyveromyces fragilis</i>		V	K	V	G	NGF	G	H
<i>Phanerochaete chrysosporium</i>	MP	V	K	V	G	NGF	G	H
<i>Phytophthora infestans</i>		M	P	V	K	V	G	NGF
<i>Podospora anserina</i>		MT	V	K	V	G	NGF	G
<i>Saccharomyces cerevisiae</i>		V	R	V	A	NGF	G	H
<i>Schizophyllum commune</i>		MA	V	K	V	G	NGF	G
<i>Streptomyces aureofaciens</i>		MT	V	K	V	G	NGF	G
<i>Streptomyces aureofaciens</i>		MT	V	K	V	G	NGF	G
<i>Trichoderma koningii</i> (gène 1)		V	P	K	V	G	NGF	G
<i>Trichoderma koningii</i> (gène 2)		A	P	I	K	V	G	NGF
<i>Ustilago maydis</i> (huilacoché)		S	Q	V	N	I	G	NGF
<i>Zygosaccharomyces rouxi</i>		V	N	V	S	V	NGF	G
Protozoaires								
<i>Entamoeba histolytica</i>	#							
<i>Giardia intestinalis</i>	#	M	P	I	R	L	G	NGF
<i>Giardia lamblia</i>	#							
<i>Leishmania mexicana</i> (cytosolique)		MV	K	V	G	NGF	G	H
<i>Leishmania mexicana</i> (glycosomal)		A	P	I	K	V	G	NGF
<i>Trichomonas vaginalis</i>		V	V	K	V	G	NGF	G
<i>Trypanoplasma borreli</i>	MA	P	I	K	V	G	NGF	G
<i>Trypanosoma brucei</i> (cytosolique)		V	I	R	V	A	NGF	G
<i>Trypanosoma brucei</i> (glycosomal)		MT	I	K	V	G	NGF	G
<i>Trypanosoma cruzi</i> (glycosomal)		M	P	I	K	V	G	NGF
Invertébrés								
<i>Caenorhabditis briggsae</i>		M	S	K	P	T	V	G
<i>Caenorhabditis elegans</i>		M	S	K	A	N	V	G
<i>Drosophila hydei</i>		M	S	K	I	G	NGF	G
<i>Drosophila melanogaster</i> (mouche de vinaigre)		M	S	K	I	G	NGF	G
<i>Homarus americanus</i> (homard)		M	S	K	A	N	V	G
<i>Schistosoma japonicum</i>		M	S	K	A	N	V	G
<i>Schistosoma mansoni</i>		M	S	K	A	N	V	G
Vertébrés								
<i>Coturnix coturnix japonica</i> (caille)		V	K	V	G	NGF	G	H
<i>Cratichneumon gressus</i> (humster)		MV	K	V	G	NGF	G	H
<i>Gallus gallus</i> (poulet)		MV	K	V	G	NGF	G	H
<i>Homo sapiens</i> (humain) (muscle)		G	K	V	K	V	G	NGF
<i>Homo sapiens</i> (humain) (foie)		G	K	V	K	V	G	NGF
<i>Mus musculus</i> (souris)		MV	K	V	G	NGF	G	H
<i>Oryctolagus cuniculus</i> (lapin)		V	K	V	G	NGF	G	H
<i>Rattus norvegicus</i> (rat)		MV	K	V	G	NGF	G	H
<i>Sus scrofa domestica</i> (porc)		V	K	V	G	NGF	G	H
Plantes (cytosoliques)								
<i>Antirrhinum majus</i>		M	A	P	I	K	V	G
<i>Arabidopsis thaliana</i>	MA	D	K	K	I	K	V	G
<i>Atriplex nummularia</i>		M	A	K	V	K	V	G
<i>Chlamydomonas reinhardtii</i>	MS	G	K	K	I	K	V	G
<i>Chondrus crispus</i>		M	T	A	P	K	V	G
<i>Croton tiglium</i> (plantaigrain)		M	A	K	V	K	V	G
<i>Dianthus carophyllus</i> (aillette)		M	A	P	I	K	V	G
<i>Gracilaria vermiculosa</i>		M	T	V	P	K	V	G
<i>Hordeum vulgare</i> (orge) (gène 1)	#	M	G	K	I	K	V	G
<i>Hordeum vulgare</i> (orge) (gène 2)		M	G	K	I	K	V	G
<i>Magnolia liliiflora</i> (magnolia)		M	A	K	V	K	V	G
<i>Neembryanthemum crystallinum</i>								
<i>Nicotiana tabacum</i> (tabac)	#							
<i>Petroselinum crispum</i> (persil)		M	K	M	K	I	K	V
<i>Petunia hybrida</i> (petunia)		M	A	K	I	K	V	G
<i>Physcomitrella patens</i>		M	G	S	A	K	I	K
<i>Pinus sylvestris</i> (pin)	MG	S	T	G	K	I	K	V
<i>Pinus sativum</i> (petit pois)		M	G	S	A	K	I	K
<i>Ranunculus acer</i> (bouton d'or)		M	T	G	K	I	K	V
<i>Sinapis alba</i> (moutarde)		A	D	K	I	K	V	G
<i>Zea mays</i> (maïs)		M	G	K	I	K	V	G
Plantes (chloroplastiques)								
<i>Arabidopsis thaliana</i> (gène A)		A	K	L	K	V	A	NGF
<i>Arabidopsis thaliana</i> (gène B)		A	K	L	K	V	A	NGF
<i>Chlamydomonas reinhardtii</i>		A	K	L	K	V	A	NGF
<i>Chondrus crispus</i>		A	K	L	K	V	A	NGF
<i>Gracilaria vermiculosa</i>		A	K	L	K	V	A	NGF
<i>Pinus sativum</i> (petit pois) (gène A)		K	L	K	V	A	NGF	G
<i>Pinus sativum</i> (petit pois) (gène B)		K	L	K	V	A	NGF	G
<i>Nicotiana tabacum</i> (tabac) (gène A)		A	K	L	K	V	A	NGF
<i>Nicotiana tabacum</i> (tabac) (gène B)		A	K	L	K	V	A	NGF
<i>Sinapis alba</i> (moutarde)	#							
<i>Spinacia oleracea</i> (épinard) (gène A)		K	L	K	V	A	NGF	G
<i>Spinacia oleracea</i> (épinard) (gène B)		K	L	K	V	A	NGF	G
<i>Zea mays</i> (maïs) (gène A)		A	K	L	K	V	A	NGF
<i>Zea mays</i> (maïs) (pseudogène)		A	K	L	K	V	A	NGF

β A

α B

β B

α C

		70	80	90	100	110	120
Bactéries							
<i>Alcaligenes eutrophus</i> (chromosomique)	G R F F P G	V V V V D D D D	A F R V N G	D R R V R R R R	Q Q Q Q Q Q Q Q		
<i>Alcaligenes eutrophus</i> (plasmidique)	G R F F P G	V V V V D D D D	A F R V N G	D R R V R R R R	Q Q Q Q Q Q Q Q		
<i>Anabaena variabilis</i> (gène 1)	G K V L K K V	V V V V A A A A	A G S R V L V	K H T K T	L V S S E A E E	P A E E E E E E	
<i>Anabaena variabilis</i> (gène 2)	G K V L K K V	V V V V A A A A	A G S R V L V	K H T K T	L V S S E A E E	P A E E E E E E	
<i>Anabaena variabilis</i> (gène 3)	G K V L K K V	V V V V A A A A	A G S R V L V	K H T K T	L V S S E A E E	P A E E E E E E	
<i>Bacillus megaterium</i>	G K V L K K V	V V V V A A A A	A G S R V L V	K H T K T	L V S S E A E E	P A E E E E E E	
<i>Bacillus stearothermophilus</i>	G K V L K K V	V V V V A A A A	A G S R V L V	K H T K T	L V S S E A E E	P A E E E E E E	
<i>Bacillus subtilis</i>	G K V L K K V	V V V V A A A A	A G S R V L V	K H T K T	L V S S E A E E	P A E E E E E E	
<i>Buchnera aphidicola</i>	G K V L K K V	V V V V A A A A	A G S R V L V	K H T K T	L V S S E A E E	P A E E E E E E	
<i>Citrobacter freundii</i>	G K V L K K V	V V V V A A A A	A G S R V L V	K H T K T	L V S S E A E E	P A E E E E E E	
<i>Corynebacterium glutamicum</i>	G K V L K K V	V V V V A A A A	A G S R V L V	K H T K T	L V S S E A E E	P A E E E E E E	
<i>Enterobacter aerogenes</i>	G K V L K K V	V V V V A A A A	A G S R V L V	K H T K T	L V S S E A E E	P A E E E E E E	
<i>Escherichia blattae</i>	G K V L K K V	V V V V A A A A	A G S R V L V	K H T K T	L V S S E A E E	P A E E E E E E	
<i>Escherichia coli</i> (gap A)	G K V L K K V	V V V V A A A A	A G S R V L V	K H T K T	L V S S E A E E	P A E E E E E E	
<i>Escherichia coli</i> (gap B)	G K V L K K V	V V V V A A A A	A G S R V L V	K H T K T	L V S S E A E E	P A E E E E E E	
<i>Escherichia coli</i> (gap C)	G K V L K K V	V V V V A A A A	A G S R V L V	K H T K T	L V S S E A E E	P A E E E E E E	
<i>Escherichia fergusonii</i>	G K V L K K V	V V V V A A A A	A G S R V L V	K H T K T	L V S S E A E E	P A E E E E E E	
<i>Escherichia hermannii</i>	G K V L K K V	V V V V A A A A	A G S R V L V	K H T K T	L V S S E A E E	P A E E E E E E	
<i>Escherichia vulneris</i>	G K V L K K V	V V V V A A A A	A G S R V L V	K H T K T	L V S S E A E E	P A E E E E E E	
<i>Klebsiella pneumoniae</i>	G K V L K K V	V V V V A A A A	A G S R V L V	K H T K T	L V S S E A E E	P A E E E E E E	
<i>Lactococcus lactis</i>	G K V L K K V	V V V V A A A A	A G S R V L V	K H T K T	L V S S E A E E	P A E E E E E E	
<i>Pseudomonas aeruginosa</i>	G K V L K K V	V V V V A A A A	A G S R V L V	K H T K T	L V S S E A E E	P A E E E E E E	
<i>Rhodobacter sphaeroides</i>	G K V L K K V	V V V V A A A A	A G S R V L V	K H T K T	L V S S E A E E	P A E E E E E E	
<i>Salmonella typhimurium</i>	G K V L K K V	V V V V A A A A	A G S R V L V	K H T K T	L V S S E A E E	P A E E E E E E	
<i>Serratia marcescens</i>	G K V L K K V	V V V V A A A A	A G S R V L V	K H T K T	L V S S E A E E	P A E E E E E E	
<i>Serratia odorifera</i>	G K V L K K V	V V V V A A A A	A G S R V L V	K H T K T	L V S S E A E E	P A E E E E E E	
<i>Streptococcus sp.</i>	G K V L K K V	V V V V A A A A	A G S R V L V	K H T K T	L V S S E A E E	P A E E E E E E	
<i>Thermotoga maritima</i>	G K V L K K V	V V V V A A A A	A G S R V L V	K H T K T	L V S S E A E E	P A E E E E E E	
<i>Thermus aquaticus</i>	G K V L K K V	V V V V A A A A	A G S R V L V	K H T K T	L V S S E A E E	P A E E E E E E	
<i>Xanthobacter flavus</i>	G K V L K K V	V V V V A A A A	A G S R V L V	K H T K T	L V S S E A E E	P A E E E E E E	
<i>Zymomonas mobilis</i>	G K V L K K V	V V V V A A A A	A G S R V L V	K H T K T	L V S S E A E E	P A E E E E E E	
Champignons							
<i>Aspergillus nidulans</i>	G Q F F K G	V V V V T Y D D E	G L I V N G	K K K R A K K K	R R R R R R R R	P A E E E E E E	
<i>Agaricus bisporus</i> (gène 1)	G Q F F K G	V V V V T Y D D E	G L I V N G	K K K R A K K K	R R R R R R R R	P A E E E E E E	
<i>Agaricus bisporus</i> (gène 2)	G Q F F K G	V V V V T Y D D E	G L I V N G	K K K R A K K K	R R R R R R R R	P A E E E E E E	
<i>Cephalosporium acremonium</i>	G Q F F K G	V V V V T Y D D E	G L I V N G	K K K R A K K K	R R R R R R R R	P A E E E E E E	
<i>Claviceps purpurea</i>	G Q F F K G	V V V V T Y D D E	G L I V N G	K K K R A K K K	R R R R R R R R	P A E E E E E E	
<i>Cochliobolus heterostrophus</i>	G Q F F K G	V V V V T Y D D E	G L I V N G	K K K R A K K K	R R R R R R R R	P A E E E E E E	
<i>Colletotrichum gloeosporioides</i>	G Q F F K G	V V V V T Y D D E	G L I V N G	K K K R A K K K	R R R R R R R R	P A E E E E E E	
<i>Cryphonectria parasitica</i>	G Q F F K G	V V V V T Y D D E	G L I V N G	K K K R A K K K	R R R R R R R R	P A E E E E E E	
<i>Curvularia lunata</i>	G Q F F K G	V V V V T Y D D E	G L I V N G	K K K R A K K K	R R R R R R R R	P A E E E E E E	
<i>Kluyveromyces fragilis</i>	G Q F F K G	V V V V T Y D D E	G L I V N G	K K K R A K K K	R R R R R R R R	P A E E E E E E	
<i>Kluyveromyces fragilis</i>	G Q F F K G	V V V V T Y D D E	G L I V N G	K K K R A K K K	R R R R R R R R	P A E E E E E E	

	100	130	140	110	150	120	160	130	170	180
Bactéries										
<i>Alcaligenes eutrophus</i> (chromosomique)	GL	FT	-	-	-	-	-	-	-	-
<i>Alcaligenes eutrophus</i> (plasmidique)	GL	FT	-	-	-	-	-	-	-	-
<i>Anabaena variabilis</i> (gène 1)	GL	FT	-	-	-	-	-	-	-	-
<i>Anabaena variabilis</i> (gène 2)	GV	FT	-	-	-	-	-	-	-	-
<i>Anabaena variabilis</i> (gène 3)	GV	FR	-	-	-	-	-	-	-	-
<i>Bacillus megaterium</i>	GF	FT	-	-	-	-	-	-	-	-
<i>Bacillus stearothermophilus</i>	GR	FT	-	-	-	-	-	-	-	-
<i>Bacillus subtilis</i>	GF	FT	-	-	-	-	-	-	-	-
<i>Buchnera aphidicola</i>	GL	FL	-	-	-	-	-	-	-	-
<i>Citrobacter freundii</i>	GL	FL	-	-	-	-	-	-	-	-
<i>Corynebacterium glutamicum</i>	GF	FT	-	-	-	-	-	-	-	-
<i>Enterobacter aerogenes</i>	GL	FL	-	-	-	-	-	-	-	-
<i>Escherichia blattae</i>	GL	FL	-	-	-	-	-	-	-	-
<i>Escherichia coli</i> (gène A)	GL	FL	-	-	-	-	-	-	-	-
<i>Escherichia coli</i> (gène B)	GV	Y G	-	-	-	-	-	-	-	-
<i>Escherichia coli</i> (gène C)	GF	Y T	-	-	-	-	-	-	-	-
<i>Escherichia fergusonii</i>	GL	FL	-	-	-	-	-	-	-	-
<i>Escherichia hermannii</i>	GL	FL	-	-	-	-	-	-	-	-
<i>Escherichia vulneris</i>	GL	FL	-	-	-	-	-	-	-	-
<i>Klebsiella pneumoniae</i>	GL	FL	-	-	-	-	-	-	-	-
<i>Lactococcus lactis</i>	GF	FA	-	-	-	-	-	-	-	-
<i>Pseudomonas aeruginosa</i>	GL	FT	-	-	-	-	-	-	-	-
<i>Rhodobacter sphaeroides</i>	GL	FT	-	-	-	-	-	-	-	-
<i>Salmonella typhimurium</i>	VI	FL	-	-	-	-	-	-	-	-
<i>Serratia marcescens</i>	GL	FL	-	-	-	-	-	-	-	-
<i>Serratia odorifera</i>	GL	FL	-	-	-	-	-	-	-	-
<i>Streptococcus sp.</i>	GF	FA	-	-	-	-	-	-	-	-
<i>Thermotoga maritima</i>	GV	FR	-	-	-	-	-	-	-	-
<i>Thermus aquaticus</i>	GV	FT	-	-	-	-	-	-	-	-
<i>Xanthobacter flavus</i>	GI	FT	-	-	-	-	-	-	-	-
<i>Zymomonas mobilis</i>	GI	FT	-	-	-	-	-	-	-	-
Champignons										
<i>Aspergillus nidulans</i>	GV	FT	-	-	-	-	-	-	-	-
<i>Agaricus bisporus</i> (gène 1)	GV	FK	-	-	-	-	-	-	-	-
<i>Agaricus bisporus</i> (gène 2)	GV	FT	-	-	-	-	-	-	-	-
<i>Cephalosporium acremonium</i>	GV	FT	-	-	-	-	-	-	-	-
<i>Claviceps purpurea</i>	GV	FT	-	-	-	-	-	-	-	-
<i>Cochliobolus heterostrophus</i>	GV	FT	-	-	-	-	-	-	-	-
<i>Colletotrichum gloeosporioides</i>	GV	FT	-	-	-	-	-	-	-	-
<i>Cryphonectria parasitica</i>	GV	FT	-	-	-	-	-	-	-	-
<i>Curvularia lunata</i>	GV	FT	-	-	-	-	-	-	-	-
<i>Kluyveromyces fragilis</i>	GV	FK	-	-	-	-	-	-	-	-
<i>Kluyveromyces fragilis</i>	GV	FK	-	-	-	-	-	-	-	-
<i>Phanerochaete chrysosporium</i>	GV	FT	-	-	-	-	-	-	-	-
<i>Phytophthora infestans</i>	GA	FT	-	-	-	-	-	-	-	-
<i>Podospora anserina</i>	GV	FT	-	-	-	-	-	-	-	-
<i>Saccharomyces cerevisiae</i>	GV	FK	-	-	-	-	-	-	-	-
<i>Schizophyllum commune</i>	GV	FT	-	-	-	-	-	-	-	-
<i>Streptomyces arenae</i>	GI	FT	-	-	-	-	-	-	-	-
<i>Streptomyces aurifaciens</i>	GR	FT	-	-	-	-	-	-	-	-
<i>Trichoderma koningii</i> (gène 1)	GV	FT	-	-	-	-	-	-	-	-
<i>Trichoderma koningii</i> (gène 2)	GV	FT	-	-	-	-	-	-	-	-
<i>Ustilago maydis</i> (fruiting body)	GV	FT	-	-	-	-	-	-	-	-
<i>Zygosaccharomyces rouxi</i>	GI	FK	-	-	-	-	-	-	-	-
Protozoaires										
<i>Entamoeba histolytica</i>	GV	FT	-	-	-	-	-	-	-	-
<i>Giardia intestinalis</i>	GR	FT	-	-	-	-	-	-	-	-
<i>Giardia lamblia</i>	GR	FT	-	-	-	-	-	-	-	-
<i>Leishmania mexicana</i> (cytosolique)	GL	FL	-	-	-	-	-	-	-	-
<i>Leishmania mexicana</i> (glycosomal)	GL	FL	-	-	-	-	-	-	-	-
<i>Trichomonas vaginalis</i>	GL	FT	-	-	-	-	-	-	-	-
<i>Trypanosoma brucei</i>	GL	FT	-	-	-	-	-	-	-	-
<i>Trypanosoma brucei</i> (cytosolique)	GL	FT	-	-	-	-	-	-	-	-
<i>Trypanosoma brucei</i> (glycosomal)	GL	FT	-	-	-	-	-	-	-	-
<i>Trypanosoma cruzi</i> (glycosomal)	GL	FT	-	-	-	-	-	-	-	-
Invertébrés										
<i>Caenorhabditis briggsae</i>	GV	FT	-	-	-	-	-	-	-	-
<i>Caenorhabditis elegans</i>	GV	FT	-	-	-	-	-	-	-	-
<i>Drosophila hydei</i>	GV	FT	-	-	-	-	-	-	-	-
<i>Drosophila melanogaster</i> (mouche de vinaigre)	GV	FT	-	-	-	-	-	-	-	-
<i>Homarus americanus</i> (homard)	GV	FT	-	-	-	-	-	-	-	-
<i>Schistosoma japonicum</i>	GV	FT	-	-	-	-	-	-	-	-
<i>Schistosoma mansoni</i>	GV	FT	-	-	-	-	-	-	-	-
Vertébrés										
<i>Coturnix coturnix japonica</i> (caille)	GV	FT	-	-	-	-	-	-	-	-
<i>Criseolus griseus</i> (hamster)	GV	FT	-	-	-	-	-	-	-	-
<i>Gallus gallus</i> (poulet)	GV	FT	-	-	-	-	-	-	-	-
<i>Homo sapiens</i> (humain) (muscle)	GV	FT	-	-	-	-	-	-	-	-
<i>Homo sapiens</i> (humain) (foie)	GV	FT	-	-	-	-	-	-	-	-
<i>Mus musculus</i> (souris)	GV	FT	-	-	-	-	-	-	-	-
<i>Oryctolagus cuniculus</i> (lapin)	GV	FT	-	-	-	-	-	-	-	-
<i>Rattus norvegicus</i> (rat)	GV	FT	-	-	-	-	-	-	-	-
<i>Sus scrofa domestica</i> (porc)	GV	FT	-	-	-	-	-	-	-	-
Plantes (cytosoliques)										
<i>Antirrhinum majus</i>	GV	FT	-	-	-	-	-	-	-	-
<i>Arabidopsis thaliana</i>	GV	FT	-	-	-	-	-	-	-	-
<i>Atriplex nummularia</i>	GV	FT	-	-	-	-	-	-	-	-
<i>Chlamydomonas reinhardtii</i>	GV	FT	-	-	-	-	-	-	-	-
<i>Chondrus crispus</i>	GV	FT	-	-	-	-	-	-	-	-
<i>Cnesterigma plantagineum</i>	GV	FT	-	-	-	-	-	-	-	-
<i>Dianthus caryophyllus</i> (aillet)	GV	FT	-	-	-	-	-	-	-	-
<i>Gracilaria vermiculosa</i>	GV	FT	-	-	-	-	-	-	-	-
<i>Hordeum vulgare</i> (orge) (gène 1)	GV	FT	-	-	-	-	-	-	-	-
<i>Hordeum vulgare</i> (orge) (gène 2)	GV	FT	-	-	-	-	-	-	-	-
<i>Magnolia liliiflora</i> (magnolia)	GV	FT	-	-	-	-	-	-	-	-
<i>Mesembryanthemum crystallinum</i>	GV	FT	-	-	-	-	-	-	-	-
<i>Nicotiana tabacum</i> (tabac)	GV	FT	-	-	-	-	-	-	-	-
<i>Petroselinum crispum</i> (persil)	GV	FT	-	-	-	-	-	-	-	-
<i>Petunia hybrida</i> (petunia)	GV	FT	-	-	-	-	-	-	-	-
<i>Physcomitrella patens</i>	GV	FT	-	-	-	-	-	-	-	-
<i>Pinus sylvestris</i> (pin)	GV	FT	-	-	-	-	-	-	-	-
<i>Pinus sativum</i> (petit pois)	GV	FT	-	-	-	-	-	-	-	-
<i>Ranunculus acer</i> (bouton d'or)	GV	FT	-	-	-	-	-	-	-	-
<i>Sinapis alba</i> (moutarde)	GV	FT	-	-	-	-	-	-	-	-
<i>Zea mays</i> (maïs)	GV	FT	-	-	-	-	-	-	-	-
Plantes (chloroplastiques)										
<i>Arabidopsis thaliana</i> (gène A)	GV	FV	-	-	-	-	-	-	-	-
<i>Arabidopsis thaliana</i> (gène B)	GV	FV	-	-	-	-	-	-	-	-
<i>Chlamydomonas reinhardtii</i>	GV	FV	-	-	-	-	-	-	-	-
<i>Chondrus crispus</i>	GV	FV	-	-	-	-	-	-	-	-
<i>Gracilaria vermiculosa</i>	GV	FV	-	-	-	-	-	-	-	-
<i>Pinus sativum</i> (petit pois) (gène A)	GV	FV	-	-	-	-	-	-	-	-
<i>Pinus sativum</i> (petit pois) (gène B)	GV	FV	-	-	-	-	-	-	-	-
<i>Nicotiana tabacum</i> (tabac) (gène A)	GV	FV	-	-	-	-	-	-	-	-
<i>Nicotiana tabacum</i> (tabac) (gène B)	GV	FV	-	-	-	-	-	-	-	-
<i>Sinapis alba</i> (moutarde)	GV	FV	-	-	-	-	-	-	-	-
<i>Spinacia oleracea</i> (épinard) (gène A)	GV	FV	-	-	-	-	-	-	-	-
<i>Spinacia oleracea</i> (épinard) (gène B)	GV	FV	-	-	-	-	-	-	-	-
<i>Zea mays</i> (maïs) (gène A)	GV	FV	-	-	-	-	-	-	-	-
<i>Zea mays</i> (maïs) (pseudogène)	GI	FI	-	-	-	-	-	-	-	-

		190	200	210	220	230	240
		150	160	170	180	190	
Bactéries							
<i>Alcaligenes eutrophus</i> (chromosomique)	D T V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Alcaligenes eutrophus</i> (plasmidique)	D T V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Anabaena variabilis</i> (gène 1)	D V I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Anabaena variabilis</i> (gène 2)	H H I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Anabaena variabilis</i> (gène 3)	H H L	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Bacillus megaterium</i>	H N V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Bacillus stearothermophilus</i>	H H V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Bacillus subtilis</i>	H D V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Buchnera aphidicola</i>	K I V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Citrobacter freundii</i>	K I V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Corynebacterium glutamicum</i>	H N V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Enterobacter aerogenes</i>	D I V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Escherichia blattae</i>	D I V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Escherichia coli</i> (gap A)	D I V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Escherichia coli</i> (gap B)	H R I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Escherichia coli</i> (gap C)	D T I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Escherichia fergusonii</i>	D I V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Escherichia hermannii</i>	D I V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Escherichia vulneris</i>	D I V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Klebsiella pneumoniae</i>	D I V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Lactococcus lactis</i>	E T V	S A G S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Pseudomonas aeruginosa</i>	D L V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Rhodobacter sphaeroides</i>	D I V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Salmonella typhimurium</i>	D I V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Serratia marcescens</i>	D I V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Serratia odorifera</i>	E I V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Streptococcus sp.</i>	E T V	S A G S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Thermotoga maritima</i>	H T I	S C A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Thermus aquaticus</i>	H H I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Xanthobacter flavus</i>	H L V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Zymomonas mobilis</i>	D K I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
Champignons							
<i>Aspergillus nidulans</i>	I Q V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Agaricus bisporus</i> (gène 1)	E V I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Agaricus bisporus</i> (gène 2)	D T I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Cephalosporium acremonium</i>	A D V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Claviceps purpurea</i>	A D V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Cochliobolus heterostrophus</i>	I E A	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Colletotrichum gloeosporioides</i>	A D V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Cryphonectria parasitica</i>	G M V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Curvularia lunata</i>	I E V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Kluyveromyces fragilis</i>	T I V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Kluyveromyces fragilis</i>	T I V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Phanerochaete chrysosporium</i>	Y K V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Rhizoglyphus microsporus</i>	M H V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Podospira anserina</i>	A A V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Saccharomyces cerevisiae</i>	L K I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Schizosaccharomyces octosporus</i>	Y Q V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Streptomyces coelicolor</i>	H T I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Streptomyces aureofaciens</i>	H T I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Trichoderma koningii</i> (gène 1)	A N V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Trichoderma koningii</i> (gène 2)	A D V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Ustilago maydis</i> (huillacoche)	A Q V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Zygosaccharomyces rouxi</i>	Q K I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
Protozoaires							
<i>Entamoeba histolytica</i>	M N I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Giardia intestinalis</i>	M K V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Giardia lamblia</i>	M K V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Leishmania mexicana</i> (cytosolique)	P I V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Leishmania mexicana</i> (glycosomal)	H H V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Trichomonas vaginalis</i>	K C I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Trypanosoma brucei</i>	H S V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Trypanosoma brucei</i> (cytosolique)	A I V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Trypanosoma brucei</i> (glycosomal)	Q H V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Trypanosoma cruzi</i> (glycosomal)	H H V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
Invertébrés							
<i>Caenorhabditis briggsae</i>	D H I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Caenorhabditis elegans</i>	D H V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Drosophila hydei</i>	M H V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Drosophila melanogaster</i> (mouche de vinaigre)	M K V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Homarus americanus</i> (homard)	M T V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Schistosoma japonicum</i>	M S V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Schistosoma mansoni</i>	M S V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
Vertébrés							
<i>Coturnix coturnix japonica</i> (caille)	L K I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Crisetulus griseus</i> (hamster)	L K I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Gallus gallus</i> (poulet)	L K I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Homo sapiens</i> (humain) (muscle)	L K I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Homo sapiens</i> (humain) (foie)	L K I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Mus musculus</i> (souris)	L K I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Oryctolagus cuniculus</i> (lapin)	L K I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Rattus norvegicus</i> (rat)	L K I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Sus scrofa domestica</i> (porc)	L K I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
Plantes (cytosoliques)							
<i>Antirrhinum majus</i>	L H I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Arabidopsis thaliana</i>	L D I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Atriplex nummularia</i>	L N I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Chlamydomonas reinhardtii</i>	D H I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Chondrus crispus</i>	L D I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Craetostigma plantaginum</i>	L D V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Dianthus caryophyllus</i> (aillet)	L D V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Gracilaria verrucosa</i>	L N V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Hordeum vulgare</i> (orge) (gène 1)	I D I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Hordeum vulgare</i> (orge) (gène 2)	V N I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Magnolia liliiflora</i> (magnolia)	I D I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Mesembryanthemum crystallinum</i>	L N I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Nicotiana tabacum</i> (tabac)	Y D I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Petroselinum crispum</i> (persil)	I D I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Petunia hybrida</i> (petunia)	L T V	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Physcomitrella patens</i>	V E I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Pinus sylvestris</i> (pin)	N D I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Pinum sativum</i> (petit pois)	F D I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Rumex crispus</i> (lardon d'or)	I D I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Sinapis alba</i> (moutarde)	L N I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Zea mays</i> (maïs)	V N I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
Plantes (chloroplastiques)							
<i>Arabidopsis thaliana</i> (gène A)	E P I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Arabidopsis thaliana</i> (gène B)	A N I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Chlamydomonas reinhardtii</i>	Y P I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Chondrus crispus</i>	Y D I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Gracilaria verrucosa</i>	F D I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Pinum sativum</i> (petit pois) (gène A)	D D I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Pinum sativum</i> (petit pois) (gène B)	A D I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Nicotiana tabacum</i> (tabac) (gène A)	E E I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Nicotiana tabacum</i> (tabac) (gène B)	A D I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Sinapis alba</i> (moutarde)	D T I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Spinacia oleracea</i> (épinard) (gène A)	D T I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Spinacia oleracea</i> (épinard) (gène B)	A N I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Zea mays</i> (maïs) (gène A)	E E I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y
<i>Zea mays</i> (maïs) (pseudogène)	E P I	S N A S	C T T N C L	A P A K V	N D N K	F G I	V E G L M T T V H S Y

Bactéries		260	310	270	320	330	340	350	360
<i>Alcaligenes eutrophus</i> (chromosomique)		NG	L	K	G	I	-	-	-
<i>Alcaligenes eutrophus</i> (plasmidique)		NG	L	K	G	I	-	-	-
<i>Anabaena variabilis</i> (gène 1)		CA	A	M	K	G	A	-	-
<i>Anabaena variabilis</i> (gène 2)		CA	A	M	K	G	A	-	-
<i>Anabaena variabilis</i> (gène 3)		CA	A	M	K	G	A	-	-
<i>Bacillus megaterium</i>		NA	A	L	K	G	A	-	-
<i>Bacillus stearothermophilus</i>		NA	A	L	K	G	A	-	-
<i>Bacillus subtilis</i>		NA	A	L	K	G	A	-	-
<i>Buchnera aphidicola</i>		CA	E	V	K	N	A	-	-
<i>Citrobacter freundii</i>		K	K	A	M	K	G	A	-
<i>Corynebacterium glutamicum</i>		N	A	A	L	K	G	A	-
<i>Enterobacter aerogenes</i>		K	K	A	M	K	G	A	-
<i>Escherichia blattae</i>		K	K	A	M	K	G	A	-
<i>Escherichia coli</i> (gap A)		K	K	A	M	K	G	A	-
<i>Escherichia coli</i> (gap B)		N	A	A	L	K	G	A	-
<i>Escherichia coli</i> (gap C)		N	A	A	L	K	G	A	-
<i>Escherichia fergusonii</i>		K	K	A	M	K	G	A	-
<i>Escherichia hermannii</i>		K	K	A	M	K	G	A	-
<i>Escherichia vulneris</i>		K	K	A	M	K	G	A	-
<i>Klebsiella pneumoniae</i>		K	K	A	M	K	G	A	-
<i>Lactococcus lactis</i>		N	E	A	M	K	G	A	-
<i>Pseudomonas aeruginosa</i>		N	A	A	L	K	G	A	-
<i>Rhodospirillum rubrum</i>		N	A	A	L	K	G	A	-
<i>Salmonella typhimurium</i>		K	K	A	M	K	G	A	-
<i>Serratia marcescens</i>		K	K	A	M	K	G	A	-
<i>Serratia odorifera</i>		K	K	A	M	K	G	A	-
<i>Streptococcus sp.</i>		N	A	A	L	K	G	A	-
<i>Thermotoga maritima</i>		N	A	A	L	K	G	A	-
<i>Thermus aquaticus</i>		N	A	A	L	K	G	A	-
<i>Xanthobacter flavus</i>		N	E	A	M	K	G	A	-
<i>Zymomonas mobilis</i>		N	E	A	M	K	G	A	-
Champignons									
<i>Aspergillus nidulans</i>		K	K	A	M	K	G	A	-
<i>Agaricus bisporus</i> (gène 1)		K	K	A	M	K	G	A	-
<i>Agaricus bisporus</i> (gène 2)		K	K	A	M	K	G	A	-
<i>Cephalosporium acremonium</i>		K	K	A	M	K	G	A	-
<i>Claviceps purpurea</i>		K	K	A	M	K	G	A	-
<i>Cochliobolus heterostrophus</i>		K	K	A	M	K	G	A	-
<i>Colletotrichum gloeosporioides</i>		K	K	A	M	K	G	A	-
<i>Cryphonectria parasitica</i>		K	K	A	M	K	G	A	-
<i>Curvularia lunata</i>		K	K	A	M	K	G	A	-
<i>Xylomyces lactis</i>		K	K	A	M	K	G	A	-
<i>Xylomyces marianus</i>		K	K	A	M	K	G	A	-
<i>Phanerochaete chrysosporium</i>		K	K	A	M	K	G	A	-
<i>Phytophthora infestans</i>		K	K	A	M	K	G	A	-
<i>Podospora anserina</i>		K	K	A	M	K	G	A	-
<i>Saccharomyces cerevisiae</i>		K	K	A	M	K	G	A	-
<i>Schizophyllum commune</i>		K	K	A	M	K	G	A	-
<i>Streptomyces arenae</i>		K	K	A	M	K	G	A	-
<i>Streptomyces aurafaciens</i>		K	K	A	M	K	G	A	-
<i>Trichoderma koningii</i> (gène 1)		K	K	A	M	K	G	A	-
<i>Trichoderma koningii</i> (gène 2)		K	K	A	M	K	G	A	-
<i>Ustilago maydis</i> (fruitilacoe)		K	K	A	M	K	G	A	-
<i>Zygosaccharomyces rouxi</i>		K	K	A	M	K	G	A	-
Protozoaires									
<i>Entamoeba histolytica</i>		K	K	A	M	K	G	A	-
<i>Giardia intestinalis</i>		K	K	A	M	K	G	A	-
<i>Giardia lamblia</i>		K	K	A	M	K	G	A	-
<i>Leishmania mexicana</i> (cytosolique)		K	K	A	M	K	G	A	-
<i>Leishmania mexicana</i> (glycosomal)		K	K	A	M	K	G	A	-
<i>Trichomonas vaginalis</i>		K	K	A	M	K	G	A	-
<i>Trypanosoma brucei</i>		K	K	A	M	K	G	A	-
<i>Trypanosoma brucei</i> (cytosolique)		K	K	A	M	K	G	A	-
<i>Trypanosoma brucei</i> (glycosomal)		K	K	A	M	K	G	A	-
<i>Trypanosoma cruzi</i> (glycosomal)		K	K	A	M	K	G	A	-
Invertébrés									
<i>Caenorhabditis briggsae</i>		K	K	A	M	K	G	A	-
<i>Caenorhabditis elegans</i>		K	K	A	M	K	G	A	-
<i>Drosophila hydei</i>		K	K	A	M	K	G	A	-
<i>Drosophila melanogaster</i> (mouche de vinaigre)		K	K	A	M	K	G	A	-
<i>Homarus americanus</i> (homard)		K	K	A	M	K	G	A	-
<i>Schistosoma japonicum</i>		K	K	A	M	K	G	A	-
<i>Schistosoma mansoni</i>		K	K	A	M	K	G	A	-
Vertébrés									
<i>Coturnix coturnix japonica</i> (caille)		K	K	A	M	K	G	A	-
<i>Crisetulus gresius</i> (hamster)		K	K	A	M	K	G	A	-
<i>Gallus gallus</i> (poulet)		K	K	A	M	K	G	A	-
<i>Homo sapiens</i> (humain) (muscle)		K	K	A	M	K	G	A	-
<i>Homo sapiens</i> (humain) (foie)		K	K	A	M	K	G	A	-
<i>Mus musculus</i> (souris)		K	K	A	M	K	G	A	-
<i>Oryctolagus cuniculus</i> (lapin)		K	K	A	M	K	G	A	-
<i>Rattus norvegicus</i> (rat)		K	K	A	M	K	G	A	-
<i>Sus scrofa domestica</i> (porc)		K	K	A	M	K	G	A	-
Plantes (cytosoliques)									
<i>Antirrhinum majus</i>		K	K	A	M	K	G	A	-
<i>Arabidopsis thaliana</i>		K	K	A	M	K	G	A	-
<i>Atriplex nummularia</i>		K	K	A	M	K	G	A	-
<i>Chlamydomonas reinhardtii</i>		K	K	A	M	K	G	A	-
<i>Chondrus crispus</i>		K	K	A	M	K	G	A	-
<i>Crotalaria plantaginea</i>		K	K	A	M	K	G	A	-
<i>Dianthus caryophyllus</i> (œillet)		K	K	A	M	K	G	A	-
<i>Gracilaria verrucosa</i>		K	K	A	M	K	G	A	-
<i>Hordeum vulgare</i> (orge) (gène 1)		K	K	A	M	K	G	A	-
<i>Hordeum vulgare</i> (orge) (gène 2)		K	K	A	M	K	G	A	-
<i>Magnolia liliiflora</i> (magnolia)		K	K	A	M	K	G	A	-
<i>Mesembryanthemum crystallinum</i>		K	K	A	M	K	G	A	-
<i>Nicotiana tabacum</i> (tabac)		K	K	A	M	K	G	A	-
<i>Petroselinum crispum</i> (persil)		K	K	A	M	K	G	A	-
<i>Petunia hybrida</i> (petunia)		K	K	A	M	K	G	A	-
<i>Physcomitrella patens</i>		K	K	A	M	K	G	A	-
<i>Pinus sylvestris</i> (pin)		K	K	A	M	K	G	A	-
<i>Pisum sativum</i> (petit pois)		K	K	A	M	K	G	A	-
<i>Rumex crispus</i> (bouillon d'or)		K	K	A	M	K	G	A	-
<i>Sinapis alba</i> (moutarde)		K	K	A	M	K	G	A	-
<i>Zea mays</i> (maïs)		K	K	A	M	K	G	A	-
Plantes (chloroplastiques)									
<i>Arabidopsis thaliana</i> (gène A)		K	K	A	M	K	G	A	-
<i>Arabidopsis thaliana</i> (gène B)		K	K	A	M	K	G	A	-
<i>Chlamydomonas reinhardtii</i>		K	K	A	M	K	G	A	-
<i>Chondrus crispus</i>		K	K	A	M	K	G	A	-
<i>Gracilaria verrucosa</i>		K	K	A	M	K	G	A	-
<i>Pisum sativum</i> (petit pois) (gène A)		K	K	A	M	K	G	A	-
<i>Pisum sativum</i> (petit pois) (gène B)		K	K	A	M	K	G	A	-
<i>Nicotiana tabacum</i> (tabac) (gène A)		K	K	A	M	K	G	A	-
<i>Nicotiana tabacum</i> (tabac) (gène B)		K	K	A	M	K	G	A	-
<i>Sinapis alba</i> (moutarde)		K	K	A	M	K	G	A	-
<i>Spinacia oleracea</i> (épinard) (gène A)		K	K	A	M	K	G	A	-
<i>Spinacia oleracea</i> (épinard) (gène B)		K	K	A	M	K	G	A	-
<i>Zea mays</i> (maïs) (gène A)		K	K	A	M	K	G	A	-
<i>Zea mays</i> (maïs) (pseudogène)		K	K	A	M	K	G	A	-

	310	370	320	380	330	390																					
Bactéries																											
<i>Alcaligenes eutrophus</i> (chromosomique)	KV	SA	WY	DNE	WGF	SNHMLDITAV	ALAHAR																				
<i>Alcaligenes eutrophus</i> (plasmidique)	KV	SA	WY	DNE	WGF	SNHMLDITAV	ALAHAR																				
<i>Anabaena variabilis</i> (gène 1)	KV	VA	WY	DNE	WGF	SNHMLDITAV	ALAHAR																				
<i>Anabaena variabilis</i> (gène 2)	KV	MA	WY	DNE	WGF	SNHMLDITAV	ALAHAR																				
<i>Anabaena variabilis</i> (gène 3)	KI	LA	WY	DNE	WGF	SNHMLDITAV	ALAHAR																				
<i>Bacillus megaterium</i>	KV	IS	WY	DNE	WGF	SNHMLDITAV	ALAHAR																				
<i>Bacillus stearothermophilus</i>	KV	IS	WY	DNE	WGF	SNHMLDITAV	ALAHAR																				
<i>Bacillus subtilis</i>	KV	IS	WY	DNE	WGF	SNHMLDITAV	ALAHAR																				
<i>Buchnera aphidicola</i>	KI	LS	WY	DNE	ETG	SS	SKVLEDLV	ELVALK																			
<i>Citrobacter freundii</i>	KI	LV																									
<i>Corynebacterium glutamicum</i>	KV	VS	WY	DNE	WGF	TQ	LLRL	TELVAE	QALN																		
<i>Enterobacter aerogenes</i>	KI	LV																									
<i>Escherichia blattae</i>	KI	LV																									
<i>Escherichia coli</i> (gap A)	KI	VS	WY	DNE	ETG	VS	NKVL	DLIAHISK																			
<i>Escherichia coli</i> (gap B)	KI	TL	V	W	Y	DNE	WGF	ANRMLD	STLAMA	VAFR																	
<i>Escherichia coli</i> (gap C)	KI	TL	V	W	Y	DNE	WGF	VTQ	ILHTEK	PAKL																	
<i>Escherichia fergusonii</i>	KI	LV																									
<i>Escherichia hermannii</i>	KI	LV																									
<i>Escherichia vulneris</i>	KI	LV																									
<i>Klebsiella pneumoniae</i>	KI	LV																									
<i>Lactococcus lactis</i>	KI	TA	A	W	Y	DNE	ME	FT	AQL	IR	TE	LY	FA	KI	A	K											
<i>Pseudomonas aeruginosa</i>	KI	LS	WY	DNE	WGF	SNHMLDITAV	AMGRLL																				
<i>Rhodobacter sphaeroides</i>	KI	LV																									
<i>Salmonella typhimurium</i>	KI	LV																									
<i>Serratia marcescens</i>	KI	LV																									
<i>Serratia odorifera</i>	KI	LV																									
<i>Streptococcus sp.</i>	KV	VS	WY	DNE	MS	Y	T	A	Q	L	V	B	T	L	E	Y	F	A	KI	A	K						
<i>Thermotoga maritima</i>	KV	SA	WY	DNE	Y	G	Y	S	A	N	R	V	D	T	L	E	L	V	L	K	M						
<i>Thermus aquaticus</i>	KV	F	A	W	Y	DNE	W	G	F	A	S	N	R	M	D	T	V	A	M	S	K	G	V				
<i>Xanthobacter flavus</i>	R	V	L	S	W	Y	DNE	W	G	F	A	S	N	R	M	D	T	V	A	M	S	K	L	G			
<i>Zymomonas mobilis</i>	R	V	V	A	W	Y	DNE	W	G	F	A	S	N	R	M	D	T	A	A	O	M	A	K	T	L		
Champignons																											
<i>Aspergillus nidulans</i>	KI	LV	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Agaricus bisporus</i> (gène 1)	KI	LV	A	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Agaricus bisporus</i> (gène 2)	KI	LV	A	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Cephalosporium acremonium</i>	KI	LV	A	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Claviceps purpurea</i>	KI	LV	A	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Cochliobolus heterostrophus</i>	KI	LV	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Colletotrichum gloeosporioides</i>	KI	LV	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Cryphonectria parasitica</i>	KI	LV	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Curvularia lunata</i>	KI	LV	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Kluyveromyces lactis</i>	KV	V	A	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Kluyveromyces fragilis</i>	KV	V	A	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Kluyveromyces marxianus</i>	KV	V	A	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Phanerochaete chrysosporium</i>	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Phytophthora infestans</i>	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Podospora anserina</i>	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Saccharomycetes cerevisiae</i>	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Schizophyllum commune</i>	KI	V	A	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Streptomyces arenae</i>	KV	V	A	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Streptomyces aureofaciens</i>	KV	V	A	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Trichoderma koningii</i> (gène 1)	KI	M	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Trichoderma koningii</i> (gène 2)	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Ustilago maydis</i> (fruiting body)	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Zygosaccharomycetes rouxi</i>	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
Protozoaires																											
<i>Entamoeba histolytica</i>	KI	V	S				N	E	F	G	Y	A	N	K	L	V	E	L	A	N	Y	V	G	S	K	G	C
<i>Giardia intestinalis</i>	KI	V	A	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Giardia lamblia</i>	KI	V	A																								
<i>Leishmania mexicana</i> (cytosolique)	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Leishmania mexicana</i> (glycosomal)	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Trichomonas vaginalis</i>	KV	L	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Trypanoplasma borreli</i>	KV	L	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Trypanosoma brucei</i> (cytosolique)	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Trypanosoma brucei</i> (glycosomal)	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Trypanosoma cruzi</i> (glycosomal)	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
Invertébrés																											
<i>Caenorhabditis briggsae</i>	KI	L	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Caenorhabditis elegans</i>	KI	L	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Drosophila hydei</i>	KI	L	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Drosophila melanogaster</i> (mouche de vinaigre)	KI	L	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Homarus americanus</i> (homard)	KV	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Schistosoma japonicum</i>	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Schistosoma mansoni</i>	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
Vertébrés																											
<i>Coturnix coturnix japonica</i> (caille)	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Crisetulus griseus</i> (hamster)	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Gallus gallus</i> (poulet)	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Homo sapiens</i> (humain) (muscle)	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Homo sapiens</i> (humain) (foie)	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Mus musculus</i> (souris)	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Oryctolagus cuniculus</i> (lapin)	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Rattus norvegicus</i> (rat)	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Sus scrofa domestica</i> (porc)	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
Plantes (cytosoliques)																											
<i>Antirrhinum majus</i>	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Arabidopsis thaliana</i>	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Atriplex nummularia</i>	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Chlamydomonas reinhardtii</i>	KI	V	A	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Chondrus crispus</i>	KI	V	A	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Craterostigma plantaginifolium</i>	KI	V	A	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Dianthus caryophyllus</i> (caillet)	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Gracilaria verrucosa</i>	KI	V	A	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Hordeum vulgare</i> (orge) (gène 1)	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Hordeum vulgare</i> (orge) (gène 2)	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Magnolia liliiflora</i> (magnolia)	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Mesembryanthemum crystallinum</i>	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Nicotiana tabacum</i> (tabac)	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Petrolellum crispum</i> (persil)	KV	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Petunia hybrida</i> (petunia)	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Physcomitrella patens</i>	KI	V	A	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Pinus sylvestris</i> (pin)	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Pinus sativum</i> (petit pin)	KI	V	S	W	Y	DNE	W	G	F	A	S	N	R	V	D	T	L	E	L	V	L	K	M				
<i>Ranunculus acer</i> (bouton d'or)	KI	V	S	W	Y	DNE	W	G	F																		

ALIGNEMENT DES STRUCTURES PRIMAIRES DE GAPDH

La numérotation du haut correspond aux positions dans l'alignement et celle du bas (deuxième ligne) est la numérotation de la GAPDH de *B. Stearothermophilus*.

Les éléments de structure secondaire de la GAPDH de *B. Stearothermophilus* sont représentés par les boîtes (gris clair pour les brins de feuillet β et gris foncé pour les hélices α) et leur nomenclature est indiquée en dessous de l'alignement.

Les résidus conservés dans au moins 95% des séquences (108 séquences sur les 114) sont indiqués en gras.

Les séquences marquées par # ne sont pas complètes.

La plupart des séquences ont été extraites des banques de données SwissProt, PIR et EMBL. Leurs codes d'accès sont les suivants :

SwissProt : *Alcaligenes eutrophus* g3pc_alceu (chromosomique) et g3pp_alceu (plasmidique) ; *Anabaena variabilis* g3p1_anava, g3p2_anava, g3p3_anava ; *Bacillus megaterium* g3p_bacme ; *Bacillus stearothermophilus* g3p_bacst ; *Bacillus subtilis* g3p_bacsu ; *Buchnera aphidicola* g3p_bucap ; *Citrobacter freundii* g3p_citfr ; *Corynebacterium glutamicum* g3p_corgl ; *Enterobacter aerogenes* g3p_entae ; *Escherichia blattae* g3p_escbl ; *Escherichia coli* g3p2_ecoli et g3p3_ecoli ; *Escherichia fergusonii* g3p_escfe ; *Escherichia hermanii* g3p_esche ; *Escherichia vulneris* g3p_escvu ; *Klebsiella pneumoniae* g3p_klepn ; *Pseudomonas aeruginosa* g3p_pseae ; *Rhodobacter sphaeroides* g3p_rhosh ; *Salmonella typhimurium* g3p_salty ; *Serratia marcescens* g3p_serma ; *Serratia odorifera* g3p_serod ; *Thermotoga maritima* g3p_thema ; *Thermus aquaticus* g3p_theq ; *Zymomonas mobilis* g3p_zymmo ; *Aspergillus nidulans* g3p_emeni ; *Agaricus bisporus* g3p1_agabi et g3p2_agabi ; *Cochliobolus heterostrophus* g3p_coche ; *Colletotrichum gloeosporoides* g3p_colgl ; *Cryphonectria parasitica* g3p_crypa ; *Curvularia lunata* g3p_curlu ; *Kluyveromyces lactis* g3p_klula ; *Phanerochaete chrysosporium* g3p_phach ; *Phytophthora infestans* g3p_phyin ; *Podospora anserina* g3p_podan ; *Sccharomyces cerevisiae* g3p1_yeast ; *Schizophyllum commune* g3p_schco ; *Trichoderma koningii* g3p1_triko et g3p2_triko ; *Ustilago maydis* g3p_ustma ; *Zygosaccharomyces rouxii* g3p_zygro ; *Leishmania*

mexicana g3pc_leime (cytosolique) ; *Trypanosoma brucei* g3pc_trybb (cytosolique) et g3pg_trycr (glycosomal) ; *Trypanosoma cruzi* g3pg_trycr ; *Caenorhabditis briggsae* g3p_caebr ; *Caenorhabditis elegans* g3p_caeel ; *Drosophila hydel* g3p_drohy ; *Drosophila melanogaster* g3p1_drome ; *Homarus americanus* g3p_homam ; *Schistosoma mansoni* g3p_schma ; *Coturnix coturnix japonica* g3p_cotja ; *Crisetelus griseus* g3p_crigr ; *Gallus gallus* g3p_chick ; *Homo sapiens* g3p1_human (muscle) et g3p2_human (foie) ; *Mus musculus* g3p_mouse ; *Oryctolagus cuniculus* g3p_rabit ; *Rattus norvegicus* g3p_rat ; *Sus scrofa domestica* g3p_pig ; *Antirrhinum majus* g3pc_antma ; *Arabidopsis thaliana* g3p_arath (cytosolique), g3pa_arath et g3pb_arath (chloroplastiques) ; *Atriplex nummularia* g3p_atrnu ; *Chlamydomonas reinhardtii* g3p_chlre (cytosolique) et g3pa_chlre (chloroplastique) ; *Chondrus crispus* g3pc_chocr (cytosolique) et g3pa_chocr (chloroplastique) ; *Dianthus caryophyllus* g3pc_diaca ; *Hordeum vulgare* g3pc_horvu et g3px_horvu ; *Magnolia liliiflora* g3pc_magli ; *Mesembryanthemum crystallinum* g3pc_mescl ; *Nicotiana tabacum* g3pc_tobac (cytosolique), g3pa_tobac et g3pb_tobac (chloroplastiques) ; *Petroselinum crispum* g3pc_petr ; *Petunia hybrida* g3pc_pethy ; *Physcomitrella patens* g3pc_phypa ; *Pinus sylvestris* g3pc_pinsy ; *Pisum sativum* g3pc_pea (cytosolique), g3pa_pea et g3pb_pea (chloroplastiques) ; *Ranunculus acer* g3pc_ranac ; *Sinapsis alba* g3pc_sinal (cytosolique) et g3pa_sinal (chloroplastique) ; *Zea mays* g3pc_maize (cytosolique) et g3pa_maize (chloroplastique) ; *Gracilaria verrucosa* g3pa_grave (chloroplastique) ; *Spinacia oleracea* g3pa_spiol et g3pb_spiol.

PIR : *Streptococcus* sp. a42963 ; *Cephalosporium acremonium* s40610 ; *Leishmania mexicana* a48445 (glycosomal) ; *Trichomonas vaginalis* a49700 ; *Craterostigma plantagineum* s42479 ; *Gracilaria verrucosa* s59579 (cytosolique)

EMBL : *Escherichia coli* ecgap ; *Lactococcus lactis* lllpagap ; *Xanthobacter flavus* xf33064 ; *Kluyveromyces marxianus* km24452 ; *Streptomyces arenae* sa44856 ; *Streptomyces aureofaciens* sa21191 ; *Entamoeba histolytica* ehglyalph ; *Giardia intestinalis* glglyalph ; *Giardia lamblia* gl31911 ; *Trypanoplasma borreli* tpgapg ; *Schistosoma japonicum* sjgl3phde.

Publications : *Cephalosporium acremonium* (Kimura et al., 1991) ; *Zea mays* (pseudogène chloroplastique) (Quigley et al., 1989).

BIBLIOGRAPHIE

Adams, M. J., Ellis, G. H., Gover, S., Naylor, C. E. & Phillips, C. (1994). Crystallographic study of coenzyme, coenzyme analogue and substrate binding in 6-phosphogluconate dehydrogenase: implications for NADP specificity and the enzyme mechanism. *Structure* **2**, 651-668.

Adams, M. J., Ford, G. C., Koekoek, R., Lentz, P. J. J., McPherson, A. J., Rossmann, M. G., Smiley, I. E., Schevitz, R. W. & Wonacott, A. J. (1970). Structure of lactate dehydrogenase at 2.8 Å resolution. *Nature* **227**, 1098-1103.

Al-Karadaghi, S., Cedergren-Zeppezauer, E. S., Hövmoller, S., Petratos, K., Terry, H. & Wilson, K. S. (1994). Refined crystal structure of liver alcohol dehydrogenase-NADH complex at 1.8 Å resolution. *Acta Crystallog. D* **50**, 793-807.

Amelunxen, R. E. & Carr, D. O. (1967). The crystallization and properties of glyceraldehyde-3-phosphate dehydrogenase isolated from rabbit muscle by a simplified procedure. *Biochim. Biophys. Acta* **132**, 256-259.

Amemiya, Y. & Miyahara, J. (1988). Imaging plate illuminates many fields. *Nature* **366**, 89-90.

Arndt, U. W. (1968). The Optimum Strategy in Measuring Structure Factors. *Acta Crystallog.* **B24**, 1355-1357.

Bailey, S., Fairlamb, A. H. & Hunter, W. N. (1994). Structure of trypanothione reductase from *Crithidia fasciculata* at 2.6 Å resolution. Enzyme-NADP interactions at 2.8 Å resolution. *Acta Crystallog.* **B50**, 139-154.

Baker, P. J., Britton, K. L., Engel, P. C., Farrants, G. W., Lilley, K. S., Rice, D. W. & Stillman, T. J. (1992a). Subunit assembly and active site location in the structure of glutamate dehydrogenase. *Proteins* **12**, 75-86.

Baker, P. J., Britton, K. L., Rice, D. W., Rob, A. & Stillman, T. J. (1992b). Structural consequences of sequence patterns in the fingerprint region of the nucleotide binding fold. Implications for nucleotide specificity. *J. Mol. Biol.* **228**, 662-671.

Bell, C. E. & Eisenberg, D. (1996). Crystal structure of diptheria toxin bound to nicotinamide adenine dinucleotide. *Biochemistry* **35**, 1137-1149.

Bernard, N., Johnsen, K., Holbrook, J. J. & Delcour, J. (1995). D175 discriminates between NADH and NADPH in the coenzyme binding site of *Lactobacillus delbrueckii* subsp. *bulgaricus* D-lactate dehydrogenase. *Biochem. Biophys. Res. Commun.* **208**, 895-900.

Berni, R., Mozzarelli, A., Rossi, G. L., Bolognesi, M. & Oberti, R. (1979). Crystallographic symmetry and coenzyme binding properties of D-glyceraldehyde-3-phosphate dehydrogenase from the tail muscle of *Palinurus vulgaris*. *J. Biol. Chem.* **254**, 8004-8006.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr, Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). The Protein Data Bank : A computer-based archival-file for Macromolecular Structures. *J. Mol. Biol.* **112**, 535-542.

Biesecker, G., Harris, J. I., Thierry, J. C., Walker, J. E. & Wonacott, A. J. (1977). Sequence and structure of D-glyceraldehyde-3-phosphate dehydrogenase from *Bacillus stearothermophilus*. *Nature* **266**, 328-333.

Blundell, T., L. & Tickle, I., J. (1985). *Proceedings of the Daresbury Study Weekend*.

- Bocanegra, J. A., Scrutton, N. S. & Perham, R. N. (1993). Creation of an NADP-dependent pyruvate dehydrogenase multienzyme complex by protein engineering. *Biochemistry* **32**, 2737-2740.
- Bolin, J. T., Filman, D. J., Matthews, D. A., Hamlin, R. C. & Kraut, J. (1982). Crystal structures of *Escherichia coli* and *Lactobacillus casei* dihydrofolate reductase refined at 1.7 Å resolution. I. General features and binding of methotrexate. *J. Biol. Chem.* **257**, 13650-13662.
- Bork, P. & Grunwald, C. (1990). Recognition of different nucleotide binding-sites in primary structures using a property-pattern approach. *Eur. J. Biochem.* **191**, 347-358.
- Brünger, A. T., Kuriyan, J. & Karplus, M. (1987). Crystallographic R-factor refinement by molecular dynamics. *Science* **229**, 458-460.
- Brünger, A. T., Karplus, M. & Petsko, G. A. (1989). Crystallographic refinement by simulated annealing : application to crambin. *Acta Crystallog.* **A45**, 50-61.
- Brünger, A. T., Krukowski, A. & Erickson, J. W. (1990). Slow-cooling protocols for crystallographic refinement by simulated annealing. *Acta Crystallog.* **A46**, 585-593.
- Brünger, A. T. (1992a). Free R value : a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472-475.
- Brünger, A. T. (1992b). *X-PLOR Manual, version 3.1*, Yale University, Newhaven.
- Buehner, M., Ford, G. C., Moras, D., Olsen, K. W. & Rossmann, M. G. (1974). Three-dimensional structure of D-glyceraldehyde-3-phosphate dehydrogenase. *J. Mol. Biol.* **90**, 25-49.
- Bystroff, C., Oatley, S. J. & Kraut, J. (1990). Crystal structures of *Escherichia coli* dihydrofolate reductase: The NADP⁺ holoenzyme and the folate NADP⁺ ternary complex. Substrate binding and a model for the transition state. *Biochemistry* **29**, 3263-3277.
- Campbell, J. W., Duée, E., Hodgson, G., Mercer, W. D., Stammers, D. K., Wendell, P. L., Muirhead, H. & Watson, H. C. (1971). *Cold Spring Harbor Symp. Quant. Biol.* **36**, 165-170.
- Cerff, R. (1978). Glyceraldehyde-3-phosphate dehydrogenase (NADP) from *Sinapsis alba* L. NAD(P)-induced conformation changes of the enzyme. *Eur. J. Biochem.* **82**, 45-53.
- Cerff, R. (1979). Quaternary structure of higher plant glyceraldehyde-3-phosphate dehydrogenase. *Eur. J. Biochem.* **94**, 243-247.
- Chakrabarti, P. & Samanta, U. (1995). CH/ π interaction in the packing of the adenine ring in protein structures. *J. Mol. Biol.* **251**, 9-14.
- Champness, J. N., Achari, A., Ballantine, S. P., Bryant, P. K., Delves, C. J. & Stammers, D. K. (1994). The structure of *Pneumocystis carinii* dihydrofolate reductase to 1.9 Å resolution. *Structure* **2**, 915-924.
- Chen, R., Greer, A. & Dean, A. M. (1995). A highly active decarboxylating dehydrogenase with rationally inverted coenzyme specificity. *Proc. Natl. Acad. Sci.* **92**, 11666-11670.

- Chen, Z., Lee, W. R. & Chang, S. H. (1991). Role of aspartic acid 38 in the cofactor specificity of *Drosophila* alcohol dehydrogenase. *Eur. J. Biochem.* **202**, 263-267.
- Chen, Z., Tsigelny, I., Lee, W. R., Baker, M. E. & Chang, S. H. (1994). Adding a positive charge at residue 46 of *Drosophila* alcohol dehydrogenase increases cofactor specificity for NADP⁺. *FEBS lett.* **356**, 81-85.
- Clermont, S., Corbier, C., Mely, Y., Gerard, D., Wonacott, A. & 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**, 10178-10184.
- Cody, V., Luft, J. R., Ciszak, E., Kalman, T. I. & Freisheim, J. H. (1992). Crystal structure determination at 2.3 Å of recombinant human dihydrofolate reductase ternary complex with NADPH and methotrexate-γ-tetrazole. *Anti-Cancer Drug Design* **7**, 483-491.
- Collaborative Computing Project Number 4. (1994). The CCP4 suite : programs for protein crystallography. *Acta Crystallog.* **D50**, 760-763.
- Corbier, C., Branlant, C., Wonacott, A. & Branlant, G. (1989). Site-directed mutagenesis of glyceraldehyde-3-phosphate reveals the role of residue Ser148. *Protein Eng.* **2**, 559-562.
- Corbier, C., Clermont, S., Billard, P., Skarzynski, T., Branlant, C., Wonacott, A. & Branlant, G. (1990a). Probing the coenzyme specificity of glyceraldehyde-3-phosphate dehydrogenases by site-directed mutagenesis. *Biochemistry* **29**, 7101-7106.
- Corbier, C., Mougin, A., Mely, Y., Adolph, H. W., Zeppezauer, M., Gerard, D., Wonacott, A. & Branlant, G. (1990b). The nicotinamide subsite of glyceraldehyde-3-phosphate dehydrogenase studied by site-directed mutagenesis. *Biochimie* **72**, 545-554.
- Corbier, C., Michels, S., Wonacott, A. J. & 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**, 3260-3265.
- Crowther, R. A. & Blow, D. M. (1967). A method of positioning a known molecule in an unknown crystal structure. *Acta Crystallog.* **23**, 544-548.
- Crowther, R. A. (1972). The fast rotation function. In *The Molecular Replacement Method* (Rossmann, M. G., ed.), pp. 173-178. Gordon & Breach Science Publishers, New York.
- Cruickshank, W. J. (1996). Protein precision re-examined : Luzzati plots do not estimate final errors. *Acta Crystallog.*, C85.
- Dauter, Z., Lamzin, V. S. & Wilson, K. S. (1995). Proteins at atomic resolution. *Current Opinion in Structural Biology* **5**, 784-790.
- Dessen, A., Quémard, A., Blanchard, J. S., Jacobs, W. R. J. & Sacchettini, J. C. (1995). Crystal structure and function of the isoniazid target of *Mycobacterium tuberculosis*. *Science* **267**, 1638-1641.
- Devedjiev, Y., Popov, A. N., Ivanova, V. & Atanasov, B. (1994). Crystallization and preliminary X-ray analysis of the complex of pig muscle D-glyceraldehyde-3-phosphate dehydrogenase with o-phenanthroline. *Crystallography Reports* **39**, 665-666.

- Diamond, R. (1969). Profile Analysis in Single Crystal Diffractometry. *Acta Crystallog.* **A25**, 43-55.
- Drenth, J. (1994). *Principles of Protein X-ray Crystallography* (Cantor, C. R., Ed.), Springer-Verlag, New York.
- Ducruix, A. & Giegé, R. (1992). *Crystallization of Nucleic Acids and Proteins : A Practical Approach*, IRL Press, Oxford.
- Duée, E., Olivier-Deyris, L., Fanchon, E., Corbier, C., Branlant, G. & 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**, 814-838.
- Dunn, C. R., Wilks, H. M., Halsall, D. J., Atkinson, T., Clarke, A. R., Muirhead, H. & Holbrook, J. J. (1991). Design and Synthesis of new enzymes based on the Lactate Dehydrogenase framework. *Philos. Trans. R. Soc. Lond. (Biol)* **332**, 177-184.
- El-Kabbani, O., Green, N. C., Lin, G., Carson, M., Narayana, S. V. L., Moore, K. M., Flynn, T. G. & DeLucas, L. J. (1994). Structures of human and porcine aldehyde reductase: An enzyme implicated in diabetic complications. *Acta Crystallog.* **D50**, 859-868.
- El-Kabbani, O., Judge, K., Ginell, S. L., Myles, D. A. A., DeLucas, L. J. & Flynn, T. G. (1995). Structure of porcine aldehyde reductase holoenzyme. *Nature Struct. Biol.* **2**, 687-692.
- Engh, R. A. & Huber, R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallog.* **A47**, 392-400.
- Eyschen, J., Vitoux, B., Rahuel-Clermont, S., Marraud, M., Branlant, G. & Cung, M. T. (1996). Phosphorous-31 nuclear magnetic resonance studies on coenzyme binding and specificity in glyceraldehyde-3-phosphate dehydrogenase. *Biochemistry* **35**, 6064-6072.
- Fabry, S., Lehmachner, A., Bode, W. & Hensel, R. (1988). Expression of the glyceraldehyde-3-phosphate dehydrogenase gene from the extremely thermophilic archaebacterium *Methanothermus fervidus* in *E. coli*. *FEBS lett.* **237**, 213-217.
- Fan, F., Lorenzen, J. A. & Plapp, B. V. (1991). An aspartate residue in yeast alcohol dehydrogenase I determines the specificity for coenzyme. *Biochemistry* **30**, 6397-6401.
- Fermi, G. (1975). The Three-dimensional Fourier synthesis of human deoxyhaemoglobin at 2.5 Å resolution : refinement of the atomic model. *J. Mol. Biol.* **97**, 237-256.
- Ferri, G., Comerio, G., Iaradola, P., Zapponi, M. C. & Speranza, M. L. (1978). Subunit structure and activity of glyceraldehyde-3-phosphate dehydrogenase from spinach chloroplasts. *Biochim. Biophys. Acta* **522**, 19-31.
- Ferri, G., Stoppini, M., Meloni, M., Zapponi, M. C. & Iaradola, P. (1990). Chloroplast glyceraldehyde-3-phosphate dehydrogenase (NADP) - Amino acid sequence of the subunits from isoenzyme I and structural relationship with isoenzyme II. *Biochim. Biophys. Acta* **1041**, 36-42.
- Filman, D. J., Bolin, J. T., Matthews, D. A. & Kraut, J. (1982). Crystal structures of *Escherichia coli* and *Lactobacillus casei* dihydrofolate reductase refined at 1.7 Å resolution. II. Environment of bound NADPH and implications for catalysis. *J. Biol. Chem.* **257**, 13663-13672.

- Fita, I., Silva, A. M., Murthy, M. R. N. & Rossmann, M. G. (1986). The refined structure of beef liver catalase at 2.5 Å resolution. *Acta Crystallog.* **42**, 497-515.
- Fox, K. M. & Karplus, P. A. (1994). Old yellow enzyme at 2 Å resolution: Overall structure, ligand binding, and comparison with related flavoproteins. *Structure* **2**, 1089-1105.
- Fujinaga, M. & Read, R. (1987). Experiences with a new translation function program, BRUTE. *J. Appl. Crystallog.* **20**, 517-521.
- Ghosh, D., Wawrzak, Z., Weeks, C. M., Duax, W. L. & Erman, M. (1994). The refined three-dimensional structure of 3 α ,20 β -hydroxysteroid dehydrogenase and possible roles of the residues conserved in short-chain dehydrogenases. *Structure* **2**, 629-640.
- Ghosh, D., Pletnev, V. Z., Zhu, D.-W., Wawrzak, Z., Duax, W. L., Pangborn, W., Labrie, F. & Lin, S.-X. (1995). Structure of human estrogenic 17 β -hydroxysteroid dehydrogenase at 2.20 Å resolution. *Structure* **3**, 503-513.
- Gouet, P., Jouve, H.-M. & Dideberg, O. (1995). Crystal structure of *Proteus mirabilis* PR catalase with and without bound NADPH. *J. Mol. Biol.* **249**, 933-954.
- Griffith, J. P., Lee, B., Murdock, A. L. & Amelunxen, R. E. (1983). Molecular symmetry of glyceraldehyde-3-phosphate dehydrogenase from *Bacillus coagulans*. *J. Mol. Biol.* **169**, 963-974.
- Grimshaw, C. E., Matthews, D. A., Varughese, K. I., Skinner, M., Xuong, N. H., Bray, T., Hoch, J. & Whiteley, J. M. (1992). Characterization and nucleotide binding properties of a mutant dihydropteridine reductase containing an aspartate 37-isoleucine replacement. *J. Biol. Chem.* **267**, 15334-15339.
- Haefner-Gormley, L., Chen, Z., Zalkin, H. & Colman, R. F. (1992). Importance of lysine-286 at the NADP site of glutamate dehydrogenase from *Salmonella typhimurium*. *Biochemistry* **31**, 7807-7814.
- Hanukoglu, I. & Gutfinger, T. (1989). cDNA sequence of adrenodoxin reductase. Identification of NADP-binding sites in oxidoreductases. *Eur. J. Biochem.* **180**, 479-484.
- Harper, E. T. & Rose, G. D. (1993). Helix stop signals in proteins and peptides : the capping box. *Biochemistry* **32**, 7605-7609.
- Harris, J. I. & Perham, R. N. (1965). Glyceraldehyde-3-phosphate Dehydrogenases - I The Protein Chains in Glyceraldehyde-3-phosphate Dehydrogenase from Pig Muscle. *J. Mol. Biol.* **13**, 876-884.
- Harris, J. I. & Waters, M. (1976). Glyceraldehyde-3-phosphate dehydrogenase. In *The Enzymes* (Boyer, P. D., ed.), Vol. 13, pp. 1-49. Academic Press, New York.
- Harrison, S. C. (1968). A Point-Focusing Camera for Single-Crystal Diffraction. *J. Appl. Crystallog.* **1**, 84-90.
- Hecht, H. J., Erdmann, H., Park, H. J., Sprinzl, M. & Schmid, R. D. (1995). Crystal structure of NADH oxidase from *Thermus thermophilus*. *Nature Struct. Biol.* **2**, 1109-1114.
- Hendrickson, W. A. (1985). Stereochemically restrained refinement of macromolecular structures. *Methods in Enzymol.* **115**, 252-270.

Hol, W. G. J., Van Duijnen, P. T. & Berendsen, H. J. C. (1978). The α -helix dipole and the properties of proteins. *Nature* **273**, 443-446.

Hol, W. G. J. (1985). The role of the α -helix dipole in protein function and structure. *Prog. Biophys. Molec. Biol.* **45**, 149-195.

Holmes, M. A., Remington, S. J., Schwendimann, B., Christie, G. E. & Matthews, B. W. (1977). Crystallographic Data for Sturgeon Holo-D-Glyceraldehyde-3-Phosphate Dehydrogenase : A Holo-D-Glyceraldehyde-3-Phosphate Dehydrogenase with Crystallographic 2-fold Symmetry. *J. Mol. Biol.* **112**, 651-652.

Hoog, S. S., Pawlowski, J. E., Alzari, P. M., Penning, T. M. & Lewis, M. (1994). Three-dimensional structure of rat liver 3 α -hydroxysteroid/dihydrodiol dehydrogenase: A member of the aldo-keto reductase superfamily. *Proc. Natl. Acad. Sci. USA* **91**, 2517-2521.

Huang, S., Appleman, J. R., Tan, X., Thompson, P. D., Blakley, R. L., Sheridan, R. P., Venkataraghavan, R. & Freisheim, J. H. (1990). Role of lysine-54 in determining cofactor specificity and binding in human dihydrofolate reductase. *Biochemistry* **29**, 8063-8069.

Hurley, J. H., Dean, A. M., Koshland, D. E. & Stroud, R. M. (1991). Catalytic mechanism of NADP⁺-dependent isocitrate dehydrogenase: Implications from the structures of magnesium-isocitrate and NADP⁺ complexes. *Biochemistry* **30**, 8671-8678.

Hurley, J.H. & Dean, A.M. (1994). Structure of 3-isopropylmalate dehydrogenase in complex with NAD⁺: Ligand-induced loop closing and mechanism for cofactor specificity. *Structure* **2**, 1007-1016.

Hurley, J. H., Chen, R. & Dean, A. M. (1996). Determinants of cofactor specificity in isocitrate dehydrogenase: Structure of an engineered NADP⁺->NAD⁺ specificity-reversal mutant. *Biochemistry* **35**, 5670-5678.

IUBAC-IUB, Joint, Commission, on, Biochemical & Nomenclature. (1983). Abbreviations and Symbols for the description of conformations of polynucleotide chains. *Eur. J. Biochem.* **131**, 9-15.

Jancarik, J. & Kim, S. H. (1991). Sparse matrix sampling : a screening method for crystallization of proteins. *J. Appl. Crystallog.* **24**, 409-411.

Kabsch, W. (1988). Evaluation of single-crystal X-ray diffraction data from a position-sensitive detector. *J. Appl. Crystallog.* **21**, 916-924.

Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallog.* **26**, 795-800.

Karplus, P. A. & Schulz, G. E. (1987). Refined structure of glutathione reductase at 1.54 resolution. *J. Mol. Biol.* **195**, 701-729.

Kelly, C. A., Nishiyama, M., Ohnishi, Y., Beppu, T. & Birktoft, J. J. (1993). Determinants of protein thermostability observed in the 1.9-Å crystal structure of malate dehydrogenase from the thermophilic bacterium *Thermus flavus*. *Biochemistry* **32**, 3913-3922.

- Kim, H., Feil, I. K., Verlinde, C. L. M. J., Petra, P. H. & Hol, W. G. J. (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**, 14975-14986.
- Kimura, H., Sumino, Y. & Suzuki, M. (1991). Cloning and sequencing of the putative glyceraldehyde-3-phosphate dehydrogenase gene from *Cephalosporium acremonium* and its application to heterologous gene expression. *J. Ferment. Bioeng.* **71**, 145-150.
- Korndörfer, I., Steipe, B., Huber, R., Tomschy, A. & 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**, 511-521.
- Kraulis, P. J. (1991). Molscript: A program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946-950.
- Labesse, G., Vidal-Cross, A., Chomilier, J. & Mornon, J. P. (1994). Structural comparisons lead to the definition of a new superfamily of NAD(P)(H)-accepting oxidoreductases: The single-domain reductases/epimerases/dehydrogenases (the 'RED' family). *Biochem. J.* **304**, 95-99.
- Lamzin, V. S., Dauter, Z., Popov, V. O., Harutyunyan, E. H. & Wilson, K. S. (1994). High resolution structures of holo and apo formate dehydrogenase. *J. Mol. Biol.* **236**, 759-785.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK : a program to check the stereochemical quality of protein structures. *J. Appl. Crystall.* **26**, 283-291.
- Lauvergeat, V., Kennedy, K., Feuillet, C., McKie, J. H., Gorrichon, L., Baltas, M., Boudet, A. M., Grima-Pettenati, J. & Douglas, K. T. (1995). Site-directed mutagenesis of a serine residue in cinnamyl alcohol dehydrogenase, a plant NADPH-dependent dehydrogenase, affects the specificity for the coenzyme. *Biochemistry* **34**, 12426-12434.
- Lawrence, C. M., Rodwell, V. W. & Stauffacher, C. V. (1995). Crystal structure of *Pseudomonas mevalonii* HMG-CoA reductase at 3.0 angstrom resolution. *Science* **268**, 1758-1762.
- Lee, B., Griffith, J. P., Park, C. H., Sheldon, R. I., McLinden, J., Murdock, A. L. & Amelunxen, R. E. (1982). Preliminary cristallographic data for glyceraldehyde-3-phosphate dehydrogenase from the thermophile *Bacillus coagulans*. *J. Mol. Biol.* **158**, 153-156.
- Lesk, A. M. (1995). NAD-binding domains of dehydrogenases. *Curr. Opin. Struct. Biol.* **5**, 775-783.
- Leslie, A. G. W. & Wonacott, A. J. (1983). Coenzyme binding in crystals of glyceraldehyde-3-phosphate dehydrogenase. *J. Mol. Biol.* **165**, 375-391.
- Levy, H. R., Vought, V. E., Yin, X. & Adams, M. J. (1996). Identification of an arginine residue in the dual coenzyme-specific glucose-6-phosphate dehydrogenase from *Leuconostoc mesenteroides* that plays a key role in binding NADP⁺ but not NAD⁺. *Arch. Biochem. Biophys.* **326**, 145-151.
- Lin, Z.-J., Li, J., Zhang, F.-M., Song, S.-Y., Yang, J., Liang, S.-J. & Tsou, C.-L. (1993). Structure of D-glyceraldehyde-3-phosphate dehydrogenase from *Palinurus versicolor* carrying the fluorescent NAD derivatives at 2.7 Å resolution. *Arch. Biochem. Biophys.* **302**, 161-166.

Luzzati, V. (1952). Traitement statistique des erreurs dans la détermination des structures cristallines. *Acta Crystallog.* **5**, 802-810.

Martin, W. & Cerff, R. (1986). Prokaryotic features of a nucleous-encoded enzyme- cDNA sequences for chloroplast and cytosolic glyceraldehyde-3-phosphate dehydrogenase from mustard. *Eur. J. Biochem.* **159**, 323-331.

Martin, W., Brinkmann, H., Savonna, C. & Cerff, R. (1993). Evidence for a chimeric nature of nuclear genomes: Eubacterial origin of eukaryotic glyceraldehyde-3-phosphate dehydrogenase genes. *Proc. Natl. Acad. Sci. USA* **90**, 8692-8696.

Mathews, F. S. (1991). New flavoenzymes. *Current Op. Struct. Biol.* **1**, 954-967.

Mattevi, A., Obmolova, G., Sokatch, J.R., Betzel, C. & Hol, W. G. J. (1992). The refined crystal structure of *Pseudomonas putida* lipoamide dehydrogenase complexed with NAD⁺ at 2.45 Å resolution. *Proteins* **13**, 336-351.

Matthews, D. A., Alden, R. A., Bolin, J. T., Filman, D. J., Freer, S. T., Hamlin, R., Hol, W. G. J., Kisliuk, R. L., Pastore, E. J., Plante, L. T., Xuong, N.-H. & Kraut, J. (1978). Dihydrofolate reductase from *Lactobacillus casei*. X-ray structure of the enzyme-methotrexate-NADPH complex. *J. Biol. Chem.* **253**, 6946-6954.

McRee, D. E. (1993). *Practical Protein Crystallography*, Academic Press, San Diego.

McTigue, M. A., Davies II, J. F., Kaufman, B. T. & Kraut, J. (1992). Crystal structure of chicken liver dihydrofolate reductase complexed with NADP⁺ and bipterin. *Biochemistry* **31**, 7264-7273.

Mercer, W. D., Winn, S. I. & Watson, H. C. (1976). Twinning in crystals of human skeletal muscle D-glyceraldehyde-3-phosphate dehydrogenase. *J. Mol. Biol.* **104**, 277-283.

Metzger, M. H. & Hollenberg, C. P. (1995). Amino acid substitutions in the yeast *Pichia stipitis* xylitol dehydrogenase coenzyme-binding domain affect the coenzyme specificity. *Eur. J. Biochem.* **228**, 50-54.

Mittl, P. R. E., Berry, A., Scrutton, N. S., Perham, R. N. & Schulz, G. E. (1994). Anatomy of an engineered NAD-binding site. *Protein Science* **3**, 1504-1514.

Mittl, P. R. E. & Schulz, G. E. (1994). Structure of glutathione reductase from *Escherichia coli* at 1.86 Å resolution: Comparison with the enzyme from human erythrocytes. *Protein Science* **3**, 799-809.

Miyazaki, K. & Oshima, T. (1994). Co-enzyme specificity of 3-isopropylmalate dehydrogenase from *Thermus thermophilus* HB8. *Protein Eng.* **7**, 401-403.

Moodie, S. L., Mitchell, J. B. O. & Thornton, J. M. (1996). Protein Recognition of Adenylate : An Example of a Fuzzy Recognition Template. *J. Mol. Biol.* **263**, 486-500.

Moras, D., Olsen, K. W., Sabesan, M. N., Buehner, M., Ford, G. C. & Rossmann, M. G. (1975). Studies of asymmetry in the three-dimensional structure of lobster D-glyceraldehyde-3-phosphate dehydrogenase. *J. Biol. Chem.* **250**, 9137-9162.

Mougin, 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. *Prot. Eng.* **2**, 859-872.

Murthy, M. R. N., Garavito, R. M., Johnson, J. E. & Rossmann, M. J. (1980). Structure of lobster apo-D-glyceraldehyde-3-phosphate dehydrogenase at 3.0 Å resolution. *J. Mol. Biol.* **138**, 859-872.

Navaza, J. (1987). On the Fast Rotation Function. *Acta Crystallog.* **A43**, 645-653.

Navaza, J. (1994). AMoRe : an Automated Package for Molecular Replacement. *Acta Crystallog.* **A50**, 157-163.

Niefind, K., Hecht, H.-J. & Schomburg, D. (1995). Crystal structure of L-2-hydroxyisocaproate dehydrogenase from *Lactobacillus confusus* at 2.2 Å resolution. An example of strong asymmetry between subunits. *J. Mol. Biol.* **251**, 256-281.

Nishiyama, M., Birktoft, J. J. & Beppu, T. (1993). Alteration of coenzyme specificity of malate dehydrogenase from *Thermus flavus* by site-directed mutagenesis. *J. Biol. Chem.* **268**, 4656-4660.

Olivier, L., Buisson, G., Fanchon, E., Corbier, C., Branlant, G. & Dideberg, O. (1995). Crystallization and preliminary X-ray diffraction studies of *Escherichia coli* glyceraldehyde-3-phosphate dehydrogenase. *Acta Cryst.* **D51**, 245-247.

Otwinowski, Z. (1993). Oscillation data reduction program. In *Proceedings of the CCP4 study weekend : data collection and processing* (Sawyer, L., Isaacs, N. & Bailey, S., eds.), pp. 56-62, SERC Daresbury Laboratory, Warrington.

Pai, E. F. (1991). Variations on a theme: the family of FAD-dependent NAD(P)H-(disulfide)-oxidoreductases. *Curr. Opin. Struct. Biol.* **1**, 796-803.

Quigley, F., Brinkmann, H., Martin, W. F. & Cerff, R. (1989). Strong functional GC pressure in a light-regulated maize gene encoding subunit GAPA of chloroplast glyceraldehyde-3-phosphate dehydrogenase: Implications for the evolution of GAPA pseudogenes. *J. Mol. Evol.* **29**, 412-421.

Racker, E. & Krimsky, I. (1952). The mechanism of oxidation of aldehydes by glyceraldehyde-3-phosphate dehydrogenase. *J. Biol. Chem.* **198**, 731-743.

Rafferty, J. B., Simon, J. W., Baldock, C., Artymiuk, P. J., Baker, P. J., Stuitje, A. R., Slabas, A. R. & Rice, D. W. (1995). Common themes in redox chemistry emerge from the structure of oilseed rape (*Brassica napus*) enoyl acyl carrier protein reductase. *Structure* **3**, 927-938.

Rahuel-Clermont, S. (1994). Etude des facteurs moléculaires responsables de la spécificité de cofacteur des déshydrogénases à NAD(P). Thèse de l'Université H. Poincaré, Nancy I.

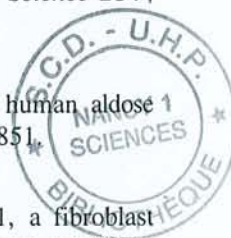
Ramakrishnan, C., Ramachandran, G. N. & Sasisekharan, V. (1965). Stereochemical criteria for polypeptide and protein chain conformations. II Allowed conformations for a pair of peptide units. *Biophys. J.* **5**, 909-933.

Reddy, S. G., Scapin, G. & Blanchard, J. S. (1996). Interaction of Pyridine Nucleotide Substrates with *Escherichia coli* Dihydropyridine Reductase : Thermodynamic and Structural Analysis of Binary Complexes. *Biochemistry* **35**, 13294-13302.

- Rodgers, D. W. (1994). Cryocrystallography. *Structure* **2**, 1135-1140.
- Rondeau, J.-M., Tête-Favier, F., Podjarny, A., Reymann, J.-M., Barth, P., Biellmann, J.-F. & Moras, D. (1992). Novel NADPH-binding domain revealed by the crystal structure of aldose reductase. *Nature* **355**, 469-472.
- Rossmann, M. G., Adams, M. J., Buehner, M., Ford, G. C., Hackert, M. L., Liljas, A., Rao, S. T., Banaszak, L. J., Hill, E., Tsernoglou, D. & Webb, L. (1973). Molecular symmetry axes and subunit interface in certain dehydrogenases. *J. Mol. Biol.* **76**, 533-537.
- Rossmann, M. G., Moras, D. & Olsen, K. W. (1974). Chemical and biological evolution of a nucleotide-binding protein. *Nature* **250**, 194-199.
- Rossmann, M. G., Liljas, A., Brändén, C.-I. & Banaszak, L. J. (1975). Evolutionary and structural relationships among dehydrogenases. In *The Enzymes* (Boyer, P. D., ed.), Vol. 11, pp. 61-102. Academic Press, New York.
- Roussel, A. & Cambillau, C. (1989). TURBO-FRODO. In *Silicon Graphics Geometry Partner Directory* (SiliconGraphics, ed.), pp. 77-78. Silicon Graphics, Mountain View.
- Sawaya, M. R. & Kraut, J. (1997). Loop and Subdomain Movements in the Mechanism of *Escherichia Coli* Dihydrofolate Reductase : Crystallographic Evidence. *Biochemistry* **36**, 586-603.
- Scapin, G., Blanchard, J. S. & Sacchettini, J. C. (1995). Three-dimensional structure of *Escherichia coli* dihydrodipicolinate reductase. *Biochemistry* **34**, 3502-3512.
- Schuller, D. J., Grant, G. A. & Banaszak, L. J. (1995). The allosteric ligand site in the Vmax-type cooperative enzyme phosphoglycerate dehydrogenase. *Struct. Biol.* **2**, 69-76.
- Schulz, G. E. (1992). Binding of nucleotides by proteins. *Curr. Op. Struct. Biol.* **2**, 61-67.
- Scrutton, N. S., Berry, A. & Perham, R. N. (1990). Redesign of the coenzyme specificity of a dehydrogenase by protein engineering. *Nature* **343**, 38-43.
- Seale, J. W., Srinivasan, R. & Rose, G. D. (1994). Sequence determinants of the capping box, a stabilizing motif at the N-termini of α -helices. *Protein Sci.* **3**, 1741-1745.
- Segal, H. L. & Boyer, P. D. (1953). The role of sulfhydryl groups in the activity of D-glyceraldehyde-3-phosphate dehydrogenase. *J. Biol. Chem.* **204**, 265-281.
- Sem, D. S. & Kasper, C. B. (1993a). Enzyme-substrate binding interactions of NADPH-cytochrome P-450 oxidoreductase characterized with pH and alternate substrate/inhibitor studies. *Biochemistry* **32**, 11539-11547.
- Sem, D. S. & Kasper, C. B. (1993b). Interaction with arginine 597 of NADPH-cytochrome p-450 oxidoreductase is a primary source of the uniform binding energy used to discriminate between NADPH and NADH. *Biochemistry* **32**, 11548-11558.
- Shih, M. C., Lazar, G. & Goodman, H. M. (1986). Evidence in favor of the symbiotic origin of chloroplasts - Primary structure and evolution of tobacco glyceraldehyde-3-phosphate dehydrogenases. *Cell* **47**, 73-80.

- Skarzynski, T., Moody, P. C. E. & Wonacott, A. J. (1987). Structure of holo-glyceraldehyde-3-phosphate dehydrogenase from *Bacillus stearothermophilus* at 1.8 resolution. *J. Mol. Biol.* **193**, 171-187.
- Skarzynski, T. & Wonacott, A. J. (1988). Coenzyme-induced conformational changes in glyceraldehyde-3-phosphate dehydrogenase from *Bacillus stearothermophilus*. *J. Mol. Biol.* **203**, 1097-1118.
- Song, S. Y., Gao, Y. G., Zhou, J. M. & Tsou, C. L. (1983). Preliminary Cristallographic Studies of Lobster D-Glyceraldehyde-3-phosphatenDehydrogenase and the Modified Enzyme Carrying the Fluorescent Derivative. *J. Mol. Biol.* **171**, 225-228.
- Soukri, A., Mougin, A., Corbier, C., Wonacott, A., Branlant, C. & Branlant, G. (1989). Role of the histidine 176 residue in glyceraldehyde-3-phosphate dehydrogenase as probed by site-directed mutagenesis. *Biochemistry* **28**, 2586-2592.
- Stanton, M., Phillips, W., Li, Y. & Kalata, K. (1992). The Detective Quantum Efficiency of CCD and Vidicon-Based Detectors for X-ray Cristallographic Applications. *J. Appl. Crystallog.* **25**, 638-645.
- Stehle, T., Claiborne, A. & Schulz, G. E. (1993). NADH binding site and catalysis of NADH peroxidase. *Eur. J. Biochem.* **211**, 221-226.
- Stoddard, B. L., Dean, A. & Koshland, D. E. J. (1993). Structure of isocitrate dehydrogenase with isocitrate, nicotinamide adenine dinucleotide phosphate, and calcium at 2.5-Å resolution: A pseudo-Michaelis ternary complex. *Biochemistry* **32**, 9310-9316.
- Stryer, L. (1995). *Biochemistry*, W.H. Freeman and Company, New York.
- Su, Y., Varughese, K. I., Xuong, N. H., Bray, T. L., Roche, D. J. & Whiteley, J. M. (1993). The crystallographic structure of a human dihydropteridine reductase NADH binary complex expressed in *Escherichia coli* by a cDNA constructed from its rat homologue. *J. Biol. Chem.* **268**, 26836-26841.
- Sundaralingam, M. & Sekharudu, Y. C. (1989). Water-inserted α -helical segments implicate reverse turns as folding intermediates. *Science* **244**, 1333-1337.
- Swindells, M. B. (1993). Classification of doubly wound nucleotide binding topologies using automated loop searches. *Protein Science* **2**, 2146-2153.
- Tanaka, N., Nonaka, T., Nakanishi, M., Deyashiki, Y., Hara, A. & Mitsui, Y. (1996). Crystal structure of the ternary complex of mouse lung carbonyl reductase at 1.8 Å resolution: the structural origin of coenzyme specificity in the short-chain dehydrogenase/reductase family. *Structure* **4**, 33-45.
- Tanner, J. J., Hecht, R. M. & K.L., F. (1996). Determinants of enzyme thermostability observed in the molecular structure of *Thermus aquaticus* D-glyceraldehyde-3-phosphate dehydrogenase at 2.5 Å resolution. *Biochemistry* **35**, 2597-2609.
- Thoden, J. B., Frey, P. A. & Holden, H. M. (1996). Crystal structures of the oxidized and reduced forms of UDP-galactose 4-epimerase isolated from *Escherichia coli*. *Biochemistry* **35**, 2557-2566.
- Thomas, J. A., Arnold, J. R. P., Basran, J., Andrews, J., Roberts, G. C. K., Birdsall, B. & Feeney, J. (1994). Effects of substitution of Thr63 by alanine on the structure and function of *Lactobacillus casei* dihydrofolate reductase. *Protein Eng.* **7**, 783-792.

- Thorn, J. M., Barton, J. D., Dixon, N. E., Ollis, D. L. & Edwards, K. J. (1995). Crystal structure of *Escherichia coli* QOR quinone oxidoreductase complexed with NADPH. *J. Mol. Biol.* **249**, 785-799.
- Varughese, K. I., Skinner, M. M., Whiteley, J. M., Matthews, D. A. & Xuong, N. H. (1992). Crystal structure of rat liver dihydropteridine reductase. *Proc. Natl. Acad. Sci. USA* **89**, 6080-6084.
- Vellieux, F. M. D., Hajdu, J. & Hol, W. G. J. (1995). Refined 3.2 Å structure of glycosomal holo-glyceraldehyde phosphate dehydrogenase from *Trypanosoma brucei brucei*. *Acta Cryst.* **D51**, 575-589.
- Waksman, G., Krishna, T. S. R., Williams, C. H. J. & Kuriyan, J. (1994). Crystal structure of *Escherichia coli* thioredoxin reductase refined at 2 Å resolution. Implications for a large conformational change during catalysis. *J. Mol. Biol.* **236**, 800-816.
- Watson, H. C. & Banaszak, L. J. (1964). Structure of Glyceraldehyde-3-phosphate Dehydrogenase. Structural Symmetry within the molecule. *Nature* **204**, 918-920.
- Watson, H. C., Duée, E. & Mercer, W. D. (1972). Low Resolution Structure of Glyceraldehyde-3-phosphate Dehydrogenase. *Nature New Biol.* **240**, 130-139.
- Wierenga, R. K., De Maeyer, M. C. H. & Hol, W. G. J. (1985). Interaction of pyrophosphate moieties with α -helices in dinucleotide binding proteins. *Biochemistry* **24**, 1346-1357.
- Wierenga, R. K., Terpstra, P. & Hol, W. G. J. (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**, 101-107.
- Wilson, D. K., Bohren, K. M., Gabbay, K. H. & Quioco, F. A. (1992). An unlikely sugar substrate site in the 1.65 Å structure of the human aldose reductase holoenzyme implicated in diabetic complications. *Science* **257**, 81-84.
- Wilson, D. K., Tarle, I., Petrash, J. M. & Quioco, F. A. (1993). Refined 1.8 Å structure of human aldose reductase complexed with the potent inhibitor zopolrestat. *Proc. Natl. Acad. Sci. USA* **90**, 9847-9851.
- Wilson, D. K., Nakano, T., Petrash, J. M. & Quioco, F. A. (1995). 1.7 Å structure of FR-1, a fibroblast growth factor-induced member of the aldo-keto reductase family, complexed with coenzyme and inhibitor. *Biochemistry* **34**, 14323-14330.
- Wlodawer, A. & Hodgson, K. O. (1975). Crystallization and Crystal Data of Monellin. *Proc. Natl. Acad. Sci.* **72**, 398.
- Yaoi, T., Miyazaki, K. & Oshima, T. (1994). Role of Arg²³¹ and Tyr²⁸⁴ of *Thermus thermophilus* isocitrate dehydrogenase in the coenzyme specificity. *FEBS Lett.* **355**, 171-172.
- You, K.-S. (1982). Stereospecificities of the pyridine nucleotide-linked enzymes. *Methods in Enzymol.* **87**, 101-126.



Nom : DIDIERJEAN

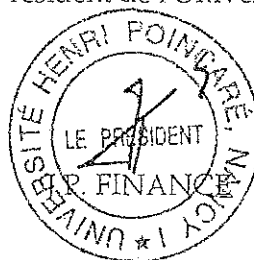
Prénom : Claude

DOCTORAT de l'UNIVERSITÉ HENRI POINCARÉ, NANCY-I
en SCIENCES DES MATERIAUX

VU, APPROUVÉ ET PERMIS D'IMPRIMER

Nancy, le 30 AVR 1997 UHP 37

Le Président de l'Université



3C-F9