

Université de Montréal

La Fructose-1,6-bisphosphate aldolase de *Thermus aquaticus*:  
séquence du gène, caractérisation et cristallisation de la protéine  
surexprimée.

par

Véronique Sauvé

Département de biochimie

Faculté de médecine

Mémoire présenté à la faculté des études supérieures  
en vue de l'obtention du grade de  
Maître ès sciences (M.Sc.)  
en biochimie

Décembre, 1999

©Véronique Sauvé, 1999



2.5952.11m2

W  
4  
U58  
2000  
V.065

Université de Montréal

La Fonction-1- $\beta$ -glycosyltransférase hépatique de *Rattus norvegicus*  
régule la glycosylation et cristallisation de la protéine  
sucrolyse.

Vétérinaire  
Département de biochimie  
Faculté de médecine

Mémoire présenté à la Faculté des études supérieures  
en vue de l'obtention du grade de  
Maître en sciences (M.Sc.)  
en biochimie



Décembre 1999  
Suzanne Gauthier, M.Sc.

Université de Montréal  
Faculté des études supérieures

Ce mémoire intitulé:

La Fructose-1,6-bisphosphate aldolase de *Thermus aquaticus*:  
séquence du gène et cristallisation de la protéine surexprimée.

présenté par:

Véronique Sauvé

a été évalué par un jury composé des personnes suivantes:

Président-rapporteur: Crine, Philippe

Directeur de recherche: Sygusch, Jurgen

Membre du jury: Morais, Réjean

Mémoire accepté le: 00-05-03

## Sommaire

Les fructose-1,6-bisphosphate (FBP) aldolases sont des enzymes qui catalysent le clivage réversible de la fructose-1,6-bisphosphate en dihydroxyacétone phosphate et en glycéraldéhyde-3-phosphate dans la voie de la glycolyse, de la gluconéogenèse et le cycle de Calvin. Il en existe deux classes distinctes. Les enzymes de classe I présents chez les eucaryotes supérieurs, les algues vertes et certaines bactéries, forment un intermédiaire base de Schiff entre le substrat et une lysine de leur site actif. Les FBP aldolases de classe II, subdivisées en deux groupes (A et B), sont retrouvées chez les eucaryotes inférieurs et les procaryotes et utilisent un ion métallique bivalent dans leur mécanisme réactionnel. Jusqu'à maintenant la seule structure tridimensionnelle de FBP aldolases de classe II à avoir été résolue est celle de l'enzyme de classe II A du mésophile *E.coli*.

Les thermophiles qui pourraient être des reliques des organismes primitifs des débuts de la vie sur la Terre, sont composés de macromolécules adaptées aux hautes températures. L'étude de leurs protéines permet de comprendre quelles modifications, le plus souvent mineures, leur confèrent leur thermostabilité. Cette caractérisation est d'autant plus intéressante que l'utilisation d'enzymes thermostables comporte des avantages industriels et commerciaux importants. La technique de PCR en est le meilleur exemple.

Ce mémoire présente la caractérisation des séquences nucléique et protéique de la fructose-1,6-bisphosphate aldolase de l'eubactérie thermophile *Thermus aquaticus* YT-1. Le gène a d'abord été repêché par l'hybridation de l'ADN génomique de *Taq* avec une sonde dégénérée correspondant à l'extrémité N-terminale de la protéine. Sa séquence nucléique très riche en G/C code pour une protéine de 305 acides aminés ayant un poids moléculaire de 33 092 Da. La FBP aldolase de *Taq* fait partie de la classe II B et est très similaire à celle de *Thermotoga maritima*, d'*Aquifex aeolicus* et d'*Helicobacter pylori*. Elle contient les acides aminés identifiés chez l'enzyme de classe II A d'*E.coli* comme étant impliqués dans la réaction aldolasique. Mis à part quelques raccourcissement de boucles et d'hélices, la structure prédite de la FBP aldolase de *Taq* concorde avec la structure de baril  $(\alpha/\beta)_8$  d'*E.coli*. Fait à noter, une insertion de 21 acides aminés au milieu du site de liaison du DHAP pourrait jouer un rôle dans la régulation de l'activité et/ou la tétramérisation de l'enzyme.

Le gène de la FBP aldolase de *Taq* a été amplifié et ses régions flanquantes modifiées par PCR. Le fragment a été inséré dans pGEM-T, digéré, puis transféré dans le vecteur d'expression pKK223-3. La protéine a été surexprimée dans des *E.coli* JM109 suite à une induction à l'IPTG. Elle a ensuite été purifiée jusqu'à homogénéité en trois étapes, soit par un traitement à la chaleur, deux colonnes, une Phosphocellulose-P11 et une Q-Sepharose. La protéine purifiée s'avère être un homotétramère peu actif à moins de 60°C et stable à 80°C pour au moins 2 heures. La cinétique de l'enzyme recombinant démontre une coopérativité négative ( $h = 0,32$ ) avec un  $K_M$  de 305  $\mu\text{M}$  et un  $V_{\text{MAX}}$  de 46 U/mg à 70°C.

Des cristaux de FBP aldolase native de *Taq* ont été obtenus par diffusion de vapeur dans une solution de sulfate d'ammonium à pH 7,5 en présence de cobalt et de détergent. Ces cristaux en forme d'aiguille de 0,1 mm X 0,1 mm X 1,5 mm diffractent jusqu'à 2,0 Å et ont un groupe spatial monoclinique P2 avec des paramètres de maille de  $a = 99,5 \text{ \AA}$ ,  $b = 57,55 \text{ \AA}$ ,  $c = 138,6 \text{ \AA}$  et  $\beta = 90,25^\circ$ . Chaque maille asymétrique contient un tétramère. Certains de ces cristaux natifs ont été trempés dans l'yttrium. Leur résolution et leurs paramètres de maille sont demeurés semblables à ceux non-trempés. Cependant l'incorporation de l'yttrium nulle ou très faible ne permet pas l'application de la technique de MIR pour la résolution de la structure tridimensionnelle.

De la FBP aldolase de *Taq* séléno-méthionée a été produite et partiellement purifiée. Des cristaux bipyramidaux de 0,25 mm X 0,16 mm X 0,16 mm ont été obtenus par diffusion de vapeur dans une condition de cristallisation contenant du sulfate d'ammonium dans un tampon à pH 4,0. Ces cristaux dans lesquels la présence de sélénium a été détectée, ont diffracté jusqu'à 2,6 Å. Ils ont un groupe spatial tétragonal I4<sub>1</sub> avec des paramètres de maille de  $a = 88,6 \text{ \AA}$ ,  $b = 88,6 \text{ \AA}$ ,  $c = 164,1 \text{ \AA}$ . Chaque maille asymétrique contient deux sous-unités (l'équivalent d'un dimère). Récemment des données cristallographiques complètes de ces cristaux à trois longueurs d'onde spécifiques ont été récoltées et permettront de solutionner la structure tridimensionnelle de la protéine grâce à la technique de "Multiple wavelength Anomalous Dispersion" (MAD).

**Mots clés:** aldolase - *Thermus aquaticus* - thermostabilité - surexpression - purification - cinétique - cristallisation

# Table des matières

<b>Sommaire</b> .....	iii
<b>Table des matières</b> .....	v
<b>Liste des tableaux</b> .....	viii
<b>Liste des figures</b> .....	ix
<b>Liste des sigles et des abréviations</b> .....	x
<b>Remerciements</b> .....	xii
<b>Chapitre 1: Introduction</b> .....	1
1.1 Les Fructose-1,6-bisphosphate aldolases.....	2
1.1.1 Bref historique.....	2
1.1.2 Les caractéristiques générales des FBP aldolases.....	2
1.1.3 Les FBP aldolases de classe I.....	4
1.1.3.1 Les FBP aldolases de classe I des eucaryotes.....	5
1.1.3.1.1 Leur distribution.....	5
1.1.3.1.2 Leur mécanisme réactionnel.....	6
1.1.3.1.3 Leur structure tridimensionnelle.....	6
1.1.3.2 Les FBP aldolases de classe I des procaryotes.....	8
1.1.4 Les FBP aldolases de classe II.....	9
1.1.4.1 Leur structure tridimensionnelle.....	10
1.1.4.2 Leur mécanisme réactionnel.....	11
1.1.5 Sa phylogénie.....	12
1.2 Les thermophiles.....	13
1.2.1 Leurs habitats.....	13
1.2.2 Leurs adaptations à la chaleur.....	14
1.2.2.1 Leurs lipides et leurs membranes.....	16
1.2.2.2 Leurs acides nucléiques et leurs ribosomes.....	17

1.2.2.3 Leurs protéines.....	18
1.2.2.3.1 Une augmentation de la rigidité.....	19
1.2.2.3.2 La substitution de certains acides aminés.....	20
1.2.2.3.3 Les interactions aromatique-aromatique.....	21
1.2.2.3.4 La stabilisation des hélices et des feuillettes.....	22
1.2.2.3.5 Les ponts salins et l'hydratation.....	22
1.2.2.3.6 Le ratio surface par volume et la densité d'empilement.....	23
1.2.2.3.7 Le niveau d'association des sous-unités.....	23
1.2.3 Leur intérêt industriel.....	24
1.2.3.1 L'utilisation d'organismes thermophiles.....	25
1.2.3.2 L'utilisation de protéines thermophiles.....	25
1.2.4 Leur place dans l'évolution.....	26
1.2.5 L'eubactérie thermophile <i>Thermus aquaticus</i> .....	29
1.2.5.1 Ses lipides.....	29
1.2.5.2 Ses acides nucléiques et son ribosome.....	30
1.2.5.3 Ses protéines.....	30
1.2.5.4 Sa fructose-1,6-bisphosphate aldolase.....	31
1.3 Les objectifs de ce projet.....	32
<b>Chapitre 2: Article</b> .....	33
Fructose-1,6-bisphosphate aldolase from <i>Thermus aquaticus</i> : gene sequence and crystallization of the overexpressed enzyme.	
Summary.....	36
Introduction.....	37
Materials and methods.....	38
Results and discussion.....	45
Acknowledgements.....	54
References.....	54

<b>Chapitre 3: Matériel et méthodes</b> .....	74
3.1 Les souches et les produits.....	75
3.2 La production et la purification de la FBP aldolase de <i>Taq</i> séléno-méthionée.....	75
3.3 Les conditions de cristallisation et de cryoprotection pour la FBP aldolase de <i>Taq</i> native et séléno-méthionée à pH 4,0.....	76
3.4 D'autres conditions de cristallisation de la FBP aldolase de <i>Taq</i> .....	76
3.5 Le trempage de cristaux de FBP aldolase native de <i>Taq</i> dans l'yttrium et leur condition de cryoprotection.....	77
3.6 La collecte de données cristallographiques.....	77
<b>Chapitre 4: Discussion et conclusion</b> .....	78
4.1 Discussion.....	79
4.2 Conclusion.....	86
<b>Bibliographie</b> .....	87



# Liste des tableaux

## Article

Tableau I. Summary of purification of *Taq* FBP aldolase from a 1-1 of *E.coli*-pKKaldoTaq.

## Liste des figures

Figure 1. La glycolyse et la gluconéogenèse

Figure 2. Le cycle de Calvin

Figure 3. Le mécanisme réactionnel de la FBP aldolase de classe I de muscle de lapin

Figure 4. Le mécanisme réactionnel proposé pour la FBP aldolase de classe II d'*E.coli*

Figure 5. Photographie de deux cristaux de la FBP aldolase de *Taq* sélénométhionée.

### Article

Figure 1. Nucleotide sequence of the *Taq* fructose-1,6-bisphosphate aldolase and the deduced amino acid sequence of the enzyme.

Figure 2. Sequence alignment of class II FBP aldolases.

Figure 3. SDS-PAGE of the purification of *Taq* FBP aldolase produced in *E.coli* transformed by pKKaldoTaq.

Figure 4. Non-linear regression analysis of the kinetic data using the Hill relation.

Figure 5. Effect of the temperature on *Taq* FBP aldolase activity.

Figure 6. Photography of a *Taq* FBP aldolase crystal.

Figure 7. Diffraction pattern of a flash frozen monoclinic *Taq* aldolase crystal.

# Liste des sigles et abréviations

A: Adénine

ADN: Acide désoxyribonucléique

Amp: Ampicilline

ARN: Acide ribonucléique

C: Cytosine

DHAP: Dihydroxyacétone phosphate

DLS: Diffraction de lumière dynamique

DO: Densité optique

DTNB: Acide 5,5'-dithiobis(2-nitrobenzoïque)

DTT: 1,4-dithiothréitol

*E.coli*: *Escherichia coli*

EDTA: acide éthylènediamine tétraacétique

FBP: D-Fructose-1,6-bisphosphate

FPLC: Chromatographie rapide de protéines en phase liquide

G: Guanine

GAP: Glycéraldéhyde-3-phosphate

h: Coefficient de Hill

Hepes: Acide 4-(2-hydroxyéthyl)piperazine-1-éthanesulfonique

IPTG: Isopropyl- $\beta$ -D- thiogalactoside

kb: kilobase

kDa: kiloDalton

$K_M$ : Constante de Michaelis

MAD: Multiple wavelength Anomalous Dispersion

Mes: Acide 2-morpholinoéthanesulfonique monohydrate

MIR: Multiple Isomorphous Replacement

MPD: 2-méthyl-2,4-pentane-1,3-diol

PAGE: Électrophorèse en gel de polyacrylamide

PCR: Réaction de polymérisation en chaîne

PEG: Polyéthylène glycol

PGH: Phosphoglycolohydroxamate

PMSF: Fluorure de phénylméthylsulfonyl

SDS: Docécyll sulfate de sodium

*Taq*: *Thermus aquaticus*

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

U: Unité

$V_{MAX}$ : Vitesse maximale

## Remerciements

Merci à mon directeur de recherche, le Dr Jurgen Sygusch, de m'avoir permis de travailler sur ce projet dans son laboratoire.

Merci à la première équipe du laboratoire, Anne Pelletier, Danielle De Moissac, Nathalie Éthier, Steve Tétreault, Hérick Moniz et à la seconde, Jean-Pascal Viola, Vivian Saridakis, Christian Houde, Julie Allard, Christine Munger, Élisabeth Picard, Tae-Sung Yoon, Louis-David Boudrias et Pawel Grochulski, pour leur support, leurs judicieux conseils et leur précieuse aide concernant ce projet, de même que pour les diners et les cafés endiablés passés ensemble.

Merci à Nicole Varin-Dionne, Claudette Pocetti et Sylvie Beauchemin pour leur aide et leur oreille attentive.

Finalement merci à mes parents et à mon petit frère d'avoir toujours cru en moi et de m'avoir encouragé tout au long de mes études universitaires.

# **Chapitre 1: Introduction**

## 1.1 Les Fructose-1,6-bisphosphate aldolases

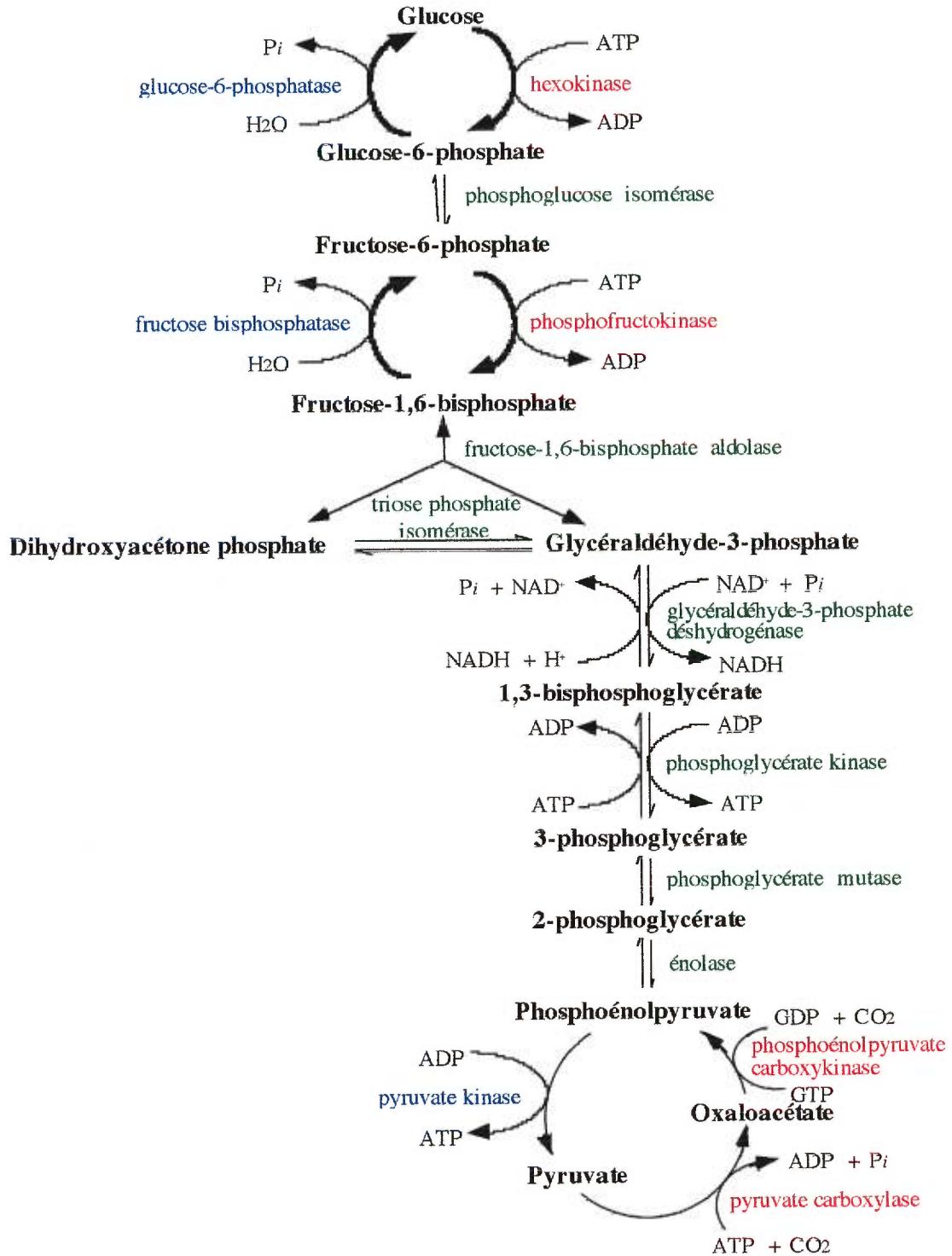
### 1.1.1 Un bref historique

En 1934 Meyerhof et Lohmann découvrirent qu'un enzyme qu'ils nommèrent *zymohexase*, permettait le clivage réversible d'une mole d'hexose bisphosphate en deux moles de dihydroxyacétone phosphate. Par la suite ils rebaptisèrent cet enzyme *aldolase* après avoir trouvé qu'il catalysait la condensation aldolique réversible du dihydroxyacétone phosphate avec une variété d'aldéhydes. Finalement en 1938 Meyerhof démontra que l'aldolase clivait selon une réaction uni-bi ordonnée, une mole de D-fructose-1,6-bisphosphate (FBP) pour produire une mole de dihydroxyacétone phosphate (DHAP) et une mole de D-glycéraldéhyde-3-phosphate (GAP)<sup>1</sup>. De nos jours cet enzyme est appelé la fructose-1,6-bisphosphate aldolase (EC 4.1.2.13; D-fructose-1,6-bisphosphate D-glycéraldéhyde-3-phosphate-lyase) étant donné la découverte de d'autres enzymes ayant également une activité aldolasique, entre autres la tagatose-1,6-bisphosphate aldolase, la 2-désoxyribose-5-phosphate aldolase et la 2-céto-3-désoxy-6-phosphogluconate aldolase.

### 1.1.2 Les caractéristiques générales des FBP aldolases

Le clivage du FBP en GAP et DHAP par la FBP aldolase prend place dans la voie de la glycolyse Embden-Meyerhof-Parnas (figure 1), tandis que la condensation du GAP et DHAP en FBP par le même enzyme, intervient dans le cycle de Calvin (figure 2) et la gluconéogenèse (figure 1).

Il existe deux classes d'aldolase différant par leur distribution à travers les espèces, leur mécanisme et leurs propriétés biochimiques, structurales et fonctionnelles. Leur manque d'homologie de séquence ( $\approx 15\%$  entre les classes I et II), de même que la localisation de leur site actif et de son contenu différent en acides aminés<sup>2,3</sup> démontrent une évolution distincte<sup>4</sup>.



**Figure 1. La glycolyse et la gluconéogenèse.**

Les enzymes partagés par les deux voies métaboliques sont en vert. Les enzymes exclusifs à la glycolyse sont en rouge et ceux exclusifs à la gluconéogenèse sont en bleu.



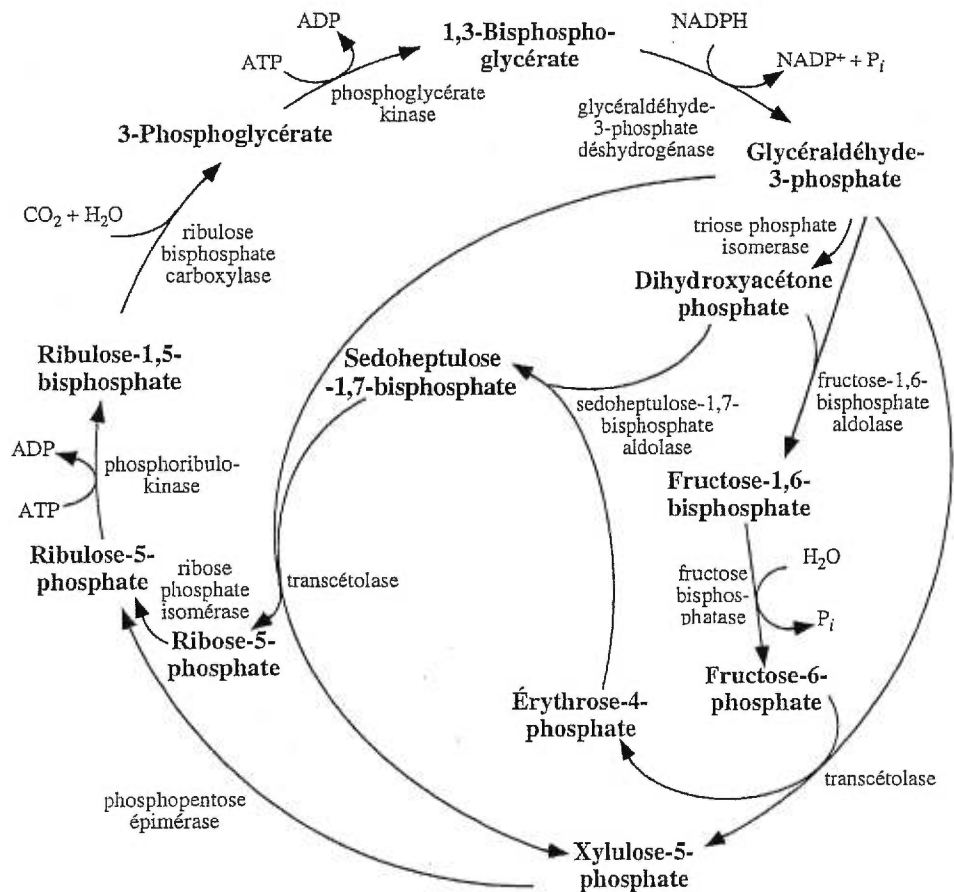


Figure 2. Le cycle de Calvin

### 1.1.3 Les FBP aldolases de Classe I

La FBP aldolase de classe I est présente chez les animaux, les plantes supérieures, les algues vertes, les métazoaires, les protozoaires et chez certains procaryotes.

Son mécanisme d'action est caractérisé par la formation d'un intermédiaire base de Schiff entre le  $\epsilon$ -amine d'une lysine du site actif et le substrat (la partie DHAP du FBP). Cette étape du mécanisme a été mise en lumière par l'inactivation de l'enzyme lors de sa réduction par le  $\text{NaBH}_4$  en présence de substrat (FDP ou DHAP). Les FBP aldolases de classe I ont le même profil de pH large pour le clivage du FBP que pour l'échange d'hydrogène du DHAP. De plus leur activité aldolasique n'est pas affectée par l'ajout d'agents chélateurs (EDTA) ou d'ions mono ou bivalents<sup>4</sup>.

### 1.1.3.1 Les FBP aldolases de Classe I des eucaryotes

#### 1.1.3.1.1 Leur distribution

Chez les vertébrés se retrouvent trois isoformes de FBP aldolase ayant chacun leurs propres propriétés catalytiques et distribution tissulaire. L'isoforme A est la seule forme présente dans le muscle et se retrouve dans la majorité des tissus. L'isoforme B est restreint au foie et aux reins où il tient un rôle important dans la gluconéogenèse et le métabolisme du fructose vue sa grande affinité pour le fructose-1-phosphate. L'isoforme C est présent dans plusieurs tissus notamment le cerveau. Dans les tissus contenant plus d'un isoforme, des formes hybrides sont retrouvées<sup>5</sup>.

Chez les plantes supérieures il existe aussi deux isoenzymes de la FBP aldolase de classe I: un d'origine plastidique et l'autre cytosolique<sup>6,7</sup>.

Certaines algues vertes ont soit de très faibles activités ou sont carrément déficientes pour quelques ou tous les isoenzymes cytosoliques des métabolismes des sucres phosphatés<sup>8</sup>. C'est le cas de *Chlamydomonas reinhardtii*, une algue verte chlorophyte, qui n'a qu'une seule FBP aldolase retrouvée dans ses chloroplastes<sup>9</sup>.

*Trypanosoma brucei*, un protozoaire parasite responsable de la maladie du sommeil, n'a qu'une FBP aldolase de classe I dans ses glycosomes<sup>10</sup>. *Plasmodium falciparum*, un sporozoïte responsable du paludisme, contient uniquement l'enzyme de classe I dans son cytosol<sup>11</sup>.

Certaines espèces eucaryotes contiennent les 2 classes de FBP aldolases. L'algue verte unicellulaire *Euglena gracilis* a une FBP aldolase de classe I dans ses chloroplastes (cycle de Calvin) et une de classe II dans son cytosol (glycolyse, gluconéogenèse)<sup>12</sup>. *Chlamydomonas mundana*, une algue verte, de même que d'autres algues, rouges et brunes<sup>4</sup>, ne produisent que la FBP aldolase de classe I dans des conditions de croissance autotrophes où le CO<sub>2</sub> sert de source de carbone et où le cycle obscur de la photosynthèse (cycle de Calvin) est sollicité. Dans des conditions hétérotrophes où l'acétate est utilisé comme source de carbone, l'enzyme de classe II est produit et utilisé via la voie de la gluconéogenèse et le cycle glyoxylate (photoassimilation)<sup>13</sup>.

### 1.1.3.1.2 Leur structure tridimensionnelle

La FBP aldolase de muscle squelettique de lapin est le principal modèle d'étude des FBP aldolases de classe I. Sa structure a été résolue par cristallographie. Comme toutes les FBP aldolases des eucaryotes, la FBP aldolase de muscle de lapin est tétramérique<sup>4</sup>. Chaque sous-unité d'environ 40 kDa est constituée d'un baril  $\beta$  simple de la famille des barils  $(\alpha/\beta)_8$  parallèles.

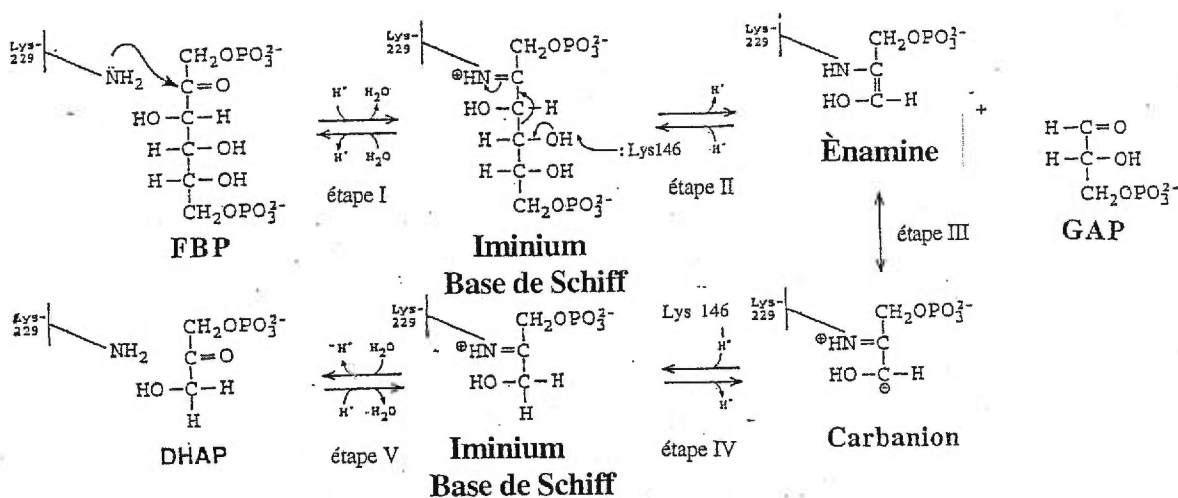
Le site actif situé au centre du baril forme une sorte de poche accessible du côté C-terminal des feuillets  $\beta$ . Cette localisation du site actif fait en sorte que l'intérieur du baril contient non seulement des résidus hydrophobes mais aussi des résidus potentiellement chargés (trois lysines (107, 146 et 229) dont la lysine 229 impliquée dans l'intermédiaire base de Schiff, l'acide aspartique 33 et l'acide glutamique 187)<sup>2</sup>. Les groupements acides et basiques de ces acides aminés sont disposés colinéairement en alternance de façon à neutraliser les charges et à permettre le repliement de ces résidus à l'intérieur du baril.

Le repliement général de la protéine est semblable dans les quatre sous-unités excepté pour la partie C-terminale. La région C-terminale forme une structure similaire à un bras ayant un coude flexible. Dans les sous-unités 1 et 2, le C-terminal se replie à la surface de sa propre sous-unité couvrant ainsi la poche du site actif. Dans les sous-unités 3 et 4, le C-terminal se replie sur sa surface en direction de son propre site actif avant de bifurquer vers une fente située entre les sous-unités de l'homotétramère, laissant ainsi son site actif totalement exposé. Cette conformation de la région C-terminale médie l'attachement du produit DHAP, contrôlant ainsi la formation de la base de Schiff et par conséquent, l'activité de l'enzyme<sup>14</sup>.

Les résidus hydrophobes et chargés du site actif sont conservés à travers les espèces. Ce n'est toutefois pas le cas pour la région C-terminale dont l'hétérogénéité contribue aux différences d'activité et de spécificité de chaque enzyme pour divers substrats<sup>2</sup>.

### 1.1.3.1.3 Leur mécanisme réactionnel

Le mécanisme général proposé pour la FBP aldolase de muscle de lapin, représentante des enzymes de classe I, est le suivant (figure 3):



**Figure 3. Le mécanisme réactionnel proposé pour la FBP aldolase de classe I de muscle de lapin (d'après Morris, A. J. & Tolan, D. R.<sup>15</sup>)**

Le groupement  $\epsilon$ -amine de la lysine 229 du site actif viendrait attaquer le groupement carbonyle  $C_2$  de la forme ouverte du FBP, formant ainsi un intermédiaire iminium (base de Schiff) tout en libérant une molécule d'eau (étape I). Une base (la lysine 146 précédemment déprotonnée par un acide aspartique) extrairait le proton de l'hydroxyle du carbone 4 du FBP. Les électrons se dirigeraient vers le puits à électrons en  $C_2$ . Il y aurait alors clivage du lien entre le carbone 3 et 4 du FBP. Le GAP serait libéré et le DHAP resterait attaché à l'enzyme sous forme d'une ènamine (étape II). L'ènamine se retrouverait dans un système de résonance entre sa forme ènamine et carbanion (étape III). Le carbanion en position  $C_3$  de la portion restante DHAP se protonerait en déprotonant l'amine de la lysine 146, pour former un ion immonium (étape IV). L'hydrolyse de l'ion iminium libérerait le DHAP de l'enzyme en permettant à la lysine de réitérer sa forme initiale non-liée (étape V)<sup>15</sup>.

Il a été démontré qu'une tyrosine de la région C-terminale joue un rôle essentiel dans la réaction aldolase des enzymes eucaryotes de classe I puisque le relâchement de ce résidu lors d'un traitement de l'enzyme à la carboxypeptidase A entraîne la perte de l'activité<sup>16</sup>. Les cystéines sont aussi impliquées dans le mécanisme de ces mêmes enzymes puisque leur modification par le DTNB inhibe également l'activité de l'enzyme<sup>17</sup>.

### 1.1.3.2 Les FBP aldolases de classe I des procaryotes

Des FBP aldolases de classe I ont été retrouvées chez plusieurs procaryotes. Certaines eubactéries comme *Mycobacterium smegmatis*<sup>18</sup>, *Staphylococcus carnosus*<sup>19</sup>, *Micrococcus aerogenes*<sup>20</sup> et certaines archéobactéries halophiles (*Halobacterium* R-113, *Halobacterium* CH-1, *H.saccharovorum*, *H.vallismortis*<sup>21</sup>) n'ont qu'une FBP aldolase de classe I.

Le plus souvent les FBP aldolases de classe I sont exprimées lorsque l'organisme procaryote croît dans des conditions nécessitant la synthèse d'hexoses via la voie de la gluconéogenèse<sup>7</sup>. Chez *E.coli* par exemple, la production de FBP de classe I est favorisée lorsque la bactérie prolifère dans un milieu composé de lactate ou de pyruvate et doit alors recourir à la gluconéogenèse. Dans des conditions de croissance contenant du glucose et où la voie de la glycolyse est sollicitée, la FBP aldolase de classe II est grandement favorisée. Ce phénomène s'apparente un peu à celui des vertébrés dont les muscles qui emploient la voie glycolytique contiennent l'isoforme A alors que le foie, siège principal de la gluconéogenèse contient l'isoforme B<sup>22</sup>.

Toutefois *Xanthobacter flavus*, une bactérie gram négatif, synthétise une FBP aldolase de classe I dans des conditions hétérotrophiques et l'enzyme de classe II dans des conditions autotrophiques nécessitant le recours au cycle de Calvin. Cette spécificité est contraire à celle de *Chlamydomonas mundana*, d'*Euglena gracilis* et des plantes supérieures qui utilisent un enzyme de classe I pour la voie métabolique du cycle de Calvin<sup>101</sup>.

Les FBP aldolases de classe I des procaryotes sont assez peu similaires à celles des eucaryotes<sup>4</sup>. Le poids moléculaire de leurs sous-unités varie entre 27 et 40 kDa<sup>23</sup>.

De plus, contrairement à leurs pendants eucaryotes, elles ont une structure quaternaire très variable qui va du monomère (*Staphylococcus aureus*)<sup>24</sup> au décamère (*E.coli*)<sup>23</sup>.

Une autre caractéristique majeure des FBP aldolases de classe I des procaryotes est leur insensibilité aux traitements à la carboxypeptidase A et aux réactifs sulfhydryles<sup>20</sup>. Leur mécanisme comprend donc la formation d'un intermédiaire base de Schiff mais diffère de celui des eucaryotes, quant au rôle des tyrosines et des cystéines dans leur activité. Seules des études de structure/fonction plus approfondies permettront de déterminer les différences entre les enzymes de classe I des eucaryotes et des procaryotes.

### 1.1.4 Les FBP aldolases de classe II

La fructose-1,6-bisphosphate aldolase de classe II est présente chez les bactéries, les champignons et les levures<sup>4</sup>. L'algue *Cyanophora paradoxa* contient même deux FBP aldolases de classe II, une dans son cytosol et l'autre dans ses cyanoplastes (qui viendraient des cyanobactéries)<sup>25</sup>.

Les FBP aldolases de classe II se différencient de celles de classe I par leur mécanisme réactionnel. Celles de classe II ne forment pas un intermédiaire base de Schiff puisque l'ajout de  $\text{NaBH}_4$  en présence du substrat<sup>4</sup>, de même que le traitement de l'enzyme à la carboxypeptidase A<sup>26</sup> ne compromettent pas leur activité aldolasique. Le mécanisme de clivage des enzymes de classe II utilise plutôt un ion métallique divalent (un ion par sous-unité) comme électrophile<sup>4</sup>. Ceci est confirmé par l'inhibition de l'activité enzymatique en présence d'agents chélateurs comme l'EDTA. Leur activité peut aussi être augmentée par l'ajout de cations monovalents ( $\text{K}^+$ ,  $\text{Na}^+$  ou  $\text{NH}_4^+$ )<sup>1,4</sup>. Les FBP aldolases de classe II ont deux profils de pH étroits décalés d'environ une unité de pH pour le clivage du FBP et l'échange d'hydrogène du DHAP<sup>4</sup>. De plus leur activité spécifique ( $V_{\text{MAX}}$ ) et leur  $K_M$  pour le fructose-1,6-bisphosphate sont plus élevés que ceux des aldolases de classe I. Pour des substrats analogues au FBP, leur vitesse est beaucoup plus affectée que les classes I, voir même pratiquement nulle pour le clivage du fructose-1-phosphate<sup>26</sup>.

Plusieurs duplications des gènes de FBP aldolases ont eu lieu durant l'évolution eubactérienne. Une duplication impliquant un ancêtre commun des bactéries Gram positif et des protéobactéries mena à la séparation des FBP aldolases de classe II en 2 types: A et B<sup>12,27</sup>.

Les FBP aldolases de classe II A forment un groupe homogène comprenant les enzymes de la glycolyse des levures et des champignons, des eubactéries Gram positif à haut contenu en G/C (*Mycobacterium* et *Corynebacterium*), des protéobactéries  $\epsilon$  et  $\gamma$  (*E.coli* et *Campylobacter*) et l'enzyme cytosolique d'*Euglenia gracilis*. Chacune de leurs sous-unités a un poids moléculaire d'environ 37 à 40 kDa.

Les FBP aldolases de classe II B constituent un groupe hétérogène comprenant les enzymes de la glycolyse des bactéries Gram positif à faible contenu en G/C (*Bacillus* et *Mycoplasma*) (environ 31 kDa/sous-unité) et d'un groupe de bactéries extrémophiles (*Helicobacter pylori*, *Aquifex aeolicus* et *Thermotoga maritima*) (33-35 kDa/sous-unité), les enzymes du cycle de

Calvin des cyanobactéries et des protéobactéries  $\alpha$  et  $\beta$  (*Synechocystis*, *Rhodobacter* et *Alcaligenes*) (37-39 kDa/sous-unité) et enfin, d'un groupe d'enzymes protéobactériens agissant sur la tagatose-1,6-bisphosphate ou dont la fonction n'est pas encore bien définie (AGAY *E.coli* et *Vibrio furnissii*) (31kDa/sous-unité)<sup>27</sup>.

#### 1.1.4.1 Leur structure tridimensionnelle

Jusqu'à maintenant tous les enzymes de classe II caractérisés sont dimériques excepté celui de *Thermus aquaticus*<sup>28,29</sup>, *Anacystis nidulans* et *Saprospira thermalis*<sup>30</sup>. La seule structure de FBP aldolase de classe II résolue à ce jour est celle de l'homodimère d'*E.coli*<sup>3</sup>. Chacune de ses sous-unités adopte la même conformation soit un baril  $(\alpha/\beta)_8$  simple avec un site actif localisé du côté carboxy-terminal des feuillets  $\beta$ . Le baril  $\beta$  est fermé de son côté N-terminal par une hélice  $\alpha$  et les motifs  $\alpha/\beta$  répétés sont reliés entre eux par des boucles de longueurs variées. Des interactions hydrophobiques permettent la formation d'une structure supersecondaire composée de deux longues hélices  $\alpha$  antiparallèles dans la région C-terminale de l'enzyme. Une forte interaction hydrophobique entre cette paire d'hélices de chaque sous-unité dont les résidus contribuent au site actif de l'autre sous-unité, constitue l'élément majeur responsable de l'homodimérisation. En plus de cette "poignée de main hydrophobique", quelques liens hydrogène entre des hélices  $\alpha$  adjacentes contribuent à l'association des sous-unités.

Chaque sous-unité comporte deux sites de liaison au zinc mutuellement exclusifs, mais structurellement reliés, soit un site enfoui dans le site actif et un second près de la surface de la sous-unité. La coordination tétraédrale du zinc est effectuée par trois histidines (110, 226 et 264) et un acide glutamique (174). L'exclusivité de la position du zinc vient du changement conformationnel induit par la rotation des chaînes latérales des histidines. La composition du site actif laisse croire que la séquestration du zinc dans le site enfoui servirait au maintien de l'intégrité du site actif lors du recrutement du substrat tandis que le site en surface serait utilisé lors de la catalyse.

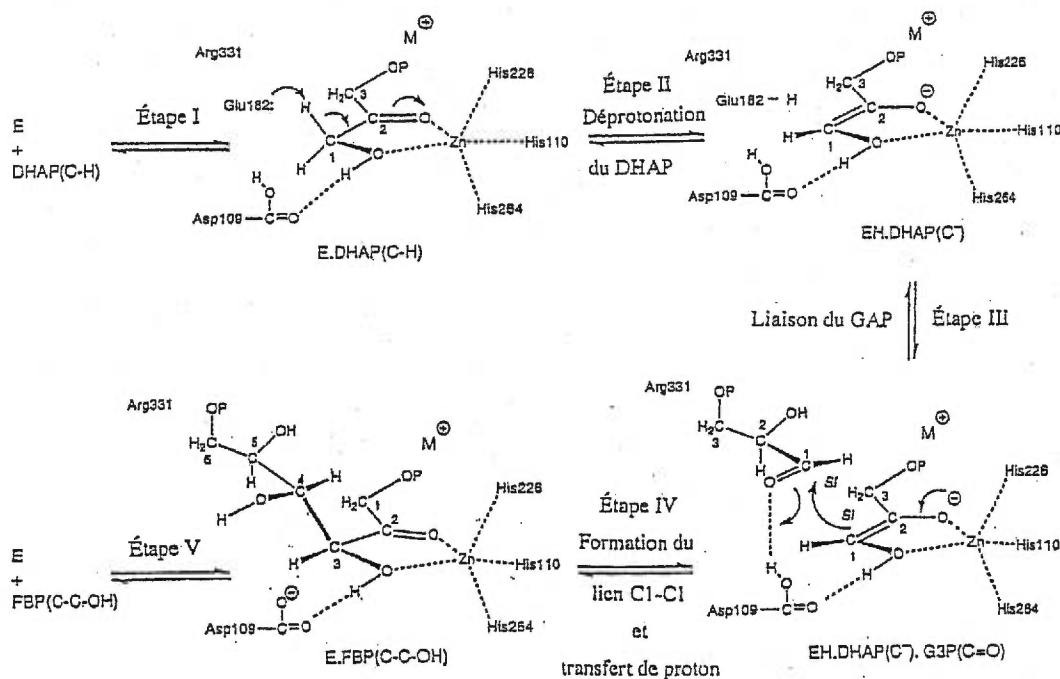
Au centre de chaque baril  $\beta$  se trouve un site de liaison pour un cation monovalent (probablement  $K^+$ ,  $Na^+$  ou  $NH_4^+$ ) constitué de l'histidine 107, l'acide glutamique 172, la lysine 284, l'asparagine 286 et l'acide aspartique 109<sup>3</sup>.

### 1.1.4.2 Leur mécanisme réactionnel

Le mécanisme réactionnel des FBP aldolases de classe II a jusqu'alors été assez peu étudié comparativement à celui de la FBP aldolase de muscle de lapin. Toutefois récemment des essais enzymatiques et des données cristallographiques<sup>3,31</sup> dont une structure de la FBP aldolase de *E.coli* en présence du PGH, un analogue de l'état de transition ène-diolate<sup>32</sup>, ont permis d'élucider un peu ce mécanisme (figure 4). D'après ces études, le zinc occuperait son site de liaison en surface lors de la liaison du substrat. Le DHAP lierait l'enzyme simultanément ou tout de suite après la liaison du cation monovalent, ce qui amènerait un réarrangement de trois boucles. Le DHAP chelaterait le zinc qui perdrait ces liaisons avec deux molécules d'eau. Une de ces molécules d'eau viendrait protoner l'acide aspartique 109 (étape I). Le zinc agirait comme acide de Lewis pour polariser le lien carbonyle de la cétone du substrat et le préparer à la condensation. La polarisation d'un groupement carbonyle cétonique faciliterait l'abstraction d'un proton du DHAP par une base (possiblement l'acide glutamique 182) pour former l'intermédiaire carbanion ène-diolate et produire une liaison insaturée à l'emplacement de l'addition (étape II). Le GAP viendrait lier l'arginine 331 afin d'être amené en position pour interagir avec l'ène-diolate (étape III). Le carbonyle C<sub>1</sub> du carbanion stabilisé électrostatiquement par l'asparagine 286, attacherait le carbonyle C<sub>1</sub> du GAP pour générer un nouveau lien carbone-carbone. L'acide aspartique 109, responsable de la polarisation du groupement carbonyle du GAP donnerait un proton pour stabiliser la charge développée (étape IV). Le FBP nouvellement formé quitterait ensuite le site actif de l'enzyme (étape V).

L'acide aspartique 109 améliorerait catalytiquement l'étape la plus lente de la réaction c'est-à-dire la formation du lien carbone-carbone. Les données cristallographiques<sup>3</sup> ont démontré que cet acide aspartique fait partie du site de liaison du cation monovalent qui augmente l'activité de l'enzyme. Il est probable que la présence de cet ion aide à positionner correctement la chaîne latérale de cet acide aspartique de manière à maximiser la polarisation du carbonyle du GAP<sup>33</sup>.





**Figure 4. Le mécanisme réactionnel proposé pour la FBP aldolase de classe II d'*E. coli* (d'après Hall, D. R. & al.<sup>32</sup>)**

Toutefois, les données cristallographiques obtenues dans notre laboratoire portent à croire que le zinc serait dans le site enfoui lors de la liaison du substrat à l'enzyme d'*E. coli* (données non-publiées). La détermination complète de la structure tridimensionnelle permettra de déterminer quelles modifications devront être apportées au mécanisme mentionné ci-haut.

### 1.1.5 Sa phylogénie

L'analyse phylogénétique des FBP aldolases de classe I suggère trois sous-classes, soit celle des eucaryotes, celle des procaryotes tels *Staphylococcus aureus* et *Synechocystis* et celle des procaryotes comme *E. coli*<sup>23</sup>.

La capacité d'exprimer les deux classes de FBP aldolases vient probablement d'organismes ancestraux qui ont donné naissance à ceux existants aujourd'hui. Les formes de vie contemporaines n'expriment exclusivement que la FBP aldolase de classe I ou de classe II, démontrent l'élimination de la redondance enzymatique et l'évolution indépendante des deux classes. L'usage exclusif de l'une ou l'autre des classes de FBP aldolases chez ces organismes montre bien que les deux classes ont des fonctions catalytiques similaires<sup>34</sup>.

## 1.2 Les thermophiles

La température est un des facteurs environnementaux les plus déterminants dans le contrôle de l'activité, la croissance, la reproduction et l'évolution des organismes vivants. Les environnements à hautes températures sont particulièrement intéressants de par leur représentation d'une limite extrême de l'évolution. Par exemple, ceux associés à l'activité volcanique (sources thermales) sont particulièrement fascinants puisqu'ils constituaient déjà des habitats naturels à l'époque où les organismes vivants sont apparus sur la Terre. Déjà dans l'Antiquité, la présence de vie dans les sources d'eau thermale avait été remarquée.

Ces environnements chauds abritent une grande diversité d'organismes parfaitement adaptés: eucaryotes, eubactéries, archéobactéries autant aérobies qu'anaérobies, autotrophes que hétérotrophes. Ces organismes vivants sont qualifiés d'hyperthermophiles ou de thermophiles extrêmes s'ils prolifèrent à une température de croissance optimale de 80°C et plus (avec une température de croissance maximale de 90°C et plus) et thermophiles modérés si leur température de croissance optimale se situe entre 55°C et 80°C<sup>35</sup>. Aucun organisme multicellulaire ne vit à plus de 50°C. Les eucaryotes thermophiles tels les levures, les champignons ou les algues, ont une température de croissance maximale variant entre 50 et 60°C. Seules les eubactéries et les archéobactéries peuvent croître entre 60 et 90°C; c'est pourquoi les études sur les thermophiles portent essentiellement sur les procaryotes. Dans le texte suivant, la mention d'organismes thermophiles réfèrera aux thermophiles procaryotiques. Les microorganismes photosynthétiques ne vivent pas à plus de 73°C. Au-delà de 90°C seules les archéobactéries survivent, à l'exception de l'eubactérie *Thermotoga*<sup>36</sup>.

### 1.2.1 Leurs habitats

Les thermophiles se retrouvent dans divers habitats géothermiques naturels. Ces habitats se retrouvent partout à travers le monde, entre autres en Islande, en Nouvelle-Zélande, aux États-Unis et en Italie. Ils sont situés dans les zones tectoniques actives où des mouvements de la croûte terrestre ont permis au magma de venir très près de la surface de la Terre et ainsi servir de source de chaleur. L'eau de mer ou les eaux souterraines infiltrées dans

la terre sont alors réchauffées jusqu'à ce que la pression force le liquide à travers les pores et les fissures pour former à la surface de la terre, un geyser ou une source thermale. Lors du passage de l'eau chauffée vers la surface, les minéraux contenus dans les roches se dissolvent, ce qui explique la forte minéralisation des sources thermales et des geysers. D'autres constituants comme le sulfure d'hydrogène, le dioxyde de carbone, le méthane, l'hydrogène, l'ammoniaque, des composés organiques de faible poids moléculaire, le fer, le bicarbonate, le chlorure et d'autres éléments en trace peuvent aussi se dissoudre dans ce liquide chaud. La composition chimique de chaque source thermale dépend donc de la température de l'eau et de la nature des roches à travers lesquelles elle passe. Dans le cas d'un faible apport d'eau à la nappe souterraine, seules des vapeurs d'eau et de gaz s'échapperont sous forme de fumerolles. Les bassins constitués de geysers ont un important débit d'eau et une température très constante à leur source. Leur eau très minéralisée (très riche en anions chlorure et bicarbonate) est tamponnée par un système carbonate-bicarbonate qui maintient son pH entre 7,5 et 9,0.

D'autres bassins tirent leur source de l'eau des montagnes avoisinantes qui coule très lentement et est réchauffée lors de son contact avec des fumerolles. Elle contient alors très peu de minéraux (excepté pour le fer) et de l'acide sulfurique d'où son pH autour de 1,8-2,2. Sa température devient dépendante de la proximité des vapeurs réchauffantes. Dans ces sources acides les bactéries croissent librement dans l'eau tandis que dans les sources neutres, les bactéries s'attachent aux parois silicieuses de la source.

Les sources hydrothermales (abyssales) des fonds marins constituent aussi un habitat idéal pour les thermophiles. Elles fonctionnent selon le même principe que les geysers excepté pour l'eau de la nappe qui est de l'eau de mer au lieu de l'eau douce. À cause de sa salinité, l'eau de mer chaude est très réactive et devient alors très minéralisée<sup>37</sup>.

## 1.2.2 Leurs adaptations à la chaleur

La stabilité des macromolécules (ADN, protéines, lipides, etc) est souvent considérée comme le facteur déterminant la plus haute température permettant la vie. Or quelques

changements mineurs dans la composition en acides aminés des protéines peuvent les rendre plus stables, les acides nucléiques peuvent être stabilisés par des forces ioniques modérées, des histones ou de polyamines<sup>38</sup> et les lipides ne sont pas particulièrement thermolabiles<sup>39</sup>. Les facteurs qui limitent la vie à hautes températures sont donc autres. Le besoin d'eau libre à hautes températures peut être comblé grâce à des pressions modérées. Dans le cas d'une théorie évolutive voulant que les thermophiles viennent des mésophiles, le temps devient un facteur limitant surtout que les mutations aléatoires ne semblent pas particulièrement tendre vers une augmentation de la thermostabilité. La limite de vie à hautes températures serait plutôt déterminée par la stabilité de petites molécules essentielles comme les acides aminés et les cofacteurs enzymatiques (NADH par exemple) puisqu'il n'y a jusqu'à maintenant aucune preuve de leur stabilisation par de grandes pressions ou par leur association avec de plus grosses molécules. Or la biochimie connue jusqu'à présent est constituée de voies métaboliques comportant des interconversions contrôlées et précises de ces petites molécules. À moins de l'existence d'une biochimie encore inconnue ou de la sous-estimation de la stabilité de ces petites molécules, la vie serait difficile à plus de 150 °C<sup>40</sup>.

L'origine des thermophiles et leur relation avec les mésophiles n'est pas complètement connue. Évidemment la transition thermophile-mésophile ne peut être attribuée à une mutation spontanée dans une seule protéine ou macromolécule. Un nombre important de mutations est indispensable pour l'adaptation et la stabilisation à d'aussi hautes températures. Les études sur les organismes thermophiles dans leur habitat naturel ont démontré que toutes leurs macromolécules étaient adaptées optimalement à la température à laquelle ils vivaient. Du point de vue évolutif, ceci indique que ces organismes ne sont pas des organismes qui ont su s'adapter graduellement aux hautes températures, mais bel et bien des organismes déjà adaptés de façon unique à ces températures<sup>37</sup>.

Toutefois les organismes vivant dans des conditions extrêmes ont mis de côté leur efficacité de croissance au profit de leur survie. Même si certains thermophiles prolifèrent plus vite que les mésophiles, cette augmentation est moindre que celle attendue par l'équation d'Arrhénius. Leur courbe d'Arrhénius pour leur croissance n'est pas linéaire mais courbe. La relation linéaire décrit plutôt la relation entre la racine carrée de la vitesse de croissance et la

température absolue<sup>41</sup>. Les thermophiles même s'ils s'ont optimalement adaptés à leur environnement, ne sont pas capables d'en utiliser pleinement le potentiel thermal. C'est la thermostabilité de leurs composants cellulaires et non l'augmentation de leur synthèse qui est responsable de l'augmentation de leur température optimale de croissance<sup>35</sup>.

Voici décrites les adaptations à la chaleur notées pour les différentes macromolécules des thermophiles.

### 1.2.2.1 Leurs lipides et leurs membranes

La membrane cytoplasmique est un élément physiologique essentiel aux cellules puisqu'elle contrôle le trafic de différentes substances entre l'environnement intra- et extracellulaire en plus de délimiter ces deux environnements. Elle est composée d'ions, de protéines implantées dans la membrane ou attachées à celle-ci, et principalement de lipides. Ces derniers lui permettent de maintenir une fluidité appropriée pour le bon fonctionnement des protéines membranaires et ce, malgré une variété de changements extérieurs dont la température.

Les membranes des thermophiles doivent donc être constituées de lipides qui permettent cette fluidité à hautes températures sans mettre en péril l'intégrité des membranes elles-mêmes. La membrane des eubactéries thermophiles est conforme au modèle général de bicouche lipidique. La membrane des archéobactéries thermophiles contient des tétraéthers qui peuvent s'étendre de la face interne à la face externe de la membrane et former structurellement l'équivalent d'une monocouche amphiphilique. Chez les deux types de bactéries thermophiles la membrane est composée d'acides gras à haut point de fusion. La majorité de leurs chaînes apolaires possède un groupement méthyle qui interviendrait dans la fluidité membranaire. La plupart de leurs lipides polaires contiennent des carbohydrates ou des groupements phosphate. Une forte concentration de glycolipides et le grand nombre de sucres présents sur ces glycoprotéines rendent la bicouche phospholipidique plus rigide. Les protéines membranaires des thermophiles sont aussi plus stables que celles des mésophiles<sup>41</sup>. Des travaux ont même démontré que la composition des membranes était modifiée selon la température de croissance de l'organisme et ce même à de hautes températures<sup>39</sup>.

### 1.2.2.2. Leurs acides nucléiques et leurs ribosomes

L'ADN, même celui des mésophiles, a une grande thermostabilité particulièrement dans le milieu ionique dans lequel il se retrouve *in vivo*. L'ADN des thermophiles n'a donc pas besoin de beaucoup de stabilisation pour rester intact à hautes températures. De façon générale son contenu en guanine (G) et cytosine (C) est plus élevé que celui des mésophiles. Étant donné que les guanines et les cytosines sont pairées par trois liens hydrogène comparativement à deux pour les thymines et les adénines, leur plus haut pourcentage chez les thermophiles fait augmenter la température de fusion de l'ADN. Les deux brins de l'ADN restent donc appariés même à haute température.

Les ARN de transfert (ARNt) sont habituellement très stables. Pour augmenter leur stabilité chez les thermophiles, ils contiennent plus de paires de bases G/C. Des méthylations augmentent aussi leur résistance à la chaleur et aux ribonucléases. De plus, des thiolations viennent renforcer l'interaction des bases de l'ARNt hélicoïdal. Le nombre de ce genre de substitutions et la température de fusion de l'ARNt augmentent avec la température de croissance du thermophile.

Les ribosomes des thermophiles sont aussi plus stables que ceux des mésophiles. Cette thermostabilité vient à la fois de l'ARN et des protéines ribosomiales. Les protéines du ribosome des thermophiles ont les mêmes adaptations à la chaleur que les autres protéines thermostables (voir le paragraphe suivant 1.2.2.3). Quant aux ARN ribosomiaux, ils sont plus stables à cause de leur plus hauts taux en G/C et grâce à leurs interactions spécifiques avec les protéines ribosomiales. L'une des boucles de la structure secondaire de l'ARN 5S des thermophiles contient quatre bases au lieu de trois chez les mésophiles. Une grande diversité de polyamines (12 au lieu de 2) aux chaînes plus longues que chez les mésophiles stabiliserait les ribosomes thermostables *in vivo* en aidant l'association des sous-unités ribosomiales et faciliterait la synthèse protéique à hautes températures<sup>38,41</sup>.

### 1.2.2.3 Leurs protéines

La caractérisation de l'adaptation des protéines à hautes et basses températures fait partie intégrale de notre compréhension du repliement des protéines, de la relation structure/fonction, de l'évolution de la vie sur Terre et permettra le design de biocatalyseurs. La thermostabilité d'une protéine peut être changée intrinsèquement par l'altération de ses acides aminés ou extrinsèquement par l'addition de molécules stabilisatrices (cations, coenzymes, membranes, peptides). La stabilité des protéines thermophiles est en grande partie ou totalement intrinsèque à leur séquence en acides aminés et à leur structure secondaire, tertiaire et quaternaire puisque ces protéines sont encore thermostables une fois purifiées<sup>40</sup>. Le remplacement (turnover) des protéines chez les thermophiles est faible durant leur croissance<sup>41</sup>. De plus la stabilité de leurs protéines est optimisée selon la température de croissance de l'organisme<sup>40</sup>. Cependant la conséquence directe de l'augmentation de la thermostabilité intrinsèque des enzymes thermophiles est l'adaptation de leur activité catalytique à de telles conditions. En tenant compte de la faible augmentation du taux de croissance des thermophiles par rapport aux mésophiles, l'activité spécifique des enzymes thermophiles peut même être considérée moindre que celle attendue par extrapolation de l'activité des mésophiles à de hautes températures.

La stabilité d'une protéine résulte d'un équilibre entre de grandes forces déstabilisatrices (1000 kJ/mol) dues à l'entropie de la chaîne peptidique et des forces stabilisatrices légèrement plus grandes dues à la somme des interactions électrostatiques, hydrophobiques et aux liens hydrogènes. L'énergie libre de stabilisation nette de la protéine est généralement de l'ordre de 30 kJ/mol. Donc de petits changements des forces stabilisatrices peuvent avoir des effets dramatiques sur toute la stabilité protéique. L'ajout de quelques ponts salins, d'interactions hydrophobes ou de liens hydrogènes peut alors facilement contribuer à l'augmentation de l'énergie libre de stabilisation nécessaire pour élever la stabilité relative de l'enzyme<sup>40,42</sup>. Par exemple, le remplacement ponctuel d'une arginine par une histidine dans le lysozyme, quoique ne provoquant aucun changement structural détectable par cristallographie, a descendu de 14°C la température de dénaturation de l'enzyme<sup>43</sup>. La comparaison de séquences entre une kanamycine nucléotidyltransférase d'un mésophile et d'un thermophile a

démontré une différence d'une seule base. Une thréonine était remplacée par une lysine ce qui permettait la formation d'un pont salin avec un acide aminé chargé négativement. Cette unique substitution augmentait la thermostabilité de l'enzyme toujours sans modifier sa structure tridimensionnelle<sup>44</sup>.

Les enzymes thermophiles en plus de résister à la chaleur, sont généralement relativement stables en présence d'agents dénaturants tels que des pH extrêmes, des concentrations élevées de sels, des détergents et des agents chaotiques comme les solvants organiques<sup>45</sup>. Ces agents contribuent habituellement à diminuer les forces stabilisantes et/ou à augmenter les forces déstabilisatrices, ce qui entraîne une réduction nette de la stabilité. Or les interactions additionnelles des protéines thermostables les rendent aussi plus stables en présence de ces agents. Cette explication de la thermostabilité par une plus grande stabilité thermodynamique générale a été observée chez la RNase H de *Thermus thermophilus* HB8<sup>46</sup> et la phosphoglycérate kinase de *Thermotoga maritima*<sup>47</sup>, mais ne fait pas l'unanimité.

Deux autres hypothèses ont été avancées pour expliquer la thermostabilité des protéines thermophiles. La première veut que la résistance des protéines thermophiles ne soit pas due à une stabilité supérieure à celle des mésophiles, mais plutôt au fait que leur énergie libre de repliement soit moins sensible à la température. Leur profil d'énergie libre serait donc plus aplati que celui des mésophiles. Ce concept est supporté par des résultats obtenus avec de petites protéines liant l'ADN chez des organismes hyperthermophiles<sup>48,49</sup>. Une autre hypothèse veut que l'énergie libre de repliement des protéines thermophiles ne soit pas plus influencée par la température que celle des mésophiles mais que la courbe soit simplement déplacée à plus hautes températures. Jusqu'à ce jour cette dernière théorie n'a pas été démontrée expérimentalement. Ces trois types de thermostabilisation seraient utilisés chez les diverses protéines thermophiles retrouvées dans la nature<sup>50</sup>.

### 1.2.2.3.1 Une augmentation de la rigidité

Une certaine mobilité moléculaire est critique pour le bon fonctionnement d'une protéine, particulièrement pour lui permettre d'interagir allostériquement avec de petites molécules<sup>35</sup>. Parallèlement une certaine rigidité est nécessaire au maintien de la structure



protéique<sup>51</sup>. À leur température optimale, les protéines thermophiles sont pratiquement aussi flexibles que les homologues mésophiles à la leur (25-37°C). Toutefois cette flexibilité diminue avec la température, ce qui les rend pratiquement inopérantes aux températures optimales pour les mésophiles<sup>40,52</sup>. Les thermophiles doivent donc sacrifier le taux d'activité de leurs enzymes au profit de leur croissance à hautes températures<sup>35</sup>. Cette adaptation souvent caractérisée par un ajout de prolines ou de boucles en surface plus courtes<sup>67</sup>, rend les enzymes thermophiles plus résistants à la protéolyse puisqu'à la température optimale de lyse (37°C) leur conformation rigide diminue l'accès des protéases<sup>40</sup>.

### 1.2.2.3.2 La substitution de certains acides aminés

Diverses études comparatives ont été effectuées entre des séquences protéiques thermophiles et mésophiles. Les protéines thermophiles ont un nombre plus restreint d'asparagines et de glutamines qui sont susceptibles aux désaminations. Ces résidus sont généralement éliminés lors du raccourcissement des boucles ou remplacés par des résidus chargés pour former de nouvelles paires d'ions<sup>67</sup>. Les méthionines et surtout les cystéines sont éliminées pour éviter leur oxydation à hautes températures. Les cystéines et méthionines restantes sont généralement enfouies à l'intérieur de la protéine afin d'être protégées des molécules d'eau entourant la protéine et qui pourraient les modifier<sup>67</sup>.

Une diminution des résidus polaires non-chargés (sérine, asparagine, glutamine, thréonine, méthionine) se fait au profit de résidus hydrophobes pour augmenter le centre hydrophobe, diminuer l'accès du solvant à l'intérieur et ainsi aider à prévenir les désaminations et le clivage de la chaîne principale, ou encore au profit de résidus chargés pour augmenter les ponts salins tout en gardant la possibilité de faire des ponts hydrogène. La substitution de résidus hydrophobes pour d'autres plus hydrophobes, augmente la densité d'empilement tout en gardant l'intégrité du centre hydrophobe. Le remplacement de résidus par d'autres de plus gros volume, permet d'exclure l'eau de l'intérieur de la protéine, de remplir les cavités et de réduire la liberté entropique du squelette de la protéine non-repliée<sup>53</sup>.

Argos a déterminé 5 échanges d'acides aminés fréquemment retrouvés chez les protéines thermostables soit la substitution des acides glutamiques en alanines, des sérines en alanines,

des sérines en thréonines, des lysines en arginines et des acides aspartiques en acides glutamiques, toutes des substitutions n'impliquant le changement que d'une seule base dans leur codon<sup>54</sup>. La substitution des isoleucines en valines est aussi fréquemment observée<sup>55</sup>.

La substitution des lysines en arginines est intéressante notamment à cause des propriétés chimiques du groupement guanidinium de cette dernière. En effet, il permet le maintien de la charge nette positive, a une réactivité chimique réduite en raison de son pKa très élevé, une possibilité de stabilisation par résonance et simultanément la possibilité de faire les liens hydrogène et des ponts salins avec plusieurs résidus. La chaîne de l'arginine est longue, ce qui lui donne une certaine hydrophobicité et lui permet d'interagir directement avec d'autres résidus situés plus loin alors qu'une lysine à la même position doit utiliser une molécule de solvant pour faire la même interaction électrostatique<sup>56</sup>. De plus l'arginine est l'acide aminé le plus hydrophile. Sa constante d'équilibre de transfert vers l'eau est de 9 à 10 fois plus grande que celle de la lysine, le deuxième acide aminé le plus hydrophile. Le mouvement de l'arginine de l'intérieur vers l'extérieur de la protéine est très favorisé. Le remplacement de lysines par des arginines garantit virtuellement que les résidus chargés positivement seront localisés à la surface de la protéine, augmentant ainsi le contraste entre l'intérieur hydrophobe et l'extérieur polaire<sup>65</sup>. Chez la lactate déshydrogénase H4 de cochon le remplacement de lysines externes par des méthyl acétimidates ressemblant à des arginines, augmente sa thermostabilité<sup>57</sup>.

### 1.2.2.3.3 Les interactions aromatique-aromatique

Une mutation introduisant un résidu aromatique dans un environnement où se trouvait déjà deux autres chaînes aromatiques donne un réseau de pairage entre les trois au lieu d'une simple paire. La diminution de l'énergie libre de la structure créée par l'introduction de deux interactions énergiquement favorables entre les chaînes latérales aromatiques amène une stabilisation sans de substantiels changements compensatoires de l'entropie<sup>58</sup>. Par exemple un mutant de la protéine répressive du phage  $\lambda$  s'est avéré avoir une température de dénaturation supérieure de 6°C à celle de la protéine native. La mutation consistait au remplacement d'une glutamine par une tyrosine dans une région contenant une autre tyrosine et une phénylalanine<sup>59</sup>.

#### 1.2.2.3.4 La stabilisation des hélices et des feuillets

L'hélice étant l'unité structurale la plus stable et se repliant la plus rapidement, elle forme les premiers centres de nucléation lors du repliement et agit ensuite comme structure rigide et exposée qui entoure et interagit avec le cœur de la protéine en lui permettant de résister à la dénaturation. Elle constitue donc un élément idéal pour augmenter la thermostabilité, c'est pourquoi les échanges d'acides aminés des protéines thermophiles s'y retrouvent le plus souvent. Chez les thermophiles, plusieurs substitutions qui augmentent notamment la quantité d'isoleucines, d'alanines et de prolines localisées dans la partie non-enfouie de l'hélice, permettent par l'augmentation de contacts hydrophobes, un empilement plus serré et une diminution de la flexibilité de l'hélice<sup>51</sup>. Les substitutions de sérines ou de valines en alanines et d'acides aspartiques en acides glutamiques aident à la stabilisation des hélices  $\alpha$ <sup>60</sup>. La thermostabilité d'une protéine peut aussi être augmentée par la stabilisation des feuillets bêta mais de façon moindre qu'avec les hélices. Ces derniers peuvent être stabilisés par le remplacement de sérines en thréonines<sup>54</sup>.

#### 1.2.2.3.5 Les ponts salins et l'hydratation

La formation d'un pont salin suite à l'association de deux molécules chargées entraîne une pénalité de désolvation qui est très rarement complètement compensée par des interactions favorables entre les ponts salins et le reste de la protéine. Les énergies libres d'hydratation des chaînes latérales chargées sont davantage diminuées par l'augmentation de la température que celles des chaînes latérales hydrophobes de même taille et forme. En effet à hautes températures la pénalité de désolvation pour la formation de ponts salins est dramatiquement réduite, ce qui mène à une augmentation de leur stabilité. Les effets de l'hydratation stabilisent donc préférentiellement les thermophiles à hautes températures alors qu'ils déstabilisent les protéines mésophiles correspondantes<sup>61</sup>.

Perutz a été le premier à attribuer l'augmentation de la thermostabilité chez la ferrédoxine et l'hémoglobine à une augmentation des ponts salins extérieurs<sup>62</sup>. Près de 25 ans plus tard, la détermination de structures tridimensionnelles de protéines hyperthermophiles a effectivement démontré un nombre élevé de paires d'ions organisées en larges réseaux<sup>60</sup>. Par

exemple les glycéraldéhyde-3-phosphate déshydrogénases thermostables contiennent plus de paires d'ions entre leurs sous-unités que chez leurs homologues mésophiles<sup>66</sup>.

#### **1.2.2.3.6 Le ratio surface par volume et la densité d'empilement**

Le point de dénaturation thermique d'une protéine est inversement proportionnel à la longueur de sa chaîne polypeptidique. Comme le ratio surface par volume pour les protéines globulaires diminue avec la longueur de leur chaîne, il a été possible d'établir la relation entre la température de dénaturation à la chaleur d'une protéine et la fraction de sa région polypeptidique accessible pour le solvant<sup>63</sup>. En général, la minimisation de la proportion surface par volume et l'augmentation d'atomes enfouis augmentent la stabilité d'une protéine en réduisant l'énergie de surface non-favorable tout en augmentant les interactions attractives internes. Un empilement serré dans le centre hydrophobique réduit les cavités<sup>60,64,65</sup> diminuant ainsi l'accès pour le solvant et par le fait même, les altérations dans le noyau hydrophobique<sup>51</sup>.

#### **1.2.2.3.7 Le niveau d'association des sous-unités**

Les structures primaires, secondaires, tertiaires et quaternaires sont généralement assez similaires entre les protéines thermophiles et celles de leurs homologues mésophiles, quoique les sous-unités des protéines thermophiles aient généralement un poids moléculaire plus petit<sup>66</sup>. Leur thermostabilité est donc attribuable à de petites modifications dans les régions vulnérables aux attaques par la température sans provoquer de changements dans la conformation de la chaîne principale. Chez les enzymes thermostables monomériques, l'énergie de stabilisation supplémentaire vient généralement de quelques ponts salins additionnels à la surface de la molécule (sans affecter la structure tertiaire) tandis que chez les enzymes oligomériques la stabilisation vient plutôt de l'ajout de quelques ponts salins, liens hydrogène et contacts hydrophobes principalement entre les sous-unités<sup>67,68</sup>. Les interactions entre les sous-unités ont un rôle très important d'autant plus qu'un événement critique avant la dénaturation irréversible est une dissociation très coopérative de l'oligomère en monomère<sup>69</sup>. Plusieurs protéines thermophiles démontrent un état d'association de leurs sous-unités plus élevé qui les stabilise en favorisant des contacts hydrophobes entre les sous-unités et en réduisant la surface exposée au solvant. Les protéines avec de courtes chaînes peptidiques,

comme la triose phosphate isomérase, ou constituées de domaines indépendants qui manquent de contacts entre eux comme la phosphoglycérate kinase, profiteront davantage d'une haute agrégation. Cependant le potentiel stabilisateur des contacts additionnels ne peut être utilisé que si les protéines permettent l'association de leurs sous-unités sans restreindre la flexibilité conformationnelle de la région du site actif ou compromettre leur structure générale<sup>59</sup>. Ce phénomène d'association est observé chez les triose phosphate isomérases des archéobactéries thermophiles *Pyrococcus woesei* et *Methanothermus fervidus*, qui sont tétramériques plutôt que dimériques comme chez les mésophiles dont l'archéobactérie *Methanobacterium bryantii*<sup>70</sup>. C'est aussi le cas de la phosphoribosyl anthranilate isomérase de *Thermotoga maritima*<sup>71</sup> et des phosphoglycérate kinases de *Pyrococcus woesei* et *Methanothermus fervidus*<sup>72</sup>, dont la forme active est dimérique alors que leurs homologues mésophiles sont monomériques. Des expériences de mutagenèse dirigée de la FBP aldolase de muscle de lapin ont aussi démontré que des contacts entre les sous-unité agissent sur la thermostabilité de la molécule sans pour autant changer la cinétique de l'enzyme<sup>73</sup>.

La contribution de la structure quaternaire à la thermostabilité est toutefois controversée puisque la chaîne polypeptidique doit d'abord résister à la chaleur lorsqu'elle se replie en structure tertiaire. C'est seulement une fois la structure quaternaire adoptée que les interactions supplémentaires entre les sous-unités peuvent augmenter la stabilité<sup>66</sup>.

### 1.2.3 Leur intérêt commercial

Les réactions effectuées à haute température comportent de nombreux avantages chimiques et physiques, entre autres une augmentation de la vitesse de diffusion, une plus grande solubilité de la plupart des composants non-gazeux, des substrats et des produits, une réduction de la tension de surface des solvants et de leur viscosité, une diminution de la contamination microbienne et une récupération simplifiée des produits volatiles<sup>37</sup>.

Les thermophiles et leurs composantes en permettant l'utilisation de hautes températures, offrent de nombreuses applications commerciales. Le meilleur exemple est la mise au point de la technique de réaction de polymérisation en chaîne (PCR) rendue possible grâce à la polymérase d'ADN thermostable de *Thermus aquaticus*. La PCR est aujourd'hui utilisée

comme technique de base pour l'amplification de séquences nucléotidiques dans les laboratoires de recherche, pour l'étude de l'évolution moléculaire, le diagnostic de maladies génétiques, la médecine légale, de même que dans le cadre du projet du génome humain. La vente de son brevet s'est chiffré à 300 millions de dollars américains<sup>74</sup>.

### **1.2.3.1 L'utilisation d'organismes thermophiles**

Les organismes thermophiles sont déjà utilisés pour certains procédés industriels. Ils servent à la filtration bactérienne, un procédé hydrométallurgique d'oxydoréduction qui consiste à solubiliser des métaux et à les séparer du minerai. Comme certaines régions deviennent chaudes à cause de l'oxydation exothermique, seules les bactéries thermophiles peuvent résister. De plus le remplacement de la transformation traditionnelle chimique par une biologique, augmente la vitesse d'exécution et diminue la pollution de l'air.

D'autres thermophiles sont aussi utilisés pour la fermentation. Certains peuvent transformer directement des polysaccharides en éthanol. Le fait d'effectuer ce procédé à haute température élimine ou diminue le refroidissement et rend la distillation de l'éthanol plus efficace.

Des thermophiles permettent le traitement anaérobie des déchets afin de produire du méthane ou du compost. La haute température permet la diminution du temps de rétention, l'augmentation de la vitesse de réaction, la destruction des micro-organismes pathogènes contenus dans les vidanges et une diminution de la viscosité des boues résiduelles ce qui rend leur mélange et leur déshydratation plus faciles<sup>37</sup>.

### **1.2.3.2 L'utilisation de protéines thermophiles**

L'utilisation d'enzymes dans les procédés industriels comporte de nombreux avantages. Contrairement aux réactions chimiques les enzymes sont spécifiques, donnent peu ou pas de produits secondaires ou toxiques et sont réutilisables pour plusieurs réactions. Leur principal désavantage est leur instabilité qui peut engendrer des coûts de remplacement élevés. L'utilisation d'enzymes thermophiles diminue ce problème puisqu'en plus d'être plus

résistants à la chaleur, ils le sont aussi aux solvants, aux détergents, aux agents dénaturants et aux procédures d'immobilisation<sup>37</sup>. La faible activité des enzymes thermophiles à la température ambiante diminuerait les problèmes de manipulation et d'entreposage. La rigidité responsable de leur inactivation moléculaire à température de la pièce diminuerait aussi la réponse immunitaire à ces protéines et par conséquent, leurs risques potentiels pour la santé. Certains enzymes thermophiles comme les protéases, ont même une plus grande activité spécifique à leur température optimale que ceux des mésophiles. Pour les autres, les pertes de rendement possibles seraient compensées par un procédé de récupération plus efficace à cause des propriétés des liquides chauds et une réduction du taux de remplacement de l'enzyme<sup>42</sup>.

Les protéines thermophiles pourraient être utiles pour diverses tâches commerciales comme pour le nettoyage dans des préparations contenant des détergents, la récupération de l'argent des émulsions photographiques, la synthèse organique de l'aspartame, la récupération ou l'hydrolyse de protéines de soya, l'hydrolyse de la gélatine, l'attendrissement de la viande<sup>75</sup>, la production de glucose et de maltose à partir d'amidon ou de saccharides<sup>36</sup> et le nettoyage des membranes de filtration utilisées lors de la production du petit-lait<sup>76</sup>.

Une compréhension des bases de la thermoadaptation des protéines thermophiles permet de modifier par génie génétique et protéique ou chimiquement, des enzymes mésophiles pour augmenter leur stabilité ou créer des catalyseurs totalement artificiels ayant des centres actifs modelés sur les sites actifs d'enzymes thermophiles<sup>42</sup>. Le génie des protéines peut aussi rendre possible la production d'enzymes thermophiles avec une pleine activité à basse température<sup>77</sup>.

### **1.2.4 Leur place dans l'évolution**

Actuellement l'origine de la vie soulève un débat quant à savoir si la vie est apparue dans les milieux chauds des sources thermales et hydrothermales des fonds marins, ou bien dans les milieux froids des bassins ou des océans. La plupart des géologues pensent qu'à ces débuts la Terre était couverte de volcans et d'impacts d'astéroïdes qui créaient un environnement chaud.

Les espèces procaryotiques thermophiles sont taxonomiquement reliées aux espèces mésophiles du même genre et ont probablement évolué d'ancêtres communs. Une hypothèse veut que les microorganismes des sources thermales soient des reliques des formes primordiales de vie. Cette hypothèse semble plausible compte tenu de la découverte dans certaines roches formées de dépôts de sources thermales, de fossiles de microorganismes ressemblant à des Flexibactéries communément retrouvées dans les sources d'eau chaude d'aujourd'hui. Les formes de vie à plus basses températures auraient alors découlé des organismes primordiaux thermophiles par mutations et sélection. Les propriétés cinétiques des enzymes et la machinerie de synthèse des protéines se seraient modifiées de sorte que les mésophiles auraient sacrifié leur thermostabilité au dépend d'une flexibilité conformationnelle leur conférant une plus grande affinité avec le substrat. Cette transformation les aurait rendus capable de croître et de fonctionner aussi, sinon plus rapidement, que les thermophiles, même en vivant à une température plus basse<sup>35</sup>.

Au point de vue évolutif, les archéobactéries hyperthermophiles et certaines eubactéries thermophiles représentent les lignages phylogénétiques les plus courts de leur genre respectif et par conséquent, les plus près de la racine de l'arbre phylogénétique universel dérivé de séquences d'ARNr<sup>41</sup>. Selon cet arbre, le plus récent ancêtre commun des formes de vie actuelles serait d'origine chaude et auxotrophe<sup>78</sup> ce qui appuieraient l'hypothèse de l'apparition de la vie à hautes températures. Ceci indique que les thermophiles seraient les organismes existants les plus primitifs et auraient subi peu de changements. L'étude de leur thermostabilité serait le point de départ de la compréhension de l'adaptation des protéines au cours du refroidissement de la Terre<sup>67</sup>.

En assumant que la température était le facteur environnemental changeant le plus vite durant la vie primitive des microorganismes, seule une adaptation des protéines simple et rapide a pu permettre la survie des organismes. Un nombre restreint de mutations auraient suffi à cette adaptation aux conditions mésophiles. Les autres mutations faisant partie d'un processus évolutif normal, auraient permis le maintien de la conformation du site actif<sup>79</sup> et permettrait d'expliquer la similarité des activités spécifiques des enzymes mésophiles et thermophiles à leur température optimale respective<sup>80</sup>.



Toutefois, les arbres phylogénétiques basés sur les protéines ne placent généralement pas les thermophiles sur les branches les plus lointaines<sup>81</sup>. Par exemple l'arbre phylogénétique des Hsp60 place les thermophiles *Thermotoga maritima* et *Aquifex aeolicus* (placés sur la branche la plus profonde des bactéries de l'arbre des ARNr<sup>78</sup>) au milieu de l'embranchement bactérien. Les thermophiles contemporains ne seraient peut être pas représentatifs des espèces bactériennes et archéobactériennes primitives. L'arbre phylogénétique des Hsp60 semble suggéré que les espèces bactériennes thermophiles auraient évolué indépendamment à différents temps à partir de divers lignages mésophiles. La protéine thermophile ne représenterait alors pas l'état ancestral de son homologue mésophile, mais une adaptation distincte<sup>82</sup>.

Récemment une étude de deux molécules d'ARN sensibles à la température, la petite sous-unité ARNr et la grosse sous-unité ARNr, a permis de construire un arbre phylogénétique à partir de 40 organismes différents allant des bactéries aux mammifères. Sachant que le contenu en G/C des ARNr d'un organisme est directement proportionnel à sa température optimale de croissance, l'application d'un modèle informatique a conclu que les ARNr ancestraux avaient un contenu en G/C correspondant à un organisme vivant à une température modérée. Cette étude appuie donc l'hypothèse selon laquelle les thermophiles seraient des mésophiles s'étant adaptés aux hautes températures et les branches courtes attribuées aux thermophiles ne reflèteraient pas une affinité avec l'ancêtre mais une conséquence de l'accroissement de la pression sélective qui agit sur les ARNr des thermophiles<sup>83</sup>.

Comme le dernier ancêtre commun semble avoir été un organisme aussi complexe que les bactéries contemporaines, il se peut qu'il ait lui-même une histoire évolutive complexe. Ses propres ancêtres ont peut être eu à s'adapter plusieurs fois à des changements de température. Ce dernier ancêtre commun ne serait alors pas nécessairement représentatif de la véritable origine de la vie<sup>84</sup>. Le débat reste ouvert tant que les géologues n'auront pas établi avec certitude les conditions prévalant à l'origine de la Terre.

## 1.2.5 L'eubactérie thermophile *Thermus aquaticus*

*Thermus aquaticus* a initialement été isolée en 1969 des sources d'eau chaude du parc de Yellowstone (États-Unis). Par la suite, cet organisme a été retrouvé dans d'autres sources en Islande<sup>85</sup>, en Australie<sup>86</sup>, dans l'eau chaude d'une rivière thermiquement polluée de Belgique<sup>87</sup> et dans des sources hydrothermales du golfe de la Californie et de la côte mid-atlantique des États-Unis<sup>88</sup>. Cette eubactérie thermophile est gram-négatif, sans flagelle, non-mobile, aérobie obligatoire, chimiohétérotrophe, non-endosporique, filamenteuse en forme de baguette et a tendance à s'agréger en larges sphères résistantes à la lyse osmotique dans de vieilles cultures<sup>89</sup>. Elle est sensible à la pénicilline, l'actinomycine D, la streptomycine, le chloramphénicol, la tétracycline et la novobiocine. Sa température de croissance se situe entre 40 et 79°C avec un optimum à 70°C. Son temps de génération à 70°C est de 50 minutes. Cette bactérie croît dans un milieu synthétique composé principalement de sels et est inhibée par un milieu contenant plus de 1% de mélange de composés organiques (extraît de levure et tryptone). Son pH de prolifération se situe entre 6,0 et 9,5 avec un optimum autour de pH 7,5-7,8. Ses conditions de croissance reflètent son adaptation aux sources thermales faibles en composés organiques et riches en sels inorganiques.

### 1.2.5.1 Ses lipides

Chez *Taq* les lipides constituent 8 à 12% du poids cellulaire sec. 60% de ses lipides membranaires sont des caroténoïdes généralement très polaires dont des phytoènes et des  $\beta$ -carotènes qui donnent à la bactérie sa couleur orangée. Ses lipides neutres sont constitués entre autres de ménaquinones (vitamine K<sub>2</sub>) à 35, 40 et 45 carbones. Son principal glycolipide est un glycosyl-glycosyl-glycosaminyl-glycosyl-diacylglycérol dont la taille du groupement hydrophobe est deux fois plus grande que celle du principal glycolipide retrouvé chez les bactéries mésophiles. Ses phospholipides qui comptent pour 30% des lipides totaux sont inhabituels, voir même uniques. Le phospholipide majeur comptant pour 80% de la quantité totale de ce type de lipide, et donc 35 à 42% des lipides totaux, n'a été que partiellement caractérisé. Il semble être un glycophospholipide contenant un glycérol, 3 acides gras dont un glucoamine N-acylé et une longue chaîne amine insaturée comprenant un hydroxyle. La

détermination de la structure de ce phospholipide si abondant jouera un rôle important dans la compréhension de la thermostabilité de la membrane de *Taq*. La plupart de ses lipides polaires contiennent des carbohydrates dont des sucres aminés, qui contribueraient à la thermostabilité de la membrane. De plus la température de croissance de *Taq* influence directement la composition de sa membrane. Le passage progressif de cette température de 50°C à 75°C se traduit sur le contenu lipidique total par une augmentation de 1,8 fois des caroténoïdes, de 2 fois des phospholipides et de 2 à 4 fois des glycolipides. Toutefois la quantité de vitamine K<sub>2</sub> reste constante malgré la variation de température. Ceci suggère que le système de transport d'électrons n'est pas modifié lors de la croissance à haute température<sup>39,90</sup>. C'est donc dire que les *Taq* sont capables d'adapter leurs membranes en fonction de la température environnante pour en optimiser le rôle et par le fait même, leur propre survie.

### 1.2.5.2 Ses acides nucléiques et ses ribosomes

La majorité des ARNt des *Thermus* ont un C ou un G comme première lettre de l'anticodon étant donné la grande fréquence de ces nucléotides dans leur génome. Leur ARNt pour la formylméthionine contient une paire de base G-C dans sa tige T au lieu du G-U retrouvé chez *E.coli*. De plus ils contiennent trois bases modifiées, soit un ribose méthylé en G-19 et en A-59 et un 2-thioribosylthymidine en position 54 qui renforce les interactions entre les bases voisines dans l'hélice<sup>38,91</sup>. Les souches *Thermus* ne contiennent pas moins de douze polyamines différentes dont la thermine et la thermospermine qui augmentent la synthèse protéique en stabilisant à haute température le complexe formé par l'ARN messenger, le ribosome et les aminoacyl-ARNt<sup>41,91</sup>. L'ADN génomique de *Taq* est constitué à 67,4% de G/C<sup>89</sup>. Dans les protéines de *Taq* séquencées jusqu'à maintenant, les codons contiennent respectivement 68,1%, 43,3% et 88,8% de G/C en première, deuxième et troisième position<sup>92</sup>.

### 1.2.5.3 Ses protéines

La vitesse de remplacement (turnover) des protéines chez *Taq* n'est pas plus grande que celle des mésophiles et n'explique donc pas sa thermostabilité<sup>93</sup>. C'est la

thermoadaptation de ses protéines qui expliquent le maintien de leur structure et de leur activité à hautes températures. Les structures tridimensionnelles de protéines de *Taq*, quoiqu'en quantité restreinte, démontrent les mêmes adaptations que les protéines thermostables d'autres organismes telles que l'augmentation du nombre de ponts salins et de l'hydrophobicité du centre de la protéine<sup>94,95</sup>. De plus, la composition des protéines de *Taq* serait en partie influencée par le haut contenu en G/C de son ADN, qui se traduirait par une augmentation des leucines, valines, alanines, prolines, arginines et des glycines, de même qu'un appauvrissement en isoleucines, méthionines, phénylalanines, sérines, thréonines, cystéines et tryptophanes. Cette tendance est aussi observée chez des mésophiles riches en G/C<sup>96</sup>.

#### 1.2.5.4 Sa fructose-1,6-bisphosphate aldolase

Peu après la découverte de *Thermus aquaticus*, sa FBP aldolase a été partiellement purifiée. L'activité enzymatique de la protéine obtenue est inhibée par l'EDTA et activée par de faibles concentrations de cobalt, de fer, de cystéine ou de fortes concentrations de  $\text{NH}_4^+$ . Son activité est très faible en deçà de 60°C et optimale à 95°C. Son poids moléculaire a été estimé à 140 kDa par gel de filtration. Ses propriétés sont comparables à celles des FBP aldolases de classe II<sup>28</sup>. Plus récemment la FBP aldolase de *Taq* a été purifiée jusqu'à homogénéité. L'enzyme a la même forme quaternaire et le même comportement en présence d'EDTA ou de cobalt que la protéine précédemment partiellement purifiée. Sa thermostabilité est indéniable étant donné que 75% de son activité initiale est encore détectée après deux heures de chauffage à 90°C. De plus son activation par des détergents et des alcools, de même que sa très forte affinité pour les colonnes hydrophobes semblent démontrer la présence de régions hydrophobes à sa surface. Son étude cinétique montre une coopérativité négative entre ses sous-unités. La partie N-terminale de l'enzyme a été séquencée et révèle une similarité significative avec la FBP aldolase de classe II de *Bacillus stearothermophilus*<sup>29</sup>.

### 1.3 Les objectifs de ce projet

Le but de ce projet de recherche est d'abord de trouver le gène codant pour la FBP aldolase de *Thermus aquaticus* et de le séquencer afin d'en déterminer la séquence protéique correspondante. Une comparaison avec les autres FBP aldolases de classe II permettra de déterminer à quelle sous-classe cette FBP aldolase appartient et quelles modifications peuvent expliquer sa thermostabilité. La protéine sera ensuite surexprimée dans *E.coli*, puis purifiée jusqu'à homogénéité. L'enzyme sera caractérisé afin d'être comparé à celui purifié directement de *Thermus aquaticus*. Le but ultime est la cristallisation de la protéine afin d'obtenir, par diffraction aux rayons-X, les données nécessaires à la détermination de sa structure tridimensionnelle.

## **Chapitre 2: Article**

**Fructose-1,6-bisphosphate aldolase from *Thermus aquaticus*.  
Gene Sequence and Crystallization of the Overexpressed Enzyme**

Véronique SAUVÉ and Jurgen SYGUSCH

Department of Biochemistry, Université de Montréal, Canada

*Subdivision:* 3. Protein chemistry and structure

*Correspondance* should be addressed to Jurgen Sygusch, Department of Biochemistry, Université de Montréal, C.P.6128, Station Centre Ville, Montréal, Québec, Canada, H3C 3J7.

Tel: 1-514-343-2389; Fax: 1-514-343-6463

E-mail: Jurgen.Sygusch@Umontreal.CA.

*Abbreviations.* BAP, Bacterial alkaline phosphatase; DHAP, Dihydroxyacetone phosphate; DLS, Dynamic light scattering; DMSO, Dimethyl sulfoxide; DTNB, 5,5'-dithiobis(2-nitrobenzoic acid); FBP, Fructose-1,6-bisphosphate; GAP, Glyceraldehyde-3-phosphate; GAPDH, Glyceraldehyde-3-phosphate dehydrogenase; GDH,  $\alpha$ -Glycerol-3-phosphate dehydrogenase; IPTG, isopropyl- $\beta$ -thiogalactoside; PGH, Phosphoglycolohydroxamate; PMSF, Phenylmethylsulfonyl fluoride; SSB, Single-stranded DNA binding proteins; *Taq*, *Thermus aquaticus*; TIM, Triose-phosphate isomerase.

*Enzymes.* Fructose-1,6-bisphosphate aldolase (EC 4.1.2.13); Triose-phosphate isomerase (EC 5.3.1.1);  $\alpha$ -Glycerol-3-phosphate dehydrogenase (EC 1.1.1.8); Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)

*Running title.* Fructose-1,6-bisphosphate aldolase from *Thermus aquaticus*.



## Summary

Gene sequence coding for fructose-1,6-bisphosphate aldolase from the thermophile eubacteria *Thermus aquaticus* YT-1 was cloned using a degenerated oligonucleotide probe corresponding to the N-terminal amino acid sequence of the enzyme. Nucleotide-sequence analysis revealed an open reading frame coding for a protein of 305 amino acids (M<sub>r</sub> 33092 Da). Sequence comparisons indicate greatest similarity with aldolases from extremophiles *Thermotoga maritima*, *Aquifex aeolicus* and *Helicobacter pylori* (49-54% identity, 68-81% homology). The coding sequence for *Taq* FBP aldolase was overexpressed under the control of the *tac* promoter in *E.coli* and purified to homogeneity by heat treatment followed by two chromatographic columns: a Phosphocellulose P11 and a Q-Sepharose. Non-denaturing PAGE, gel filtration and DLS analysis were consistent with a tetrameric quaternary structure. The recombinant enzyme is thermostable for at least 2 hours at 80°C with little residual activity below 60°C. Kinetic characterisation at 70°C indicates extensive negative subunit cooperativity ( $h = 0,32$ ) with a limiting  $K_m$  of 305  $\mu$ M. Maximal specific activity ( $V_{max}$ ) is 46 U/mg at 70°C. Single crystals of *Taq* FBP aldolase were grown from ammonium sulfate solution by vapor-diffusion. The crystals diffract beyond 2Å using synchrotron radiation and have space group P2 with unit-cell parameters of  $a = 99.5\text{\AA}$ ,  $b = 57.5\text{\AA}$  and  $c = 138.6\text{\AA}$ , and  $\beta = 90.25^\circ$  that corresponds to a tetramer in the asymmetric unit cell.

**Keywords:** *Thermus aquaticus*, class II aldolase, overexpression, kinetics, crystallization

## Introduction

Thermophilic organisms are fascinating by their successful adaptation to a hostile high temperature environment. *Thermus aquaticus* YT-1 (*Taq*) is a gram-negative, non-sporulating, non-mobile, aerobic, rod-shaped filamentous eubacteria that grows naturally at 70°C in thermal springs in Yellowstone Park [1]. Proteins derived from *Thermus aquaticus* as well as other thermophiles possess not only enhanced stability with respect to high temperature but also in presence of detergents, alcohols and other organic solvents [2]. Thermophilic enzymes appear to have no major differences in amino acids sequence, 3D-structure or catalytic mechanism compared to their mesophilic homologues. Thermophilic enzymes usually have only a few additional ion pairs, hydrogen bonds or hydrophobic patches to strengthen the subunit interactions and to keep the protein core more compact and less accessible to solvent [3]. The understanding of the thermal adaptation is useful in the design of new thermostable proteins. *Taq* DNA polymerase used in PCR represents a good example of the potential commercial applicability of the thermophilic proteins.

In living organisms, fructose-1,6-bisphosphate aldolase is an ubiquitous enzyme essential for glycolysis, gluconeogenesis and Calvin cycle. In glycolysis, it catalyses the reversible cleavage of fructose-1,6-bisphosphate into dihydroxyacetone phosphate and glyceraldehyde-3-phosphate. Two distinct classes of FBP aldolase have been characterized [4].

Class I FBP aldolases form a Schiff base intermediate between the substrate's carbonyl group and the  $\epsilon$ -amine group of a lysine residue in the active site. Class II FBP aldolases are EDTA sensitive using a divalent metal cation as electrophile in their mechanism and their activity can be enhanced by monovalent cations. Class II enzymes are found in

bacteria (eubacteria and archaeobacteria), fungi and in some green algae grown in heterotrophic conditions [4,5] and are usually dimers except for those from *Taq* [6,7], *Anacystis nidulans* and *Saprospira thermalis* [8] which are tetramers. Alignment of class II aldolases reveals two subdivisions, referred to as A and B [5]. So far, only the structure of the class II A FBP aldolase from *E. coli* is known [5]. In this aldolase, the polypeptide folding of each subunit corresponds to that of a  $(\alpha/\beta)_8$  barrel with an active site on the C-terminal side of the  $\beta$  strands that contains a cation and two zinc binding sites [9]. Kinetics were determined recently of point mutants of the *E. coli* enzyme [10,11] and a structure in presence of the transition state analogue PGH [12] was solved to address the reaction mechanism in class II aldolases. To date, no structure of a thermophilic or class II B FBP aldolase has been determined.

In this study, the gene of the FBP aldolase from *Thermus aquaticus* was cloned and overexpressed in *E. coli*. Purification, characterisation, and crystallization of the overexpressed protein, essential for structural determination, are described.

## **Materials and methods**

### **Bacteria, culture conditions and plasmids**

*Thermus aquaticus* YT-1 (ATCC 25104) were grown in media 481 at 70°C. *E. coli* DH5 $\alpha$ MCR were purchased from Gibco BRL. *E. coli* JM109 and pGEM-T were acquired from Promega. *E. coli* strains were grown in DYT-Ampicillin media (100 $\mu$ g/ml). The cloning vectors pUC18 *Bam*HI BAP and pUC18 *Sma*I BAP and the expression plasmid pKK223-3 were obtained from Amersham Pharmacia Biotech.

### **Enzymes and chemicals**

All reagents were purchased from Fisher. Tris, IPTG, PMSF, FBP, TIM-GDH and

Expand™ Hi Fidelity PCR system were obtained from Roche Molecular Biochemicals. Enzymes for DNA manipulation, SSB, deaza G/A T7 sequencing kit, [ $\gamma$ - $^{32}$ P]ATP, [ $^{35}$ S]dATP $\alpha$ S, DNase I, RNase I A, catalase, rabbit muscle aldolase, bovine serum albumin, ovalbumin, Q-Sepharose, HiLoad 16/60 Superdex 200 and FPLC were provided from Amersham Pharmacia Biotech. Alumina and Mes were purchased from Sigma, ammonium sulfate from Gibco BRL, Phosphocellulose P11 and alumina matrix filters (0,02  $\mu$ m) from Whatman, bicinchoninic acid (BCA) protein assay reagent kit was obtained from Pierce, and mid-range protein molecular weight marker from Promega.

### **Molecular cloning of the FBP aldolase from *Thermus aquaticus***

Genomic DNA was obtained from *Thermus aquaticus* using standard extraction methods [13] for extraction of bacterial genomic DNA. A degenerated oligonucleotide probe corresponding to the N-terminal amino acid sequence of *Taq* FBP aldolase [7] was 5'-labelled using [ $\gamma$ - $^{32}$ P]ATP and T4 polynucleotide kinase and used for Southern hybridization of genomic DNA previously digested by restriction enzymes. Probe hybridization was performed with 250 mM PO<sub>4</sub> pH 7,0 buffer and 7% SDS at low stringency conditions (overnight at 50°C). Filters were washed with 0,2X SSC (30 mM NaCl; 3 mM Na citrate pH 7,0); 0,1% SDS at room temperature. *Bam*HI genomic fragments that hybridized with the oligonucleotide probe were subcloned into pUC18 *Bam*HI/BAP and used to transform *E. coli* DH5 $\alpha$ MCR. Colony hybridization was performed using the same N-terminal probe under identical conditions. Plasmids containing the N-terminal gene portion were sequenced with the deaza G/A T7 sequencing kit using the “troublesome” template protocol in presence of DMSO and with addition of single-strand binding proteins [14]. The restriction fragment sequence obtained did not correspond to the entire gene product.

The screening procedure was repeated with other restriction enzyme digests. Southern blots were screened using two probes: sequence 5'CTTCATGGCCCCCTCGGAGAGGGCC 3' corresponded to the sequenced N-terminal region while probe sequence 5'AGCCTGGCGATCCTTTCCAGCCG 3' corresponded to the previously sequenced C-terminal region. Both probe hybridizations were done at high stringency (overnight at 60°C) and the membranes were washed in 0,1X SSC (15 mM NaCl; 1,5 mM Na citrate pH 7,0); 0,1% SDS at 60°C. *StuI* genomic fragments were subcloned in pUC18 *SmaI* /BAP and transformed in *E. coli* DH5 $\alpha$ MCR. Recombinant plasmids were screened with the most proximal C-terminal probe at high stringency. The pUC-*Taq StuI* clones containing the entire FBP aldolase gene were sequenced as described previously. Standard methods were used for the DNA manipulation [15].

#### **Expression vector construction**

PCR was performed with the Expand Hi Fidelity DNA polymerase to amplify the gene and to introduce new restriction sites at each end of the coding sequence. Two oligonucleotide primer sequences were used: sequences 5' TAA TCC GGG CCG GGG GAA TTC ATG CTG GTT ACG GGT CTA GAG 3' and 5'GCC ACG GCG AAG CTT TTA GGC CCG GCC CAC GGA GCC3' where underlined nucleotides correspond respectively to *EcoRI* and *HindIII* restriction sites while the bold letters designate nucleotides coding for the first and the last amino acid of the gene respectively. The 954pb A-tailed PCR product was ligated in the pGEM-T PCR vector. The plasmid pGEM-T/aldoTaq was digested with *EcoRI* and *HindIII*. The 924 pb fragment containing the *Taq* FBP aldolase gene was ligated into the expression vector pKK223-3 yielding expression plasmid pKKaldoTaq. The recombinant plasmid was used to transform *E. coli* JM109 to create an inducible gene product.

### **Overexpression and Purification of *Thermus aquaticus* aldolase**

*Overexpression.* A 11-culture of *E.coli* JM109 transformed with plasmid pKKaldoTaq was grown at 37°C in DYT-Ampicillin to an OD<sub>600</sub> of 0,6. IPTG was then added to a final concentration of 1mM and incubation was continued overnight.

*Crude extract and delipidation.* The cells were harvested by centrifugation (8000 g, 15 min at 4°C). The wet pellet was ground with twice its weight of alumina in presence of DNase I and RNase IA (0,2 mg/g of cells). The cellular paste was resuspended in 50 ml of lysis buffer (50 mM Tris-HCl pH 7,3; 0,1 mM CoCl<sub>2</sub>; 1 mM PMSF) and left 15 min on ice. The lysate was centrifuged at 35000g for 20 min at 4°C to remove alumina and other cellular debris. To delipidate the crude cell-free extract, the pH was lowered to 5,4 with 1 M acetic acid and then followed by centrifugation at 35000g for 15 min at 4 °C. The pellet was discarded and the pH of the recovered supernatant was restored to 7,3 with 1 M Tris.

*Heat treatment.* The protein solution was two-fold diluted with lysis buffer and heated to 80 °C for 45 min while gently agitating at 75 rpm. Heat-denatured proteins were removed by harvesting at 40000g for 25 min at 4°C. The recovered supernatant was then dialyzed against 20 mM Mes-NaOH pH 6,4 at room temperature.

*Phosphocellulose Chromatography.* The dialysate was loaded onto a phosphocellulose P11 column (2g/liter of culture) previously equilibrated with the same buffer. The column was first rinsed with loading buffer followed by an additional rinse at 20mM Mes-NaOH pH 6,6 until no further protein eluted. *Taq* FBP aldolase was eluted using 20mM Mes-NaOH pH 6,6 in presence of 1mM FBP. Active fractions were pooled and dialyzed against 50 mM Tris-HCl pH 7,3 at room temperature.

*Q-Sepharose chromatography.* The protein solution was applied to a Q-sepharose

column equilibrated with the dialyzing buffer. After washing of the column with the loading buffer, the protein was eluted using a linear gradient of NaCl (0-600 mM in 50 mM Tris-HCl pH 7,3). The active fractions were dialyzed against 10 mM Tris-HCl pH 7,3 and stored at room temperature.

### **Protein concentration**

Protein concentration was determined with the BCA protein assay reagent kit at 60°C and using Bovine Serum Albumin as reference standard.

### **SDS and non-denaturing gel electrophoresis**

SDS-PAGE analysis was performed using a 15 % acrylamide gel run in a Laemmli buffer system [16]. Samples of ~ 3 µg of protein were used and boiled 2 minutes before loading. The mid-range protein markers from Promega were used as relative molecular mass standards. A non-denaturing gel of 7,5 % acrylamide was used to analyze the quaternary form of *Taq* FBP aldolase. Catalase (232 kDa), rabbit muscle aldolase (158 kDa), bovine serum albumin (67 kDa) and ovalbumin (43 kDa) served as reference for molecular weight estimation. Proteins were visualized by staining with Coomassie blue.

### **Colorimetric detection of FBP aldolase on non-denaturing gel**

A colorimetric reaction based on a histochemical assay [17] specific for FBP aldolase activity was used to detect migration of FBP aldolase in non-denaturing gels. Following electrophoretic migration, the polyacrylamide gel was placed at room temperature on 3M paper saturated with a colorimetric solution made up of 50 mM Tris-HCl pH 7,3; 0,1 mM CoCl<sub>2</sub>, 10 mM Na<sub>2</sub>HAsO<sub>4</sub>; 2,2 mM NAD<sup>+</sup>; 10 mM FDP; 125 µg/ml GAPDH; 1 mg/ml Nitro Blue tetrazolium; 0,1 mg/ml Phenazine methosulfate. A purple coloration corresponds to aldolase activity.

### **Dynamic light scattering (DLS)**

DLS experiments were performed using a DynaPro-801 TC (Protein Solutions) apparatus. The laser wavelength used was 853,4 nm and measurements were taken at a scattering angle of 90°. The sample solutions were centrifuged and filtered through inorganic alumina matrix membrane filters (0,02 µm) before taking DLS measurements.

### **Molecular weight determination by gel filtration**

1,5 mg of *Taq* FBP aldolase in 10 mM Tris-HCl pH 7,3 were injected on a gel filtration Superdex 200 previously equilibrated with 10 mM Tris-HCl pH 7,3. The column was performed on FPLC at 1 ml/min in the loading buffer. A 3-mg sample of rabbit muscle aldolase was injected on the same column in 50 mM Tris-HCl pH 7,5; 1 mM EDTA and 100 mM NaCl and was eluted from the column with the same parameter.

### **Enzymatic assays**

Aldolase cleavage activity towards the fructose-1,6-bisphosphate substrate was assayed by the spectrophotometric measurement of NADH oxidation at 340 nm by the method of Racker [18]. The enzymatic activity determination was by end point assay because of the denaturation of the coupling enzymes at high temperatures. To determine, the pH for maximal enzymatic activity at 70°C, an assay mixture was employed consisting of 5 mM FBP and 1 mM CoCl<sub>2</sub> made up in 50 mM Tris-HCl and buffered at various pH [19]. Temperature dependence in Tris pKa of -0,31/10°C was used to adjust buffer pH to the working temperature. The assay mixture was preheated for 4 min to 70°C before the enzyme was added and the reaction was then allowed to proceed for different lengths of time.



The reaction mixture was stopped by addition of EDTA to a final concentration of 100 mM and chilled on ice. NADH was next added at a final concentration of 0,142 mM to the reaction mixture at room temperature. The coupling enzymes TIM-GDH were then added to the assay and NADH oxidation was monitored at 340 nm with a Perkin Elmer UV/visible  $\lambda 2$  spectrophotometer.

Kinetic measurements were performed by the same protocol using different concentrations of FBP made up in 1 mM CoCl<sub>2</sub>, 50 mM Tris-HCl pH 6,5 at 70°C. Enzymatic reactions at different temperatures were performed in 5 mM FBP; 1 mM CoCl<sub>2</sub>; 50 mM Tris-HCl pH 6,5 buffered at the corresponding temperature using the procedure described above. One unit of activity (U) corresponds to one  $\mu$ mol of FBP cleaved per min.

### **Crystallization and X-ray diffraction of *Taq* FBP aldolase**

Various crystallization screens, based on the sparse matrix method [20], were used to determine crystallization conditions for *Taq* FBP aldolase. Screening technique for protein crystal growth conditions was by vapor diffusion and employed both hanging and sitting drops [21].

High-resolution data were collected at beam line X-8C of the National Synchrotron Light Source (NSLS) at Brookhaven National Laboratory using a Quantum 4R CCD detector. The radiation wavelength used was 1,105Å and the detector distance was set to 150 mm. Data reduction was performed using DENZO and Scalepack [22].

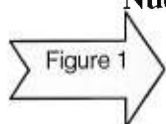
## Results and Discussion

### Molecular cloning of the FBP aldolase from *Thermus aquaticus*

An N-terminal amino acid sequence of FBP aldolase from *Thermus aquaticus* was reported previously [7]. A degenerate oligonucleotide was synthesized corresponding to the N-terminal sequence and coding for residues 10 - 29 (KAHEEGYGVGAFNV?NMGP). This region of sequence possessed low redundancy in terms of base substitution per codon. Nucleotide sequence degeneracy was allowed for in case of ambiguity between two possible bases and inosine (I) was used whenever there was four-fold codon degeneracy. At the position of the unknown amino acid (?), alignment of class II FBP aldolases indicated a large preference for asparagine and a smaller preference for threonine residues. The degenerate nucleotide coding sequence A(A/C)(C/T) used at that position accounted for all codon degeneracies of asparagine and partial codon degeneracy of threonine. The final probe sequence used was 5' AA(A/G) AA(A/G) GCI CA(T/C) GA(A/G) GA(A/G) GGI TA(T/C) GGI GTI GGI GCI TT(T/C) AA(T/C) GTI A(A/C)(C/T) AA(T/C) ATG GGI TT 3'. Because of the primer degeneracy, low stringency conditions were used for screening restriction digests of *Taq* DNA. Genomic fragments of 1,1kb from a *Bam*HI restriction digest yielded a positive signal and was subcloned in pUC18 *Bam*HI BAP. Positive clones were detected by colony hybridization with the same N-terminal probe and sequenced on both strands. Deaza A and G nucleotides, DMSO and SSB proteins were used to eliminate sequence compressions in regions of high G/C content considerably improving sequencing gel readability. Sequencing of the *Bam*HI fragment revealed only the first 703 bp of the *Taq* FBP aldolase coding sequence. *Taq* genomic DNA was digested with other restriction enzymes and restriction digests were probed under high stringency conditions using an N-terminal and C-terminal

probe from the partial gene sequence. A genomic fragment of 3kb from a *StuI* restriction digest hybridized to both N-terminal and C-terminal probes and was subcloned in pUC18 *SmaI* BAP site. The recombinant plasmid containing the 3kb fragment was detected by screening with the C-terminal probe and sequenced in both directions. The *StuI* restriction fragment contained the entire gene sequence of *Taq* FBP aldolase.

### Nucleotide sequence of the *Taq* FBP aldolase



The gene sequence of *Thermus aquaticus* FBP aldolase has an ORF of 918 bp coding for a protein of 305 amino acids shown in figure 1. The relative molecular weight per subunit was determined to be 33092 Da which correspond to the previous value estimated on SDS/PAGE for the native enzyme purified from *Thermus aquaticus* [7]. The G/C content for the coding region is 67,3%. The G/C content of the gene at first, second and third codon positions are respectively 69,0%, 42,2%, and 90,8%. These percentages are comparable with G/C values of 68,1%, 43,3%, and 88,8% usually found in *Thermus aquaticus* genes [23]. The exceptional incidence of G/C in third position constitutes a significant codon bias. The high content of G/C most likely stabilizes the duplex form of *Taq* DNA at high temperatures.

### Amino acid composition and thermophilic adaptations

The deduced N-terminal amino acid sequence of the protein differs slightly from the reported sequence [7]. Amino acid 13 is an arginine instead of histidine residue while at position 28 there is a glutamic acid instead of a glycine residue. The 25th (unknown) amino acid was determined to be an asparagine consistent with other class II FBP aldolases.

*Taq* FBP aldolase has a somewhat larger content of hydrophobic residues than in other class II FBP aldolases (47,2% compared to 39,9 - 45,7%). The number of charged side chains are however comparable to those found in other extreme thermophiles; 30,8% compared to

31,9% and 32,0% in *Aquifex aeolicus* and *Thermotoga maritima* respectively. This number is somewhat higher than that found in other class II FBP aldolases (22,1-28,9%) and would favor formation of an increased number of salt bridges and hydrogen bond networks. Preponderance of these substitutions is however made to the detriment of a reduced number of uncharged polar residues; 13,5% compared to 16,1-27,4%.

The *Taq* FBP aldolase contains no tryptophan similar to other class II enzymes. As in many other thermophilic proteins, cysteine residues are absent and there are fewer methionines because of their greater susceptibility to oxidation at high temperatures [24]. In proteins coded by the *Thermus aquaticus* genome, cysteines are usually replaced by alanine or valine residues. The low number of glutamine and asparagine residues present in *Taq* aldolase reflects the propensity by these amino acids to undergo deamination at high temperatures [24]. The large number of alanine residues, which are found frequently in helices [25], represent a recurring substitution for serine residues in thermostable proteins including *Taq* aldolase and are thought to enhance helix packing thereby increasing protein rigidity [26,27]. The replacement of isoleucine residues for valines, another frequent substitution in thermophiles, is surprising as it decreases protein hydrophobicity [27]. This unexpected replacement may be explained by the bias resulting from the high G/C content of the genomic DNA. The codon first letter is particularly important because it reflects the genomic G/C content bias in terms of protein amino acid composition. The nucleotide representations for isoleucine and valine codons are AT(T/C/A) and GT(T/C/A/G), respectively. Thus G in first position favors valines and the substitution of valine for isoleucine then implicates change of only one base [28].

An increase in the number of glutamic acid residues and a decrease in aspartic acid residues, frequently found in helices [25], maintains the number of negative charges constant

in *Taq* aldolase. The longer lateral side chain of glutamic acid residues is able to interact with further residues while simultaneously increasing protein surface hydrophobicity [26].

The *Taq* enzyme contains an abnormally large number of arginine residues mainly attributable to substitution for lysine, a frequent replacement, and which is characteristic of thermophilic proteins [26]. This common exchange has the advantage of preserving the positive charge on the protein, while the longer arginine side chain can interact with further amino acids without solvent molecule intermediate. The resonance capacity of the arginine guanidinium moiety particularly favors formation of salt bridges and hydrogen bonding with more than one residue. This moiety has also a reduced chemical reactivity because of its high pKa [29,30].

#### Comparison between *Taq* FBP aldolase and the other class II FBP aldolases

Figure 2

The alignment of class II FBP aldolases, shown in figure 2, reveals two distinct groups [5]. Group A contains most class II FBP aldolases that function in glycolysis and gluconeogenesis and includes as members gram positive bacteria with high G/C content,  $\epsilon$ - and  $\gamma$ -proteobacteria and eukaryotes. In contrast, group B is more heterogeneous and contains glycolytic and gluconeogenetic class II aldolases from gram positive bacteria with low G/C content, *Giardia lamblia* (an amichondriate protist) and extremophiles including *Helicobacter pylori*, *Aquifex aeolicus*, *Thermotoga maritima*. Other members in group B have diverse metabolic roles and include FBP aldolases functioning in Calvin cycle of cyanobacteria,  $\alpha$ - and  $\beta$ -proteobacteria using sedoheptulose substrates [31] as well as proteobacteria aldolases with other functions.

On basis of sequence alignment shown in figure 2, *Taq* FBP aldolase belongs to group B. It is very similar in amino acid sequence to aldolases from *Thermotoga maritima*,

*Helicobacter pylori* and *Aquifex aeolicus* having respectively 54%, 51%, 49% identity and 81%, 71%, 68% homology. Similarity with respect to group A FBP aldolase from *E. coli*, the homology drops to 41% with an identity of 27%.

The sequence alignment of *Taq* FBP aldolase, shown in figure 2, indicates numerous conserved residues: H81, E132, H178 and H208 that chelate with the divalent metal ion [9], H78, D80, E130, K249, N251 that form the monovalent cation site [9], D80 which may polarize the C<sub>1</sub> carbonyl group of GAP [11] and R278 interacting with the C<sub>6</sub> phosphate of FBP (or C<sub>3</sub> phosphate of GAP) [10]. Invariant residues D102 and E130 interact with histidine residues coordinating the divalent ion while conserved residues G179, S211, D253, T254 would bind to the PGH phosphate [12].

The *Taq* FBP aldolase secondary structure predicted by NNPREPDICTION [32] was compared with the reported secondary structures of *E. coli* in PDB database (accession number 2392284). The secondary structure prediction is consistent with subunit polypeptide folding corresponding to that of an ( $\alpha/\beta$ )<sub>8</sub> barrel found in the *E. coli* crystal structure [9] and the alignment of secondary structure elements is shown in figure 2. The  $\beta$ -barrel commences at  $\beta$ -strand 1 in *E. coli* and ends at  $\alpha$ -helix 10 with secondary structure predictions for corresponding helices 4, 5 and 8 found in *E. coli* being in poorer agreement due to their apparent shortness. The shorter loop between  $\alpha$ -helix 10 and 11 could modify the hydrophobic handshake interaction reported between *E. coli* aldolase subunits [9].

An insertion of 21 residues in *Taq* aldolase, shown by a banner in figure 2, is found immediately after Ser-211 in the *Taq* aldolase sequence. In *E. coli*, an equivalent insertion of such a 21 residue sequence would map to Ser-267 which is part of the GGSG motif that

interacts with DHAP in the structure of *E. coli* aldolase (J. Sygusch, private communication).

This 21 residue sequence insertion is unique to two sub-classes of the class II B group: the sub-class of cyanobacteria,  $\alpha$ - and  $\beta$ -proteobacteria using Calvin cycle for fixation of carbon dioxide as well as *Giardia* (a amitochondrial diplomonad protist which shared an ancestry with  $\alpha$ -proteobacteria [33]) and the heterogeneous sub-class of extremophiles consisting of *Thermus*, *Thermotoga*, *Aquifex* and *Helicobacter*. Secondary structure prediction suggests considerable  $\alpha$ -helical content by the insertion sequence. The placement of this sequence insertion in the active site locus suggests a role for activity regulation of these enzymes.

#### Expression of *Taq* FBP aldolase in *E.coli*

For study of structure/function relationships, overexpression of *Taq* FBP aldolase was carried out using the pKK223-3 expression vector in *E. coli*. *Taq* FBP aldolase obtained by PCR was cloned in the PCR vector pGEM-T. The 924pb *EcoRI-HindIII* fragment of pGEM-T/aldoTaq was then subcloned into pKK223-3 that had been previously digested by *EcoRI-HindIII* to yield the construct pKKaldoTaq. The recombinant plasmid was transformed into *E. coli* JM109 cells to place it under control of the *tac* promoter. After overnight induction with IPTG, SDS denaturing gel electrophoresis of *E. coli* cell lysates yielded a strong band at 33



Figure 3

kDa relative molecular weight corresponding to the overexpressed *Taq* FBP aldolase (figure 3, lane 1). Under conditions of constitutive expression, that is, in absence of promoter control by *lac* repressor using *E. coli* DH5 $\alpha$ MCR cells, *Taq* FBP aldolase strongly inhibited cell growth.

### **Purification of recombinant *Taq* FBP aldolase**

Table 1

The summary of the purification protocol is shown in table 1. *Taq* FBP aldolase was purified to near homogeneity by heat treatment (figure 3, lane 3). The remaining minor contaminants were removed using the Phosphocellulose P11 column (figure 3, lane 4). Affinity elution with the natural substrate (FBP) dissociated the recombinant enzyme from the column and yielded a virtually homogeneous protein. A Q-Sepharose step (figure 3, lane 5) was used to concentrate FBP aldolase and remove minor contaminants including FBP to yield a homogeneous preparation. A 1-l culture corresponding to ~5 g of bacteria yields 40-50 mg of pure *Taq* FBP aldolase. Dynamic Light Scattering gave linear logarithmic plots for the autocorrelation function and confirmed the purified protein solution to be monodisperse.

### **Subunit composition and molecular mass**

From SDS-PAGE, purified *Taq* FBP aldolase migrates as a single band of ~33kDa (figure 3, lane 5) in agreement both with the molecular weight calculated from the predicated amino acid sequence and with the value reported for the native enzyme purified from *Thermus aquaticus* [7]. On non-denaturing PAGE gel, the estimated molecular mass for the native *Taq* FBP aldolase was 138 kDa (data not shown). Colorimetric detection yielded an intense purple band migrating with the same electrophoretic mobility as revealed from Coomassie Blue staining. Molecular weight estimation from DLS of *Taq* FBP aldolase yielded a value of 113 kDa at 22°C and 114 kDa at 45°C. On gel filtration the *Taq* FBP aldolase eluted at same place as the tetrameric rabbit muscle aldolase of 158 kDa (data not shown). These data support a homotetrameric quaternary structure of 132 kDa for the *Taq* FBP aldolase as was previously reported [6,7]. The reduction of the surface area caused by the tetramerization

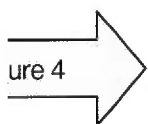


could contribute to the thermostability of the *Taq* protein [34].

It is intriguing that colorimetric detection shows the FBP aldolase from *Helicobacter pylori* to migrate at the same electrophoretic mobility with *Taq* aldolase on non-denaturing gels (Julie Allard, private communication). This result indicates that the quaternary structure of the *H.pylori* protein, which also contains a 21 amino acids insertion sequence, is tetrameric. Further the cyanobacteria *Anacystis nidulans* (now classified as *Synechococcus* sp. PCC6301) is phylogenetically related to *Synechocystis* whose FBP aldolase contains the 21 residue insertion sequence (figure 2) and which also has a tetrameric quaternary structure [8]. *Saprospira thermalis*, a flexibacteria that was found in a hot spring in Iceland and have evident affinities with cyanobacteria [35], also has a tetrameric class II FBP aldolase [8]. In contrast, FBP aldolases of class II A (*E. coli* [36], *S. cerevisiae* [37]) as well as class II B (*B. stearothermophilus* [38]) that do not contain this sequence insertion are dimers. It may be speculated that this sequence insertion stabilizes the tetrameric quaternary structure of FBP aldolases.

#### **Kinetic properties of *Taq* FBP aldolase**

Kinetic measurements were performed at 70°C, the natural growth temperature of *Thermus aquaticus*. At 70°C, the optimal pH is 6,5 (pH corrected to 70°C) as was reported previously [7]. The Co<sup>2+</sup> metal ion cofactor is readily exchangeable, enzymatic activity is lost upon dialysis of the recombinant enzyme in presence of EDTA, and can be restored upon addition of CoCl<sub>2</sub> following dialysis to remove EDTA. Activity in presence of ZnCl<sub>2</sub> instead of CoCl<sub>2</sub> is reduced by ~10 fold. This is consistent with the reported ion preference of *Taq* aldolase [6,7]. Non-linear regression analysis of the kinetic data using the Hill relation, shown in figure 4, yields a maximal specific activity,  $V_{\max}$ , for *Taq* FBP aldolase of 46 U/mg, a  $K_m$  of



$K_m$  of 305  $\mu\text{M}$  and a Hill coefficient of 0,32 at 70°C. Double reciprocal plot of enzyme kinetic data, shown in inset of figure 4, is consistent with strong negative cooperativity. The high negative cooperativity may be related to the 21 amino acid insertion observed near the potential DHAP binding site of *Taq* FBP aldolase.

The activity of *Taq* FBP aldolase is negligible below 60°C and at temperatures exceeding 80°C, instability of the cleavage products (especially GAP) compromises the kinetic data (figure 5). The Arrhenius plot of the enzyme (inset of figure 5) corresponds to an activation energy of 84 kJ/mol.



Figure 5

#### **Crystallization FBP aldolase from *Thermus aquaticus***

The *Taq* FBP aldolase has a strong tendency to form needle-like crystals in many precipitating agents such as ammonium sulfate, PEG of various sizes, MPD, ethanol, methanol, isopropanol. These needles are often observed aggregated as larger clusters. To improve protein crystal growth and crystal morphology, use of various additives were examined.

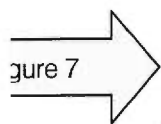
The best single crystals appeared within 2-3 weeks at room temperature (23°C) in a final solution containing 3mg/ml protein; 1,7 M ammonium sulfate; 0,1 M Tris-HCl pH 7,5; 1 mM  $\text{CoCl}_2$  and 0,4 mM sucrose monolaurate. A resultant crystal was washed, dissolved, and loaded onto a SDS-PAGE gel. One major band corresponding to the *Taq* FBP aldolase was



Figure 6

observed (data not shown) indicating crystallization of *Taq* FBP aldolase. Crystals of 0,1 mm X 0,1 mm X 1,5 mm in dimension (figure 6) were cryoprotected in mother liquor to which had been added 20% glycerol before freezing in liquid nitrogen.

### Preliminary crystallographic data



Diffraction data were collected at the NSLS synchrotron at 100 °K. The crystals diffract beyond 2Å, as shown in figure 7, and have space group P2 with unit-cell parameters of  $a = 99.5\text{Å}$ ,  $b = 57.6\text{Å}$  and  $c = 138.6\text{Å}$ , and  $\beta = 90.25^\circ$ . The unit-cell volume is  $793647\text{Å}^3$  and corresponds to a tetramer in the asymmetric unit cell. The specific volume of the unit cell ( $V_M$  - Matthews' factor) is  $3,0\text{Å}^3/\text{Da}$ , which corresponds to a solvent content of 59%(v/v) [39]. The structural analysis should lead to the determination of the first structure of a class IIB FBP aldolase novel because of its tetrameric quaternary structure and thermostability.

### Acknowledgements

This work was supported by the Medical Research Council of Canada. The sequence in this paper has been submitted to Genbank under the accession number AF052204.

### References

- [1] Brock, T. D. & Freeze, H. (1969) *Thermus aquaticus* gen. n. and sp. n., a non-sporulating extreme thermophile. *J. Bacteriol.*, **98**, 289-297.
- [2] Brock, T. D. (1986) Introduction: an overview of the thermophiles. In *Thermophiles: General, molecular, and applied microbiology* (Thomas D. Brock edition) pp1-16, J. Wiley & Sons, NY.
- [3] Haney, P. J., Badger, J. H., Buldak, G. L., Reich, C. I., Woese, C. R. & Olsen, G. J. (1999) Thermal adaptation analyzed by comparison of protein sequences from mesophilic and extremely thermophilic *Methanococcus* species. *Proc. Natl. Acad. Sci.*, **96**, 3578-3583.

- [4] Rutter, W.J. (1964) Evolution of aldolase. *Fed. Proc.*, **23**, 1248-57.
- [5] Plaumann, M., Pelzer-Reith, B., Martin, W. F. & Schnarrenberger, C. (1997) Multiple recruitment of class-I aldolase to chloroplasts and eubacterial origin of eukaryotic class-II aldolases revealed by cDNAs from *Euglena gracilis*. *Curr. genet.*, **31**, 430-438.
- [6] Freeze, H. & Brock, T. D. (1970) Thermostable aldolase from *Thermus aquaticus*. *J. Bacteriol.*, **101**, 541-550.
- [7] De Montigny, C & Sygusch, J. (1996) Functional characterization of an extreme thermophilic class II fructose-1,6-bisphosphate aldolase. *Eur. J. Biochem.*, **241**, 243-248.
- [8] Willard, J. M. & Gibbs, M. (1968) Purification and characterization of the fructose diphosphate aldolases from *Anacystis nidulans* and *Saprospira thermalis*. *Biochim. Biophys. Acta*, **151**, 438-448.
- [9] Blom, N. S., Tétreault, S., Coulombe, R. & Sygusch, J. (1996) Novel active site in *Escherichia coli* fructose 1,6-bisphosphate aldolase. *Nature struct. biol.*, **3**, 856-862.
- [10] Qamar, S. Marshal, K. & Berry, A. (1996) Identification of arginine 331 as an important active site residue in the class II fructose-1,6-bisphosphate aldolase of *Escherichia coli*. *Protein Sci.*, **5**, 154-161.
- [11] Plater, A. R., Zgiby, S. M., Thomson, G. J., Qamar, S., Wharton, C. W. & Berry, A. (1999) Conserved residues in the mechanism of the *E.coli* class II FBP-aldolase. *J. Mol. Biol.*, **285**, 843-855.
- [12] Hall, D. R., Leonard, G. A., Reed, C. D., Watt, C. I., Berry, A & Hunter, W. N. (1999) The crystal structure of *Escherichia coli* class II fructose-1,6-bisphosphate aldolase in complex with phosphoglycolohydroxamate reveals details of mechanism and specificity. *J. Mol. Biol.*, **287**, 383-394.

- [13] Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., Struhl, K. S. (1987-1988) *Current protocols in molecular biology*, 1<sup>st</sup> edn, Greene publishing associates and Wiley-Interscience, John Wiley & Sons , NY.
- [14] Myers, T.W. and Romano, L.J. (1988) Mechanism of stimulation of T7 DNA polymerase by *Escherichia coli* single-stranded DNA binding protein (SSB). *J. Biol. Chem.*, **263**, 17006-17015.
- [15] Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989) *Molecular cloning: a laboratory manual*, 2nd edn, Cold Spring Harbor Laboratory Press, Plainview NY.
- [16] Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680-685.
- [17] Meijer, A. E. F. H. (1985) The histochemical demonstration of fructose diphosphate aldolase activity using a semipermeable membrane technique. *Histochem. J.*, **17**, 1271-1275.
- [18] Racker, E. (1947) Spectrophotometric measurement of hexokinase and phosphohexokinase activity. *J. Biol. Chem.*, **167**, 843-854.
- [19] Good, N. E., Winget, G. D., Winter, W., Connolly, T. N., Izawa, S. & Singh, R. M. M. (1966) Hydrogen ion buffers for biological research. *Biochemistry*, **5**, 467-477.
- [20] Jancarik, J. & Kim, S.H. (1991) Sparse-matrix sampling- a screening method for crystallization of proteins. *J. Appl. Crystal.*, **24**, 409-411.
- [21] Ducruix, A & Giegé, R. (1992) Methods of crystallization. In *Crystallization of nucleic acids and proteins: A practical approach* (Ducruix, A & Giegé, R., eds), pp73-98. IRL Press, NY.
- [22] Otwinowski, Z. & Minor, W. (1997) Processing of X-rays diffraction data collected in oscillation mode, in *Methods in enzymology*, Vol. 276: Macromolecular crystallography, part

A (C. W. Carter, Jr & R. M. Sweet, eds) pp.307-326, Academic Press, New York.

[23] <http://www.dna.affrc.go.jp/nakamura/codon.html>

[24] Danson, M. J. & Hough, D. W. (1998). Structure, function and stability of enzymes from the Archea. *Trends microbiol.*, **6**, 307-314.

[25] Branden, C. & Tooze, J. (1991) *Introduction to protein structure*. Garland publishing, Inc. NY.

[26] Argos, P., Rossmann, M. G., Grau, U. M., Zuber, H., Frank, G. & Tratschin, J. D. (1979) Thermal stability and protein structure. *Biochemistry*, **18**, 5698-5703.

[27] Menendez-Arias, L. & Argos, P. (1989) Engineering protein thermal stability: sequence statistics point to residue substitutions in  $\alpha$ -helices. *J. Mol. Biol.*, **206**, 397-406.

[28] Wilquet, V. & Van de Casteele, M. (1999) The role of the first letter in the relationship between genomic GC content and protein amino acid composition. *Res. Microbiol.*, **150**, 21-32.

[29] Mrabet, N. T., Van den Broeck, A., Van den Brande, I., Stanssens, P., Laroche, Y., Lambeir, A.-M., Matthijssens, G., Jenkins, J., Chiadmi, M., Van Tilbeurgh, H., Rey, F., Janin, J., Quax, W. J., Lasters, I., De Maeyer, M. & Wodak, S. J. (1992) Arginine residues as stabilizing elements in proteins. *Biochemistry*, **31**, 2239-2253.

[30] Merkler, D. J., Farrington, G. K. & Wedler, F. C. (1981) Protein thermostability. *Int. J. Pept. Prot. Res.*, **18**, 430-442.

[31] Flechner, A., Gross, W., Martin, W. F. & Schnarrenberger (1999) Chloroplast class I and class II aldolases are bifunctional for fructose-1,6-biphosphate and sedoheptulose-1,7-biphosphate cleavage in the Calvin cycle. *FEBS letters*, **447**, 200-202.

[32] <http://www.cmpharm.ucsf.edu/~nomi/nnpredict.html>

- [33] Roger, A. J., Morrison, H. G. & Sogin, M. C. (1999) Primary structure and phylogenetic relationship of a malate dehydrogenase gene from *Giardia lamblia*. *J. Mol. Evol.*, **48**, 750-755.
- [34] Steelwagen, E. & Wilgus, H. (1978) Relationship of protein thermostability to accessible surface area. *Nature*, **275**, 342-343.
- [35] Lewin, R. A. (1965) Isolation and some physiological features of *Saprospira thermalis*. *Can. J. Microbiol.*, **11**, 77-86.
- [36] Baldwin, S. A., Perham, R. N. & Stribling, D. (1978) Purification and characterization of the class-II D-fructose 1,6-bisphosphate aldolase from *Escherichia coli* (Crookes' strain). *Biochem. J.*, **169**, 633-641.
- [37] Jack, R. S. & Harris, J. I. (1971) The primary structure of yeast aldolase. *Biochem. J.*, **124**, 68P-69P.
- [38] Sugimoto, S. & Nosoh, Y. (1971) Thermal properties of fructose-1,6-diphosphate aldolase from thermophilic bacteria. *Biochim. Biophys. Acta*, **235**, 210-221.
- [39] Matthews, B. W. (1968) Solvent content of protein crystals. *J. Mol. Biol.*, **33**, 491-497.
- [40] Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, **22**, 4673-4680.

**Table 1. Summary of purification of *Taq* FBP aldolase from a 1-l culture of *E. coli*-pKKaldoTaq.** Activity was assayed at 70°C in 5 mM FBP; 1 mM CoCl<sub>2</sub>; 50 mM Tris-HCl pH 6,5 at 70°C, according to protocol described in the Material and Methods section.

Purification step	Total protein (mg)	Total activity (U)	Specific activity (U/mg)	Yield (%)
Crude extract	387,4	5418,5	14,0	100
Delipidation	253,0	4385,2	17,3	80,9
Heat treatment	79,2	1915,4	24,2	35,3
Phosphocellulose	57,4	1871,7	32,6	34,5
Q-Sepharose	47,5	1845,0	38,8	34,0



**Figure 1. Complete nucleotide sequence and deduced amino acids sequence of the *Taq* fructose-1,6-bisphosphate aldolase.** Numbers in parentheses represent the nucleotide sequence and regular numbers refer to the amino acids sequence of the enzyme. The initiation and the termination codons are underlined. The "stop" codon is shown by three stars (\*\*\*)

(1) GGCCTACTGAGTTTCCACGCTTCCATAAGGGTCCGGGTATAATCCGGGCCGGGGGTGCTT

(61) ATGCTGGTTACGGGTCTAGAGATCCTGAAGAAGGCGCGGGAAGAGGGCTACGGGGTGGGG  
 1 M L V T G L E I L K K A R E E G Y G V G

(121) GCCTTTAACGTGAACAACATGGAGTTCCTCCAGGCGGTCTCGAGGCCGCCGAGGAGCAA  
 21 A F N V N N M E F L Q A V L E A A E E Q

(181) AGAAGCCCGGTCATCCTGGCCCTCTCCGAGGGGGCCATGAAGTACGGGGGGCGGGCCCTC  
 41 R S P V I L A L S E G A M K Y G G R A L

(241) ACCCTCATGGCCGTGGAGCTGGCCAAGGAGGCCCGGGTGCCGGTGGCCGTGCACCTGGAC  
 61 T L M A V E L A K E A R V P V A V H L D

(301) CACGGCAGTAGCTACGAGAGCGTGCTGAGGGCCCTTCGGGCGGGCTTCACCAGCGTCATG  
 81 H G S S Y E S V L R A L R A G F T S V M

(361) ATTGACAAGTCCCACGAGGACTTTGAGACCAACGTGCGGGAGACCCGCGGGGTGGTGGAG  
 101 I D K S H E D F E T N V R E T R R V V E

(421) GCGGCCACGCCGTGGGGGTACGGTGGAGGCGGAGCTGGGGCGGCTTGCTGGCATAGAG  
 121 A A H A V G V T V E A E L G R L A G I E

(481) GAGCACGTGGCCGTGGACGAGAAGGACGCCCTCCTCACCAACCCCGAGGAGGCCCGCATC  
 141 E H V A V D E K D A L L T N P E E A R I

(541) TTCATGGAGAGGACGGGGGCCGACTACCTGGCGGTGGCCATCGGCACCAGCCACGGGGCC  
 161 F M E R T G A D Y L A V A I G T S H G A

(601) TACAAGGGGAAGGGCAGGCCCTTCATTGACCACGCCCGGCTGGAAAGGATCGCCAGGCTG  
 181 Y K G K G R P F I D H A R L E R I A R L

(661) GTGCCCCTCCTTTGGTCCTCCACGGGGCCAGCGCCGTGCCCCCTGAGCTGGTGGAGCGC  
 201 V P A P L V L H G A S A V P P E L V E R

(721) TTCCGAGCTTCGGGAGGGGAGATCGGCGAGGCGGCGGGGATCCACCCCGAGGACATAAAG  
 221 F R A S G G E I G E A A G I H P E D I K

(781) AAGGCCATCTCCCTGGGTATCGCCAAGATCAACACCGACACCGACCTGCGCCTGGCCTTC  
 241 K A I S L G I A K I N T D T D L R L A F

(841) ACCGCCCTCATCCGGGAGGCCCTCAACAAGAACCCCAAGGAGTTTGACCCCGGAAGTAC  
 261 T A L I R E A L N K N P K E F D P R K Y

(901) CTGGGCCACGCCGGGAGGCGGTGAAGGAGGTGGTGAAGGCGCATGGAGCTCTTCGGC  
 281 L G P A R E A V K E V V K S R M E L F G

(961) TCCGTGGGCCGGGCCTAAGTACCCCGCGGTGGCGAGAGCCCCAAAGAAGCTATGGGCGGC  
 301 S V G R A \*\*\*

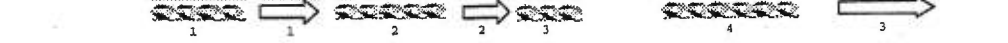
**Figure 2. Sequence alignment of class II FBP aldolases.** The sequence alignment was performed using CLUSTAL W [40]. Sequence were obtained from different databases; *Thermotoga maritima* AAD35362, *Aquifex aeolicus* AE000737, *Helicobacter pylori* J99 AAD05734, *Giardia lamblia* O97447, *Rhodobacter sphaeroides* 2 P29271, *Synechocystis* Q55664, *Xanthobacter flavus* U29134, *Alcaligenes eutrophus* 1 Q59101, *Bacillus subtilis* 1 P13243, *Bacillus stearothermophilus* Y11135, *Mycoplasma pneumoniae* P75089, *E.coli* GatY P37192, *E.coli* AgaY P42908, *Mycobacterium tuberculosis* Z95324, *Corynebacterium glutamicum* X17313, *Streptomyces galbus* CAA10483, *Saccharomyces cerevisiae* P14540, *Euglena gracilis* X89769, *Campylobacter jejuni* P53818, *Haemophilus influenzae* P44429, *E.coli* fba P11604. Dashes indicate gaps. The  $\alpha$ -helices and  $\beta$ -strands from the structure of *E. coli* FBP aldolase are shown by ribbons and arrows. The banner represents a 21 amino acids insertion. The stars indicate the conserved amino acids mentioned in the text.

Class II B

Thermus aquaticus
Thermotoga maritima
Aquifex aeolicus
Helicobacter pylori J99
Giardia lamblia
Rhodobacter sphaeroides 2
Synecocystis
Xanthobacter flavus
Alcaligenes eutrophus 1
Bacillus subtilis 1
Bacillus stearothermophilus
Mycoplasma pneumoniae
Escherichia coli GatY
Escherichia coli AgaY

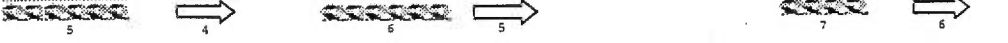
Class II A

Mycopacterium tuberculosis
Corynebacterium glutaminum
Streptomyces galbus
Saccharomyces cerevisiae
Euglena gracilis
Campylobacter jejuni
Haemophilus influenzae
Escherichia coli fda



Thermus aquaticus
Thermotoga maritima
Aquifex aeolicus
Helicobacter pylori J99
Giardia lamblia
Rhodobacter sphaeroides 2
Synecocystis
Xanthobacter flavus
Alcaligenes eutrophus 1
Bacillus subtilis 1
Bacillus stearothermophilus
Mycoplasma pneumoniae
Escherichia coli GatY
Escherichia coli AgaY

Mycopacterium tuberculosis
Corynebacterium glutaminum
Streptomyces galbus
Saccharomyces cerevisiae
Euglena gracilis
Campylobacter jejuni
Haemophilus influenzae
Escherichia coli fda



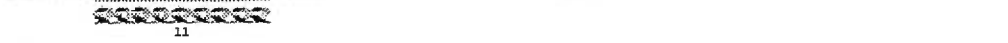
Thermus aquaticus
Thermotoga maritima
Aquifex aeolicus
Helicobacter pylori J99
Giardia lamblia
Rhodobacter sphaeroides 2
Synecocystis
Xanthobacter flavus
Alcaligenes eutrophus 1
Bacillus subtilis 1
Bacillus stearothermophilus
Mycoplasma pneumoniae
Escherichia coli GatY
Escherichia coli AgaY

Mycopacterium tuberculosis
Corynebacterium glutaminum
Streptomyces galbus
Saccharomyces cerevisiae
Euglena gracilis
Campylobacter jejuni
Haemophilus influenzae
Escherichia coli fda

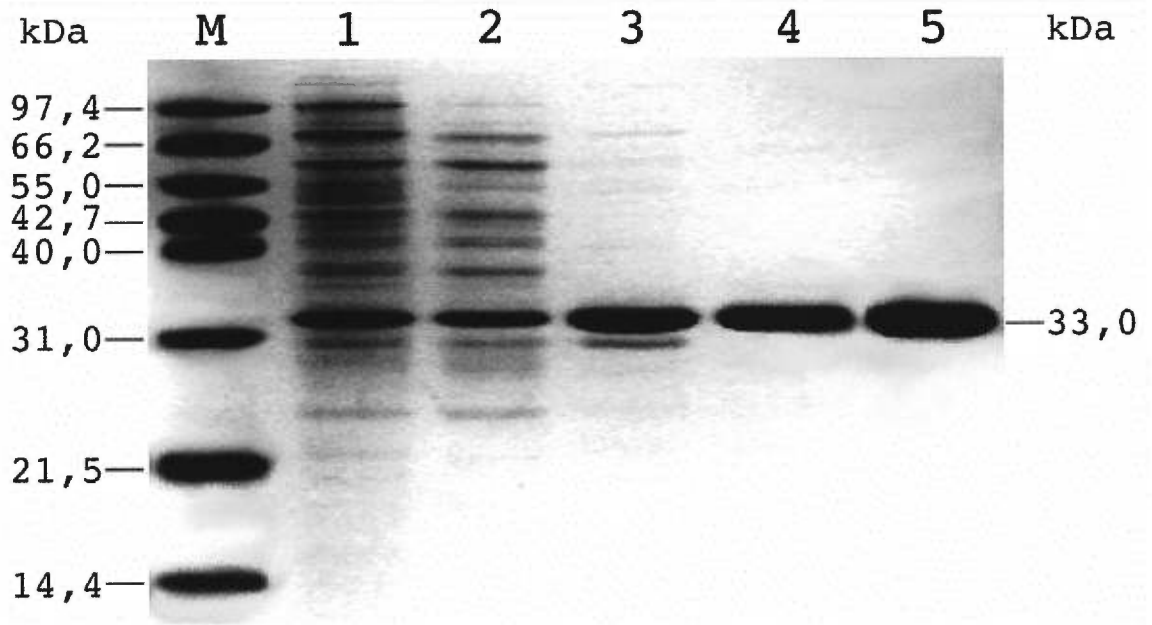


Thermus aquaticus
Thermotoga maritima
Aquifex aeolicus
Helicobacter pylori J99
Giardia lamblia
Rhodobacter sphaeroides 2
Synecocystis
Xanthobacter flavus
Alcaligenes eutrophus 1
Bacillus subtilis 1
Bacillus stearothermophilus
Mycoplasma pneumoniae
Escherichia coli GatY
Escherichia coli AgaY

Mycopacterium tuberculosis
Corynebacterium glutaminum
Streptomyces galbus
Saccharomyces cerevisiae
Euglena gracilis
Campylobacter jejuni
Haemophilus influenzae
Escherichia coli fda

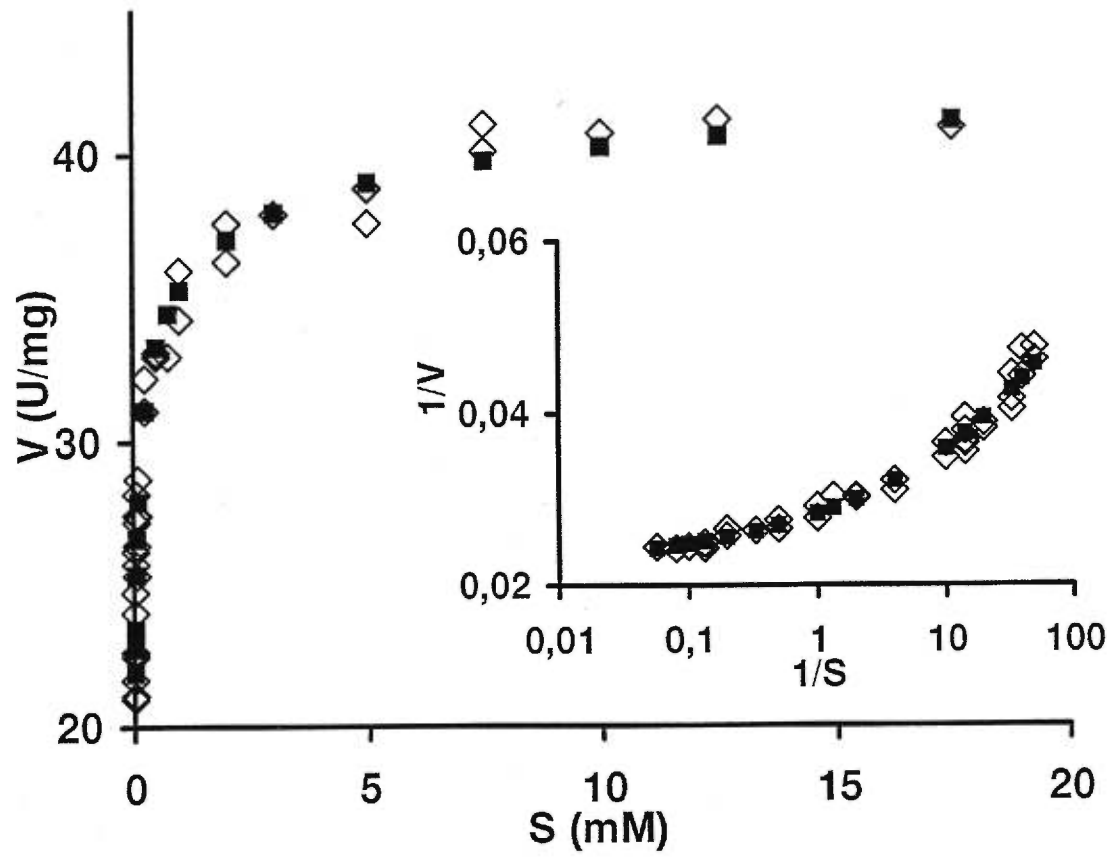


**Figure 3. SDS/PAGE of the purification of *Taq* FBP aldolase produced in *E.coli* transformed by pKKaldoTaq.** Lane M, contains protein molecular weight markers; lane 1, crude extract; lane 2, crude extract after delipidation; lane 3, crude extract after heat treatment; lane 4, active fraction following affinity elution from phosphocellulose P11; lane 5, active fraction after Q-Sepharose.



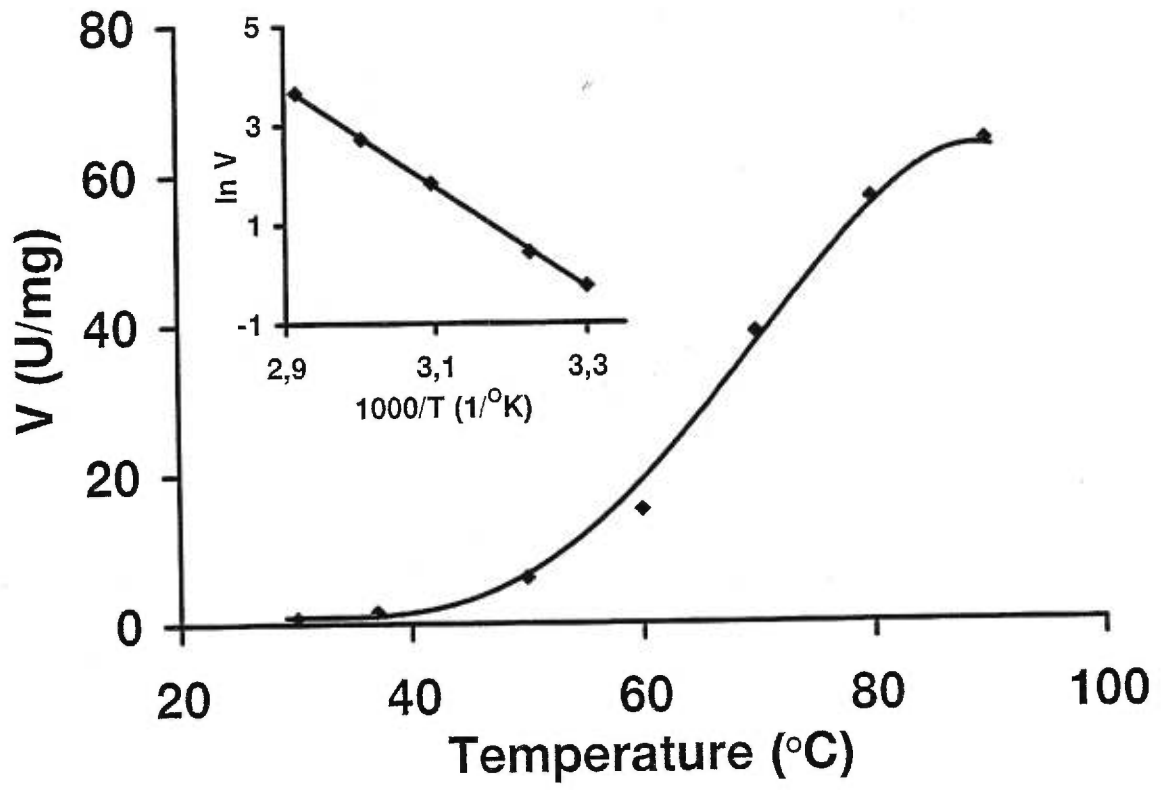
**Figure 4. Non-linear regression analysis of the kinetic data using the Hill**

**relation.** The reaction was performed at 70°C in 50 mM Tris-HCl pH 6,5 (at 70°C); 1 mM CoCl<sub>2</sub> containing different concentration of FBP. The experimental values are represented by the symbol  $\diamond$ , and calculated values by  $\blacksquare$ . **Inset**, double reciprocal plot of enzyme kinetic data using a logarithmic scale for the substrate concentration values.





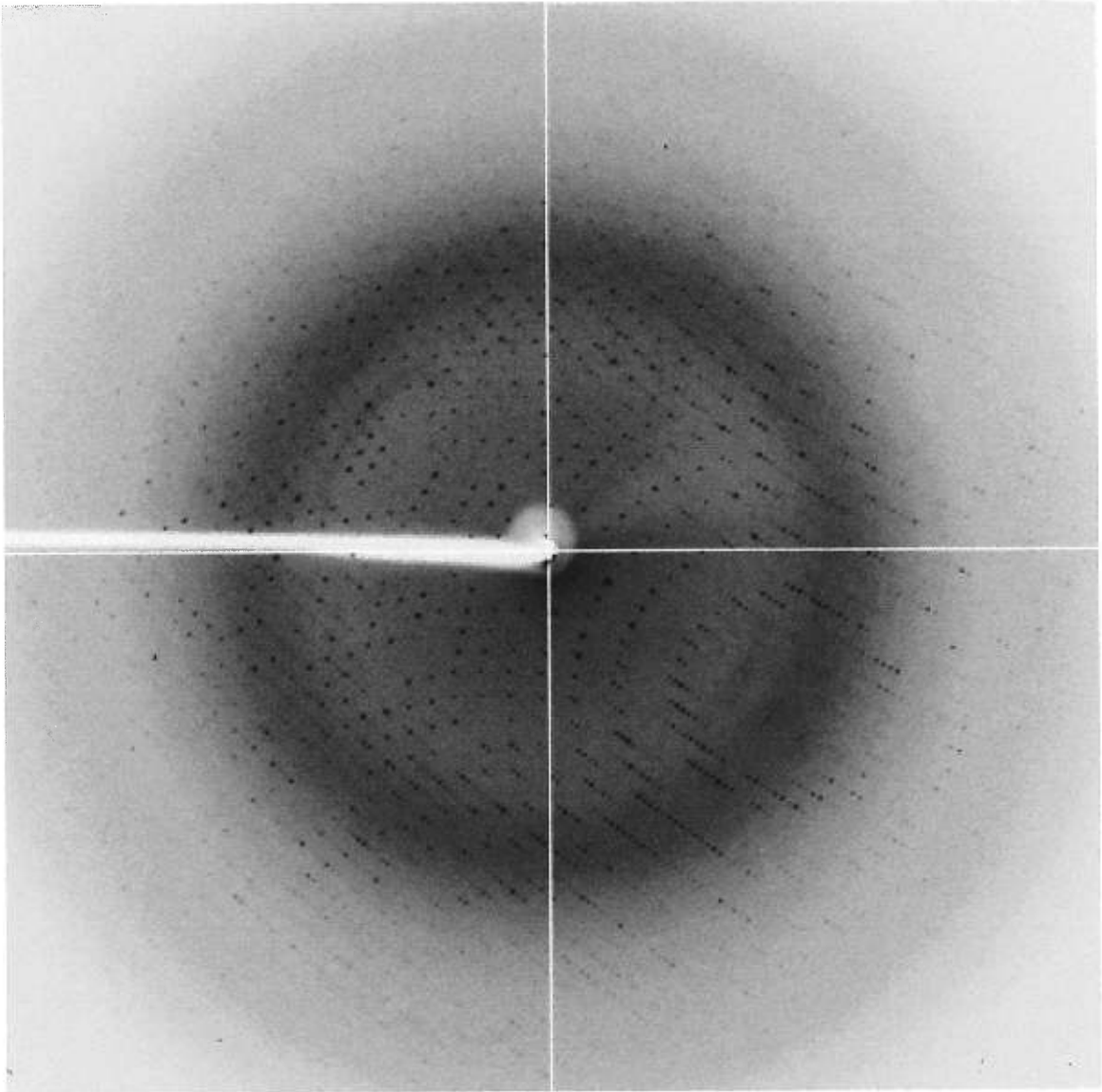
**Figure 5. Effect of temperature on *Taq* FBP aldolase activity.** The reaction was performed as described in Material and methods in 5 mM FBP; 1 mM CoCl<sub>2</sub> and 50 mM Tris-HCl pH 6,5 buffered at different temperatures. **Inset**, Arrhenius plot of the activity of the enzyme used to calculate the activation energy. The values at 80°C and 90°C were removed because of their underestimation due to GAP instability at those temperatures.



**Figure 6. Photography of a *Taq* FBP aldolase crystal.** A crystal of dimensions 0,1 mm X 0,1 mm in cross-section and 1,5 mm in length was grown in a sitting drop by vapour diffusion and equilibrated against 1,7 M  $\text{AmSO}_4$ ; 100 mM Tris-HCl pH 7,5; 10 mM  $\text{CoCl}_2$  and 0,4 mM sucrose monolaurate.



**Figure 7. Diffraction pattern of a flash frozen monoclinic *Taq* aldolase crystal.** The data was collected on a Quantum-4 CCD detector (ADSC) using synchrotron radiation of X-8C beam line at Brookhaven National Laboratory. The exposure time was 40 seconds with an oscillation angle of  $0,66^\circ$  at a detector distance of 150 mm. Intensity data extends to within  $2\text{\AA}$  resolution and represented by the edge of the film.



## **Chapitre 3: Matériel et méthodes**

### 3.1 Les souches et les produits

Les bactéries compétentes *E.coli* JM109 proviennent de Promega. L'ampicilline et la séléno-L-méthionine ont été obtenues de Sigma, l'IPTG et le DTT, de Roche Diagnostics, le sulfate d'ammonium de Gibco BRL, le chlorure d'Yttrium de Fluka, le sucrose monolaurate de Hampton Research et tous les autres réactifs de même que les Microcons 3 de Fisher Scientifique ltée. La colonne HiLoad 16/60 Superdex 200 et le FPLC viennent de Amersham-Pharmacia.

### 3.2 La production et la purification de la FBP aldolase de *Taq* séléno-méthionée

Le plasmide pKKaldoTaq contenant le gène de la FBP aldolase de *Taq*, a été introduit par choc thermique, dans des *E.coli* JM109 compétentes qui ont ensuite été étalées sur des pétris DYT-Ampicilline (100 µg/ml) placés à 37°C pour toute la nuit. Une culture de 30 ml de DYT-Amp (100 µg/ml) a été inoculée avec une colonie transformante puis incubée à 37°C avec agitation jusqu'à ce que la DO à 600 nm atteigne 0,8. Elle a alors été centrifugée à 3000g à 4°C pendant 15 min. Le culot de bactéries a été resuspendu dans 15 ml de milieu minimal NMM. Chaque cinq ml de cette resuspension ont été utilisés pour partir au total trois cultures de 500 ml de milieu M9 additionné de 1 mM MgSO<sub>4</sub>; 21,6 µM CaCl<sub>2</sub>; 20 mM glucose; 17,8 µM (NH<sub>4</sub>)<sub>2</sub>Fe(SO<sub>4</sub>)<sub>2</sub>; 29,6 µM thiamine (vitamine B1); 40,9 µM biotine; 17,0 nM CuCl<sub>2</sub>; 18,2 nM MnCl<sub>2</sub>; 15,4 nM ZnCl<sub>2</sub>; 10,3 nM Na<sub>2</sub>MoO<sub>4</sub>; 50mg/l de valine, d'isoleucine et de leucine; 100 mg/l de thréonine, phénylalanine et lysine et 60 mg/l de séléno-L-méthionine. Les trois bouillons ont été placés toute la nuit à 37°C avec agitation. Le lendemain matin, 500 ml de milieu M9 frais contenant les composants cités précédemment, ont été divisés entre les trois cultures et de l'ampicilline (100 µg/ml) leur a été ajoutée. La croissance des bactéries à 37°C a été poursuivie jusqu'à ce que la DO à 600 nm atteigne 0,8. L'IPTG a alors été ajouté à une concentration finale de 0,5 mM. La production de protéines a été induite pendant sept heures. Les cultures ont ensuite été centrifugées à 8000g à 4°C pendant 15 minutes. Le culot a été resuspendu dans 100 ml de tampon de lyse (50 mM Tris-



HCl pH 7,3; 0,1 mM PMSF; 1 mM DTT (pour éviter l'oxydation)), puis soniqué treize fois 30 secondes avec 20 secondes de repos entre chaque sonication. L'extrait brut a été centrifugé à 35000g à 4°C pendant 15 minutes. Le pH du surnageant a été abaissé à 5,4 avec de l'acide acétique glacial. La solution turbide a été centrifugée à 35000g à 4°C pendant 15 minutes. Le pH du surnageant a été ramené à 7,3 avec du NaOH. L'extrait protéique a été dilué jusqu'à 200 ml avec du tampon de lyse, puis chauffé 45 minutes à 80°C avec une légère agitation. La solution a ensuite été centrifugée à 40000g à 4°C pendant 20 minutes et conservée à 4°C. À l'aide d'un Microcon 3, 100 µg de la protéine modifiée ont été concentrés à 1 mg/ml et transférés dans un tampon de 10 mM de bicarbonate d'ammonium pour une analyse par spectrométrie de masse (analyse réalisée à l'Université de Leichester en Angleterre par le laboratoire de Mme T. Izard).

### **3.3 Les conditions de cristallisation et de cryoprotection pour la FBP aldolase de *Taq* native et séléno-méthionée à pH 4,0**

La technique de goutte suspendue a permis de cristalliser la FBP aldolase de *Taq* dans des conditions où le puits contenait 1 mL de 0,6 M de sulfate d'ammonium; 20 mM d'acide citrique-NaOH pH 4,0 et la goutte, 2 µL de solution précipitante et 5 µL de FBP aldolase de *Taq* (native ou séléno-méthionée) à 0,25 mg/ml à la température de la pièce. Les cristaux obtenus dans cette condition ont été cryoprotégés avant leur congélation dans l'azote liquide, par un trempage d'environ 15 minutes dans une solution contenant 0,6 M de sulfate d'ammonium; 20 mM d'acide citrique-NaOH pH 4,0 et 15 % de glycérol.

### **3.4 D'autres conditions de cristallisation de la FBP aldolase de *Taq***

D'autres conditions de cristallisation ont permis d'obtenir de petits cristaux de FBP aldolase de *Taq* native à température de la pièce grâce à la technique de goutte suspendue. Le

réservoir contenait 1 ml d'une des solutions précipitantes suivantes:

48% éthanol; 0,1M Hepes pH 7,75

ou 55% méthanol; 0,1M Hepes pH 7,5; 200 mM d'acétate d'ammonium

ou 12,5% PEG 3350; NaI

ou 35-37% isopropanol; 0,1 M Hepes pH 7,25; 0,3M d'acétate d'ammonium

ou 26-29% MPD; 0,1 M Mes pH 6,5; 5 mM d'acétate d'ammonium

et les gouttes suspendues étaient constituées de 2 $\mu$ l de solution précipitante et 2 $\mu$ l de FBP aldolase de Taq à 3,2 mg/ml.

### **3.5 Le trempage de cristaux de FBP aldolase native de Taq dans l'yttrium et leur condition de cryoprotection**

Certains cristaux obtenus en présence de 1,675 M de sulfate d'ammonium; 0,1 M Tris-HCl pH 7,5; 10 mM CoCl<sub>2</sub> et 0,4 mM de sucrose monolaurate (pour plus de détails voir l'article page 53) ont été trempés dans une solution de 1,675 M de sulfate d'ammonium; 0,1 M Tris-HCl pH 7,5 et 20 mM YCl<sub>3</sub> pendant 30 minutes. Avant leur congélation dans l'azote liquide, ces cristaux ont été trempés dans 1,675 M de sulfate d'ammonium; 0,1 M Tris-HCl pH 7,5; 20 mM YCl<sub>3</sub> et 20 % de glycérol pendant 15 minutes

### **3.6 La collecte de données cristallographiques**

Les données cristallographiques ont été collectées à la station X-8C du "National Synchrotron Light Source" au "Brookhaven National Laboratory" avec un détecteur CCD Quantum 4R. Pour les cristaux de FBP aldolase de Taq sélénométhionée, la distance du détecteur utilisée était de 240 mm et le temps d'exposition de 70 secondes avec une oscillation de 0,66°. Les longueurs d'onde utilisées étaient de 0,978318 Å au point de fluorescence maximale du sélénium, de 0,979056 Å au point d'inflexion et de 0,930002 Å au point d'absence de fluorescence du sélénium. Pour les cristaux trempés dans l'yttrium, les conditions de collecte utilisées sont celles décrites dans l'article (voir page 44) excepté le temps d'exposition qui a été de 30 secondes au lieu de 40. Les données ont été traitées avec les programmes DENZO et Scalepack<sup>97</sup>.

## **Chapitre 4: Discussion et conclusion**

## 4.1 Discussion

Tel que décrit dans l'article, le gène de la FBP aldolase de *Thermus aquaticus* a été identifié. Conformément à la tendance thermophile, son pourcentage en G/C est très élevé. Ses 67,3% de G/C ont même causé des problèmes de compression lors du séquençage. Du DMSO, des déaza A et G et des SSB ont du être ajoutés à la réaction pour défaire les structures secondaires. La séquence nucléotidique déterminée code pour une protéine de 305 acides aminés ayant un poids moléculaire apparent de 33 092 Da.

La plupart des substitutions d'acides aminés observées chez les protéines thermophiles sont retrouvées dans la FBP aldolase de *Taq*. Ainsi l'absence de cystéine, un nombre restreint de méthionines, de glutamines et d'asparagines, de même que les substitutions de sérines en alanines, d'isoleucines en valines, d'acides aspartiques en acides glutamiques et de lysines en arginines dénotent de sa thermostabilité. Son contenu réduit en résidus polaires non-chargés au profit des résidus chargés et hydrophobes devrait lui permettre à la fois la formation de ponts salins à sa surface pour une meilleure interaction entre les sous-unités, et l'ajout de liens hydrophobes dans les régions enfouies afin de les rendre plus compactes et moins accessibles pour le solvant.

L'alignement de la FBP aldolase de *Taq* avec les autres FBP aldolases de classe II montre qu'elle appartient à la sous-classe II B<sup>12</sup> et partage une forte homologie avec les FBP aldolases du groupe d'organismes extrémophiles. Comme les autres enzymes de classe II B, exceptés ceux des bactéries utilisant le cycle de Calvin, elle est plus courte que ceux des aldolases de classe II A à cause de raccourcissements de ses hélices et de ses boucles. Les FBP aldolases des cyanobactéries et des  $\alpha$  et  $\beta$ -protéobactéries qui sont de longueur similaire aux enzymes de classe II A, contiennent deux insertions (entre le feuillet  $\beta$  5 et l'hélice  $\alpha$  7 et entre le feuillet  $\beta$  4 et l'hélice  $\alpha$  6) qui leur sont exclusives et qui pourraient expliquer leur affinité pour le sédoheptulose-1,7-bisphosphate<sup>98</sup>, un sucre aussi impliqué dans le cycle de Calvin (figure 2). Elles contiennent une troisième insertion (entre le feuillet  $\beta$  7 et l'hélice  $\alpha$  9) aussi retrouvée chez le groupe d'extrémophiles dont fait partie *Thermus aquaticus*. Cette insertion de 21 acides aminés contiendrait une hélice  $\alpha$  selon des prédictions de structures secondaires<sup>99</sup>.

D'après l'alignement avec la structure de la FBP aldolase de classe II de *E.coli*, les acides aminés impliqués dans l'activité sont conservés dans l'enzyme de *Taq*.

L'hydrophobicité totale des résidus présumément situés en surface (établis par alignement avec la structure d'*E.coli* et en se fiant au tableau d'hydrophobicité établi par Kyte et Doolittle<sup>100</sup>) de la FBP aldolase de *Taq* est plus élevée que celle d'*E.coli*. Toutefois le même taux d'hydrophobicité de surface est retrouvé chez les autres classes II B. Cette augmentation d'hydrophobicité même si elle peut contribuer à la thermostabilité de l'enzyme, ne suffit pas à l'expliquer.

La présence de la FBP aldolase II A ou B chez un organisme pourrait dépendre des voies métaboliques énergétiques qu'il utilise. Les classes II A sont essentiellement employées dans la voie glycolytique Embden-Meyerhof-Parnas, puisque dans certains cas comme ceux d'*E.coli*<sup>22</sup> et d'*Euglena gracilis*<sup>12</sup>, le recours à la gluconéogenèse conduit à la production d'un enzyme de classe I. *E.coli* utilise aussi une voie Entner-Doudoroff linéaire inductible. Les classes II B sont principalement composée de bactéries autotrophes utilisant le cycle de Calvin, d'extrémophiles dont certains (*Aquifex*) sont autotrophes, de bactéries formant des endospores (les Bacilles par exemple) adaptées pour survivre à de longues périodes de latence sans apport énergétique, et finalement de protéines utilisant la tagatose-1,6-bisphosphate ou dont l'activité n'est pas encore bien définie. Ces organismes semblent tous utiliser leur FBP aldolase à la fois dans la voie glycolytique Embden-Meyerhof-Parnas et dans d'autres voie comme la gluconéogenèse, le cycle de Calvin ou la voie des tagatoses. Même que *Xanthobacter flavus* ne produit l'enzyme de classe II B qu'en conditions autotrophes<sup>101</sup>. De plus, *Alcaligenes eutrophus* a recours à la voie Entner-Doudoroff, mais sous la forme cyclique couplée au cycle du TCA et *Rhodobacter sphaeroides* emploierait une forme linéaire modifiée. Certaines expériences suggèrent que la FBP aldolase et/ou la fructose bisphosphatase seraient la différence-clé entre les organismes utilisant la voie métabolique Entner-Doudoroff linéaire et ceux employant la forme cyclique<sup>102</sup>.

Dans le but d'obtenir des quantités importantes de FBP aldolase de *Taq*, son gène a été sous-cloné dans le vecteur d'expression pKK223-3 pour ensuite surexprimer cette protéine dans

*E.coli*. Toutefois une concentration trop élevée de FBP aldolase de *Taq* dans les bactéries *E.coli* inhibe la croissance de celles-ci. L'enzyme de *Taq* réquisitionne peut-être le cobalt présent dans les bactéries-hôtes, n'en laissant plus pour les enzymes d'*E.coli* qui en requièrent. Étant donné son effet inhibiteur, la surexpression de la protéine thermophile a été placée sous le contrôle du promoteur inductible *tac* dans la souche d'*E.coli* JM109 produisant le répresseur de lactose *lacI<sup>q</sup>*. Ce système d'expression a une bonne efficacité puisque 40 à 50 mg de protéine purifiée sont obtenus à partir d'un litre de culture.

La purification de l'enzyme de *Taq* a été simplifiée par l'étape de traitement à la chaleur qui donne une protéine pure à 60% en moins d'une heure. La colonne de Phosphocellulose P11 est utilisée comme colonne d'affinité puisque la protéine de *Taq* est éluée avec son substrat, le FBP. Après cette étape la protéine est pure mais très diluée et contient du FBP ou ses produits de réaction, le DHAP et le GAP. La colonne échangeuse d'ions Q-Sepharose permet de concentrer la FBP aldolase de *Taq* et d'éliminer toutes traces de produits non-désirés. L'enzyme ainsi purifié est très pur comme en témoigne la présence d'une seule espèce homogène observée par DLS. Ce niveau de pureté est d'autant plus important que 81% des protéines déterminées monodisperses par DLS cristalliseraient. Des solutions protéiques polydisperses peuvent aussi donner des cristaux mais le taux de succès est de beaucoup réduit<sup>103</sup>.

Sur le tamis moléculaire, la FBP aldolase de *Taq* sort à la même position que la FBP aldolase de muscle de lapin connue comme étant tétramérique et ayant une masse moléculaire de 158 kDa. De plus la méthode de DLS permet aussi une estimation du poids moléculaire apparent des protéines. Pour la FBP aldolase de *Taq* cette estimation est de 113-114 kDa. Sur gel natif l'estimation du poids moléculaire de la FBP aldolase de *Taq* est de 138 kDa. Ces trois données dénotent d'une structure quaternaire tétramérique et sont consistantes avec la forme tétramérique observée pour la protéine purifiée directement de *Taq*<sup>29</sup>.

Fait intéressant, des mutants de FBP aldolase de muscle de lapin<sup>73</sup> ont été construits de façon à obtenir l'enzyme sous une forme dimérique active. Chez ces mutants une perte de thermostabilité a été dénotée. La tétramérisation pourrait donc contribuer à la thermostabilité de la FBP aldolase de *Taq*, notamment en diminuant son ratio surface/volume et par

conséquent, son exposition au solvant.

Autre point d'intérêt, trois autres FBP aldolases de classe II sont tétramériques, soient celle de la cyanobactérie *Anacystis nidulans*, de la flexibactérie *Saprospira thermalis*<sup>30</sup> et d'*Helicobacter pylori* (Julie Allard, communication personnelle). Or *H.pylori* qui appartient à la même sous-classe d'extrémophiles que *Taq*, de même que *Synechocystis*, une cyanobactérie au même titre qu'*A.nidulans* et partageant certaines caractéristiques avec les flexibactéries<sup>104</sup>, contiennent aussi l'insertion de 21 acides aminés. Toutes les autres FBP aldolases de classe II caractérisées sont dimériques et ne contiennent pas cette insertion. Il est donc très probable que ce rajout à la structure intervienne dans la tétramérisation de l'enzyme.

L'activité de la FBP aldolases de *Taq* a été dosée à 70°C, température correspondant à celle de la croissance des bactéries *Thermus aquaticus* et à laquelle les produits de la réaction (DHAP et GAP) restent stables suffisamment longtemps pour permettre leur dosage. À cette température, l'enzyme a une activité spécifique ( $V_{max}$ ) de 46 U/mg et une constante d'affinité ( $K_M$ ) pour le FBP de 305  $\mu$ M. Sa cinétique démontre une coopérativité négative entre ses sous-unités, caractérisée par un coefficient de Hill de 0,32. Cette caractérisation enzymatique est consistante avec les résultats obtenus antérieurement pour l'enzyme non-recombinant<sup>29</sup>. Dans le cas du tétramère de FBP aldolase de muscle de lapin, il est connu que le C-terminal de certaines sous-unités forme un clapet médiant la liaison du DHAP en bloquant ou non l'entrée du site actif<sup>105</sup>. L'insertion de 21 acides aminés retrouvée dans l'enzyme de *Taq* est située directement dans le site potentiel de liaison du DHAP pourrait elle aussi former une structure mobile qui contrôlerait l'accès du substrat dans le site actif. Cette hypothèse expliquerait la forte coopérativité négative de la protéine.

Lors de la collecte de données cristallographiques, la phase, un paramètre très important pour le traitement des données, est perdue. Différents techniques permettent de retrouver cette valeur dont celles de MAD ou de MIR. Pour utiliser la technique de MAD (multiple wavelength anomalous dispersion), un cristal d'une protéine contenant un atome lourd (comme le sélénium de la séléno-méthionine) est requis. La technique comme telle consiste à collecter une série de données à trois longueurs d'onde différentes, soit à celle à laquelle la

fluorescence de l'atome lourd est à son maximum, celle correspondant au point d'inflexion de la fluorescence et finalement une où l'atome lourd n'a plus aucune fluorescence. L'avantage de cette technique est qu'avec un seul cristal, il est possible d'avoir toutes les données nécessaires à la résolution de la structure tridimensionnelle. La technique de MIR (multiple isomorphous replacement) consiste à introduire par trempage des atomes lourds dans un cristal, puis à collecter une série de données à la longueur d'onde où la fluorescence de l'atome lourd est maximale afin de la comparer à une série de données d'un cristal n'ayant subi aucune modification. Toutefois avec cette technique le taux d'incorporation du métal lourd et son emplacement dans le cristal est moins prévisible et il faut généralement des données avec différents atomes lourds pour pouvoir résoudre la structure.

La FBP aldolase de *Taq* contenant des séléno-méthionines a été produite afin d'utiliser la technique de MAD. Pour s'assurer que seules des méthionines couplées au sélénium étaient incorporées dans la protéine, les bactéries ont proliféré dans un milieu minimal contenant de la séléno-méthionine et d'autres acides aminés qui inhibent la biosynthèse de la méthionine chez les non-autotrophes. La surexpression de la FBP aldolase de *Taq* dans ces conditions est diminuée (d'environ 30%) par rapport à celle obtenue dans les conditions décrites dans l'article. La spectrométrie de masse a confirmé l'incorporation de séléno-méthionine dans la protéine. L'enzyme séléno-marqué est encore actif.

Dans la condition de cristallisation sulfate d'ammonium-acide citrique pH 4,0, des cristaux bipyramidaux de 250  $\mu\text{m}$  X 160  $\mu\text{m}$  X 160  $\mu\text{m}$  de FBP aldolase de *Taq* native ou séléno-méthionée ont été obtenus après une semaine (figure 5). Un signal de fluorescence du sélénium a été observé à 0,978 Å (longueur d'onde pour un maximum de fluorescence du sélénium) à une distance de détecteur de 240 mm. Finalement une analyse des données a permis de déterminer que la maille de ce type de cristal était tétragonale  $I4_1$  avec des paramètres de maille de  $a = 88,6$  Å,  $b = 88,6$  Å,  $c = 164,1$  Å. Des réflexions ont été observées jusqu'à 2,6 Å avec 97,1% de complétude dans l'intervalle de résolution de 90 Å à 2,6 Å. Le nombre d'atomes de sélénium retrouvés par maille asymétrique correspond à la présence d'un dimère.





**Figure 5. Photographie de deux cristaux de FBP aldolase de *Taq* sélénométhionée** Ces cristaux mesurant 250  $\mu\text{m}$  entre les pointes des deux pyramides et 160  $\mu\text{m}$  de côté à la base de la pyramide, ont été obtenus par diffusion de vapeur dans une goutte suspendue équilibrée contre 0,6 M de sulfate d'ammonium et 200 mM d'acide citrique pH 4.

Dans la condition de cristallisation de sulfate d'ammonium-Tris pH 7,5-CoCl<sub>2</sub>-sucrose monolaurate décrite dans l'article, aucun cristal de FBP aldolase de *Taq* sélénométhionée n'a été obtenu. Cette condition nécessite peut-être une protéine très pure, or la protéine modifiée n'a subi qu'une étape de purification, l'étape de chauffage. Faute de cristaux la technique de MAD n'a pu être utilisée. La technique de MIR a été essayée en trempant les cristaux natifs dans l'yttrium. La présence d'yttrium et l'absence de cobalt n'ont pas affecté l'aspect des cristaux. La fluorescence de l'yttrium n'a pas été détectée dans le cristal trempé puisque cet élément ne contient que 39 électrons. Une série de données a quand même été collectée à 1,105 Å toujours à la station X8C au Brookhaven National Laboratory. Le cristal trempé dans l'yttrium avait les mêmes paramètres de maille que le cristal natif (voir l'article page 54) et

diffractait jusqu'à 2 Å avec 92,1% de complétude dans l'intervalle de résolution de 90 Å à 2,0 Å. La comparaison des résultats obtenus pour ce cristal et ceux du cristal natif non-trempé, a démontré un écart de 11,906 pour le chi carré et de 0,091 pour le facteur R. Toutefois ces différences trop mineures et le faible signal de l'yttrium ne permettent pas de résoudre la structure par la technique de MIR. Les légères différences observées seraient plutôt dues à une liaison non-spécifique de l'yttrium, à une légère rotation de la protéine dans le cristal ou au relâchement du cobalt du site actif lors du trempage dans la solution d'yttrium.

## 4.2 Conclusion

Maintenant que la séquence nucléotidique et protéique de la FBP aldolase de *Taq* sont connues, des études de mutagenèse pourront être effectuées pour mieux caractériser les acides aminés impliqués dans la catalyse et dans le maintien de la thermostabilité. La délétion de l'insertion de 21 acides aminés au milieu du site de liaison du DHAP permettrait de vérifier son implication dans la tétramérisation et la coopérativité négative de l'enzyme.

Les données récoltées pour les cristaux de FBP aldolase séléno-méthionée sont présentement analysées selon la technique de MAD par Mme Izard (University of Leicester, G-B) et devraient permettre d'obtenir dans un premier temps, une structure d'une résolution d'environ 2,6 Å. Cette structure pourra ensuite être appliquée aux données obtenues pour les cristaux obtenus dans la condition à pH 7,5 afin d'augmenter la résolution à environ 2 Å.

La détermination de la structure tridimensionnelle de la FBP aldolase de *Taq* permettra d'établir les différences entre les FBP aldolases de classe II A (à laquelle appartient *E.coli*) et celles de classe II B, d'expliquer la préférence pour le cobalt, de déterminer plus spécifiquement le rôle de l'insertion de 21 acides aminés et d'identifier les interactions supplémentaires responsables de la thermostabilité et la tétramérisation.

La cocrystallisation de la FBP aldolase de *Taq* avec son substrat, ses produits ou leurs analogues, rendra possible la compréhension de son mécanisme catalytique. Des inhibiteurs spécifiques aux FBP aldolases de classe II pourront ensuite être synthétisés. Étant donné que les FBP aldolases humaines et animales sont de classe I, et par conséquent que leur mécanisme est totalement différent de l'enzyme retrouvé chez les bactéries et les levures, ces inhibiteurs pourront être utilisés dans les cas d'infections bactériennes ou fongiques. Ce type de traitement est d'autant plus intéressant à l'heure où les antibiotiques sont de moins en moins efficaces contre les bactéries.

Au point de vue industriel, la FBP aldolase de *E.coli* est déjà utilisée en synthèse organique pour catalyser des condensations aldoliques<sup>104</sup>. Son remplacement par la FBP aldolase de *Taq* permettrait d'effectuer les mêmes réactions à plus haute température ou en présence de certains solvants organiques.

---

## Bibliographie

- <sup>1</sup> Horecker, B.L., Tsolas, O. & Lai, C.Y. (1972) Aldolases, dans *The Enzymes*, vol. VII, 3e éd (Boyer, P. D., ed.) pp. 213-258, Academic Press, New York.
- <sup>2</sup> Sygusch, J., Beaudry, D. & Allaire, M. (1987) Molecular architecture of rabbit skeletal muscle aldolase at 2,7-Å resolution, *Proc. Natl. Acad. Sci.* *84*, 7846-7850.
- <sup>3</sup> Blom, N. S., Tétreault, S., Coulombe, R. & Sygusch, J. (1996) Novel active site in *Escherichia coli* fructose-1,6-bisphosphate aldolase, *Nature structural biology* *3*, 856-862.
- <sup>4</sup> Rutter, W.J. (1964) Evolution of aldolase, *Fed. Proc.* *23*, 1248-1257.
- <sup>5</sup> Rutter, W.J. & Lebherz, H.G. (1969) Distribution of Fructose Diphosphate aldolase variants in biological systems, *Biochemistry* *8*, 109-121.
- <sup>6</sup> Lebherz, H.G., Leadbetter, M.M. & Bradshaz, R.A. (1984) Isolation and characterization of the cytosolic and chloroplast forms of spinach leaf fructose diphosphate aldolase, *J. Biol. Chem.* *259*, 1011-1017.
- <sup>7</sup> Marsh, J. J., Wilson, K. J. & Lebherz, H.G. (1989) Structural similarities between spinach chloroplast and cytosolic class I fructose 1,6-bisphosphate aldolases, *Plant physiol.* *91*, 1393-1401.
- <sup>8</sup> Schnarrenberger, C., Jacobshagen, S., Müller, B. & Krüger, I. (1990) dans *Isoenzymes: Structure, function, and use in Biology and medicine* (Ogita, Z.-I. & Markert, C. L., Eds.) 743-764, Wiley-Liss, New York.
- <sup>9</sup> Schnarrenberger, C., Pelzer-Reith, B., Yatsuki, H., Freund, S., Jacobshagen, S. & Hori, K. (1994) Expression and sequence of the only detectable aldolase in *Chlamydomonas reinhardtii*, *Arch. Biochem. Biophys.* *313*, 173-178.
- <sup>10</sup> Chevalier, N., Callens, M., Michels, P. A. M. (1995) High-level expression of *Trypanosoma brucei* fructose-1,6-bisphosphate aldolase in *Escherichia coli* and purification of the enzyme, *Prot. Expr. Purif.* *6*, 39-44.
- <sup>11</sup> Knapp, B., Hundt, E. & Kupper, H. A. (1990) *Plasmodium falciparum* aldolase: gene structure and localization, *Mol. Biochem. Parasitol.* *40*, 1-12.
- <sup>12</sup> Plaumann, M., Pelzer-Reith, B., Martin, W. F. & Schnarrenberger, C. (1997) Multiple recruitment of class-I aldolase to chloroplasts and eubacterial origin of eukaryotic class-II

---

aldolases revealed by cDNAs from *Euglena gracilis*, *Curr. genet.* 31, 430-438.

<sup>13</sup> Russel, G. K. & Gibbs, M. (1967) Partial purification and characterization of two fructose diphosphate aldolases from *Chlamydomonas mundana*, *Biochim. Biophys. Acta* 132, 145-154.

<sup>14</sup> Blom, N. & Sygusch, J. (1997) Product binding and role of the C-terminal region in Class I D-fructose-1,6-bisphosphate aldolase, *Nature structural biology* 4, 36-39.

<sup>15</sup> Morris, A. J. & Tolan, D. R. (1994) Lysine-146 of rabbit muscle aldolase is essential for cleavage and condensation of the C3-C4 bond of fructose 1,6-Bis(phosphate), *Biochemistry* 33, 12291-12297.

<sup>16</sup> Schwartz, E. & Horecker, B. L. (1966) Purification and properties of fructose diphosphate aldolase from *Boa constrictor constrictor*, *Arch. Biochem. Biophys.* 115, 407-416.

<sup>17</sup> Lai, C. Y., Chen, C., Smith, J. D. & Horecker, B. L. (1971) The number, distribution and functional implication of sulfhydryl groups in rabbit muscle aldolase, *Biochem. Biophys. Res. Commun.* 45, 1497-1505.

<sup>18</sup> Bai, N.J., Pai, M.R., Murthy, P.S. & Venkitasuramanian, T.A. (1975) Fructose Diphosphate aldolase from *Mycobacterium smegmatis* : Functional similarities with rabbit muscle aldolase, *Arch. Biochem. Biophys.* 168, 235-239.

<sup>19</sup> Witke, C & Götz, F (1993) Cloning, sequencing, and characterization of the gene encoding the class-I fructose-1,6-bisphosphate aldolase of *Staphylococcus carnosus*, *J. Bacteriol.* 175, 7495-7499.

<sup>20</sup> Lebherz, H.G. & Rutter, W. J. (1973) A class I (Schiff base) fructose diphosphate aldolase of prokaryotic origin : purification and properties of *Micrococcus aerogenes* aldolase, *J. Biol. Chem.* 248, 1650-1659.

<sup>21</sup> Dhar, N. M. & Altekar, W. (1986) Distribution of class I and class II fructose biphosphate aldolases in halophilic archaeobacteria. *FEMS microbio. letters* 35, 177-181.

<sup>22</sup> Stribling, D. & Perham, R. N. (1973) purification and characterization of two fructose diphosphate aldolases from *Escherichia coli* (Crookes' strain), *Biochem. J.* 131, 833-841.

<sup>23</sup> Thomson, G. J., Howlett, G. J., Ashcroft, A. E. & Berry, A. (1998) The dhna gene of *Escherichia coli* encodes a Class I fructose bisphosphate aldolase, *Biochem. J.* 331, 437-445.

<sup>24</sup> Götz, F., Fischer, S. & Schleifer, K.-H. (1980) Purification and characterization of an unusually Heat-stable and acid/base-stable Class I fructose-1,6-bisphosphate aldolase from

---

*Staphylococcus aureus*, *Eur. J. Biochem.* 108, 295-301.

<sup>25</sup> Gross, W., Bayer, M. G., Schnarrenberger, C., Gebhart, U. B., Maier, T. L. & Schenk, H. E. A. (1994) Two distinct aldolases of class II type in the cyanoplasts and in the cytosol of the alga *Cyanophora paradoxa*, *Plant physiol.* 105, 1393-1398.

<sup>26</sup> Richards, O. C. & Rutter, W. J. (1961) Comparative properties of yeast and muscle aldolase, *J. Biol. Chem.* 236, 3185-3192.

<sup>27</sup> Henze, K., Morrison, H. G., Sogin, M. L. & Müller, M. (1998) Sequence and phylogenetic position of a class II aldolase gene in the amitochondriate protist, *Giardia lamblia*, *Gene* 222, 163-168.

<sup>28</sup> Freeze, H. & Brock, T.D. (1970) Thermostable aldolase from *Thermus aquaticus*, *J. Bacteriol.* 101, 541-550.

<sup>29</sup> De Montigny, C. & Sygusch, J. (1996) Functional characterization of an extreme thermophilic class II fructose-1,6-bisphosphate aldolase, *Eur. J. Biochem.* 241, 243-248.

<sup>30</sup> Willard, J. M. & Gibbs, M. (1968) Purification and characterization of the fructose diphosphate aldolases from *Anacystis nidulans* and *Saprospira thermalis*, *Biochim. Biophys. Acta* 151, 438-448.

<sup>31</sup> Cooper, S.J., Leonard, G. A., McSweeney, S. M., Thompson, A. W., Naismith, J. H., Qamar, S., Plater, A., Berry, A. & Hunter, W.N. (1996) The crystal structure of a class II fructose-1,6-bisphosphate aldolase shows a novel binuclear metal-binding site embedded in a familiar fold. *Structure* 4, 1303-1315.

<sup>32</sup> Hall, D. R., Leonard, G. A., Reed, C. D., Watt, C. I., Berry, A & Hunter, W. N. (1999) The crystal structure of *Escherichia coli* class II fructose-1,6-bisphosphate aldolase in complex with phosphoglycolohydroxamate reveals details of mechanism and specificity, *J. Mol. Biol.* 287, 383-394.

<sup>33</sup> Plater, A. R., Zgiby, S. M., Thomson, G. J., Qamar, S. & Wharton, C. W. (1999) Conserved residues in the mechanism of the *E.coli* class II FBP-aldolase, *J. Mol. Biol.* 285, 843-855.

<sup>34</sup> Marsh, J. J. & Lebherz, H.G. (1992) Fructose-bisphosphate aldolases: an evolutionary history. *TIBS* 17, 110-113.

<sup>35</sup> Brock, T. D. (1967) Life at high temperatures, *Science* 158, 1012-10119.

<sup>36</sup> Bergquist, P. L., Love, D. R., Croft, J. E., Streiff, M. B., Danisl, R. M. & Morgan, W. H.

---

(1987) Genetics and potential biotechnological applications of thermophilic and extremely thermophilic micro-organisms, *Biotechnol. Genet. Eng. Rev. volume 5*, 199-244.

<sup>37</sup> Brock, T. D. (1986) Introduction: an overview of the thermophiles, Dans *Thermophiles: General, molecular, and applied microbiology* (Thomas D. Brock edn) pp.1-16, J. Wiley & Sons, New York.

<sup>38</sup> Oshima, T. (1986) The genes and genetics apparatus of extreme thermophiles, Dans *Thermophiles: General, molecular, and applied microbiology* (Thomas D. Brock edn) pp.137-157, J. Wiley & Sons, New York

<sup>39</sup> Langworthy, T. A. & Pond, J. L. (1986) Membranes and lipids of thermophiles, Dans *Thermophiles: General, molecular, and applied microbiology* (Thomas D. Brock edn) pp.107-135, J. Wiley & Sons, New York.

<sup>40</sup> Daniel, R. M. (1986) The stability of proteins from extreme thermophiles, Dans *Protein structure folding and design* (Oxender D. L. ed) pp.291-296, Alan R. Liss, Inc, New York.

<sup>41</sup> Sundaram, T. K. (1986) Physiology and growth of thermophilic bacteria, Dans *Thermophiles: General, molecular, and applied microbiology* (Thomas D. Brock edn) pp.75-106, J. Wiley & Sons, New York.

<sup>42</sup> Coolbear, T., Daniel, R. M. & Morgan, H. W. (1992) The enzymes from extreme thermophiles: Bacterial sources, thermostabilities and industrial relevance, *Adv. Biochem. Engin. Biotech. 45*, 57-98.

<sup>43</sup> Grutter, M. G., Hawkes, R. B. & Matthews, B. W. (1979) Molecular basis of thermostability in the lysozyme from bacteriophage T4, *Nature 277*, 667-669.

<sup>44</sup> Matsumara, M., Kataoka, S. & Aiba, S. (1986) Single amino acid replacements affecting the thermostability of kanamycin nucleotidyltransferase, *Mol. gen. genet. 204*, 355-358.

<sup>45</sup> D'Auria, S., Rossi, M., Nucci, R., Barone, G., Catanzano, F., Del Vecchio, P. & Graziano, G. (1996) Temperature-induced denaturation of  $\beta$ -glycosidase from the archaeon *Sulfolobus solfataricus*, *J. Biochem. 120*, 292-300.

<sup>46</sup> Kanaya, S. & Itaya, M. (1992) Expression, purification, and characterization of a recombinant ribonuclease H from *Thermus thermophilus* HB8, *J. Biol. Chem. 267*, 10184-10192.

<sup>47</sup> Grattinger, M., Dankesreiter, A., Schurig, H. & Jaenicke, R. (1998) Recombinant

---

phosphoglycerate kinase from the hyperthermophilic bacterium *Thermotoga maritima*: catalytic, spectral and thermodynamic properties, *J. Mol. Biol.* 280, 525-533.

<sup>48</sup> Knapp, S., Karshikoff, A., Berndt, K. D., Christova, P., Atanasov, B. & Ladenstein, R. (1996) Thermal unfolding of the DNA-binding protein Sso7d from the hyperthermophile *Sulfolobus solfataricus*, *J. Mol. Biol.* 264, 1132-1144.

<sup>49</sup> McCrary, B. S., Edmondson, S. P. & Shriver, J. W. (1996) Hyperthermophile protein folding thermodynamics: differential scanning calorimetry and chemical denaturation of Sac7d, *J. Mol. Biol.* 264, 784-805.

<sup>50</sup> Beadle, B. M., Baase, W. A., Wilson, D. B., Gilkes, N. R. & Shoichet, B. K. (1999) Comparing the thermodynamic stabilities of a related thermophilic and mesophilic enzyme, *Biochemistry* 38, 2570-2576.

<sup>51</sup> Jaenicke, R., Schuring, H., Beaucamp, N. & Ostendorp, R. (1996) Structure and stability of hyperstable proteins: glycolytic enzymes from hyperthermophilic bacterium *Thermotoga maritima*, *Adv. Protein Chem.* 48, 181-269.

<sup>52</sup> Wallon, G., Yamamoto, K., Kirino, H., Yamagishi, A., Lovett, S. T., Petsko, G. A. & Oshima, T. (1997) Purification, catalytic properties and thermostability of 3-isopropylmalate dehydrogenase from *Escherichia coli*, *Biochim. Biophys. Acta* 1337, 105-112.

<sup>53</sup> Haney, P. J., Badger, J. H., Buldak, G. L., Reich, C. I., Woese, C. R. & Olsen, G. J. (1999) Burley, S. K. & Petsko, G. A. (1985) Aromatic-aromatic interaction: a mechanism of protein structure stabilization *Science* 229, 23-28. thermal adaptation analyzed by comparison of protein sequences from mesophilic and extremely thermophilic *Methanococcus* species, *Proc. Natl. Acad. Sci.* 96, 3578-3583.

<sup>54</sup> Argos, P., Rossmann, M. G., Grau, U. M., Zuber, H., Frank, G. & Tratschin, J. D. (1979) Thermal stability and protein structure, *Biochemistry* 18, 5698-5703.

<sup>55</sup> Menendez-Arias, L. & Argos, P. (1989) Engineering protein thermal stability: sequence statistics point to residue substitutions in  $\alpha$ -helices, *J. Mol. Biol.* 206, 397-406.

<sup>56</sup> Mrabet, N. T., Van den Broeck, A., Van den Brande, I., Stanssens, P., Laroche, Y., Lambeir, A.-M., Matthijssens, G., Jenkins, J., Chiadmi, M., Van Tilbeurgh, H., Rey, F., Janin, J., Quax, W. J., Lasters, I., De Maeyer, M. & Wodak, S. J. (1992) Arginine residues as stabilizing elements in proteins, *Biochemistry* 31, 2239-2253.



- 
- <sup>57</sup> Tuengler, P. & Pfeleiderer, G. (1977) Enhanced heat, alkaline and tryptic stability of acetamidinated pig heart lactate dehydrogenase, *Biochim. Biophys. Acta* 484, 1-8.
- <sup>58</sup> Burley, S. K. & Petsko, G. A. (1985) Aromatic-aromatic interaction: a mechanism of protein structure stabilization, *Science* 229, 23-28.
- <sup>59</sup> Hecht, M. H., Strurtevant, J. M. & Sauer, R. T. (1984) Effect of single amino acid replacements on the thermal stability of the NH<sub>2</sub>-terminal domain of phage lambda repressor, *Proc. Natl. Acad. Sci.* 81, 5685-5689.
- <sup>60</sup> Richardson, J. S. & Richardson, D. C. (1988) Amino acid preferences for specific locations at the ends of  $\alpha$  helices, *Science* 240, 1648-1652.
- <sup>61</sup> Elcock, A. H. (1998) The stability of salt bridges at high temperatures: implications for hyperthermophilic proteins, *J. Mol. Biol.* 284, 489-502.
- <sup>62</sup> Perutz, M. F. & Raidt, H. (1975) Stereochemical basis of heat stability in bacterial ferredoxins and in haemoglobin A<sub>2</sub>, *Nature* 255, 256-259.
- <sup>63</sup> Steelwagen, E. & Wilgus, H. (1978) Relationship of protein thermostability to accessible surface area, *Nature* 275, 342-343.
- <sup>64</sup> Chan, M. K., Mukund, S., Kletzin, A., Adams, M. W. & Rees, D. C. (1995) Structure of a hyperthermophilic tungstopterin enzyme, aldehyde ferredoxin oxidoreductase, *Science* 267, 1463-1469.
- <sup>65</sup> Merkler, D. J., Farrington, G. K. & Wedler, F. C. (1981) Protein thermostability, *Int. J. Pept. Prot. Res.* 18, 430-442.
- <sup>66</sup> Jaenicke, R. (1981) Enzymes under extremes of physical conditions, *Ann. Rev. Biophys. Bioeng.* 10, 1-67.
- <sup>67</sup> Danson, M. J. & Hough, D. W. (1998) Structure, function and stability of enzymes from the Archaea, *Trends microbiol.* 6, 307-314.
- <sup>68</sup> Perutz, M. F. & Raidt, H. (1975) Stereochemical basis of heat stability in bacterial ferredoxins and in haemoglobin A<sub>2</sub>, *Nature* 255, 256-259.
- <sup>69</sup> Vetriani, C., Maeder, D. L., Tolliday, N., Yip, K. S.-P., Stillman, T. J., Britton, K. L., Rice, D. W., Klump, H. H. & Robb, F. T. (1998) Protein thermostability above 100°C: a key role for ionic interactions, *Proc. Natl. Acad. Sci.* 95, 12300-12305.
- <sup>70</sup> Kohlhoff, M., Dahm, A. & Hensel, R. (1996) Tetrameric triosephosphate isomerase from

---

hyperthermophilic archaea, *FEBS Letters* 383, 245-250.

<sup>71</sup> Sterner, R., Kleemann, G. R., Szadkowschi, H., Lustig, A., Hennig, M. & Kirschner, K. (1996) Phosphoribosyl anthranilate isomerase from *Thermotoga maritima* is an extremely stable and active homodimer, *Prot. Sci.* 5, 2000-2008.

<sup>72</sup> Hess, D., Krüger, K., Knappik, A., Palm, P. & Hensel, R. (1995) Dimeric 3-phosphoglycerate kinases from hyperthermophilic archaea: Cloning, sequencing and expression of the 3-phosphoglycerate kinase gene of *Pyrococcus woesei* in *Escherichia coli* and characterization of the protein. Structural and functional comparison with the 3-phosphoglycerate kinase of *Methanothermus fervidus*, *Eur. J. Biochem.* 233, 227-237.

<sup>73</sup> Beernink, P. & Tolan, D. R. (1994) Subunit interface mutants of rabbit muscle aldolase form active dimers, *Prot.Sci.* 3, 1383-1391.

<sup>74</sup> Milstein, M. (1994) Yellowstone managers eye profits from hot microbes *Science* 264, 655.

<sup>75</sup> Cowan, D., Daniel, R. & Morgan, H. (1985) Thermophilic proteases: properties and potential applications, *Trends biotech.* 3, 68-72.

<sup>76</sup> Coolbear, T., Monk, C., Peek, K., Morgan, H. W. & Daniel, R. M. (1992) Laboratory-scale investigations into the use of extremely thermophilic proteinases for cleaning ultrafiltration membranes fouled during whey processing, *J. Memb. Sci.* 67, 93-101.

<sup>77</sup> Van den Burg, B., Vriend, G., Veltman, O. R., Venema, G. & Eijsink, V. G. H. (1998) Engineering an enzyme to resist boiling, *Proc.Natl. Acad.Sci.* 95, 2056-2060.

<sup>78</sup> Woese, C. R. (1987) Bacterial evolution, *Microbiol. Rev.* 51, 221-271.

<sup>79</sup> Shoichet, B. K., Baase, W. A., Kuroki, R. & Matthews, B. W. (1995) A relationship between protein stability and protein function, *Proc.Natl. Acad.Sci.* 92, 452-456.

<sup>80</sup> Varley, P. G. & Pain, R. H. (1991) Relation between stability, dynamics and enzyme activity in 3-phosphoglycerate kinases from yeast and *Thermus thermophilus*, *J. Mol. Biol.* 220, 531-538

<sup>81</sup> Brown, J. R. & Doolittle, W. F. (1997) *Archaea* and the prokaryote-to-eukaryotes transition, *Microbiol. Mol. Biol. Rev.* 61, 456-502.

<sup>82</sup> Brown, J. R. & Lupas, A. N. (1998) What makes a thermophile?, *Trends microbiol.* 6, 349-351.

<sup>83</sup> Galtier, N., Tourasse, N. & Gouy, M. (1999) A nonhyperthermophilic common ancestor to

---

extant life forms, *Science* 283, 220-221.

<sup>84</sup> Vogel, G. (1999) RNA study suggests cool cradle of life, *Science* 283, 155-156.

<sup>85</sup> Pask-Hughes, R. A. & William, R. A. D. (1977) Yellow-pigmented strains of *Thermus* spp. from icelandic hot spring, *J. Gen. Microbiol.* 102, 375-383.

<sup>86</sup> Denman, S., Hampson, K & Patel, B.K.C. (1991) Isolation of strains of *Thermus aquaticus* from the Australian artesian basin and a simple and rapid procedure for the preparation of their plasmids, *FEMS Microbiol. Letters* 82, 73-78.

<sup>87</sup> Degryse, E., Glansdorff, N. & Piérard, A. (1978) A comparative analysis of extreme thermophilic bacteria belonging to the genus *Thermus* *Arch. Microbiol.* 117, 189-196.

<sup>88</sup> Marteinsson, V. T., Birrien, J.-L., Kristjansson, J. K. & Prieur, D. (1995) First isolation of thermophilic aerobic non-sporulating heterotrophic bacteria from deep-sea hydrothermal vents, *FEMS Microbiol. Ecology* 18, 163-174.

<sup>89</sup> Brock, T. D. & Freeze, H. (1969) *Thermus aquaticus* gen. n. and sp. n., a non-sporulating extreme thermophile, *J. Bacteriol.* 98, 289-297.

<sup>90</sup> Ray, P. H., White, D. C. & Brock, T. D. (1971) Effect of growth temperature on the lipid composition of *Thermus aquaticus*, *J. Bacteriol.* 108, 227-235.

<sup>91</sup> Watanabe, K., Oshima, T., Iijima, K., Yamazumi, Z. & Nishimura, S. (1980) Purification and thermal stability of several specific tRNAs from an extreme thermophile, *Thermus thermophilus* HB, *J. Biochem.* 87, 1-13.

<sup>92</sup> <http://www.dna.affrc.go.jp/nakamura/codon.html>

<sup>93</sup> Kenkel, T. & Trela, J. M. (1979) Protein turnover in the extreme thermophile *Thermus aquaticus*, *J. Bacteriol.* 140, 543-546.

<sup>94</sup> Korolev, S., Nayal, M., Barnes, W. M., Di Cera, E. & Waksman, G. (1995) Crystal structure of the large fragment of *Thermus aquaticus* DNA polymerase I at 2.5-Å resolution: structural basis for thermostability, *Proc. Natl. Acad. Sci.* 92, 9264-9268.

<sup>95</sup> 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.

<sup>96</sup> Wilquet, V. & Van de Casteele, M (1999) The role of the codon first letter in the relationship between genomic GC content and protein amino acid composition, *Res.*

---

*Microbiol.* 150, 21-32.

<sup>97</sup> Otwinowski, Z. & Minor, W. (1997) Processing of X-rays diffraction data collected in oscillation mode. Dans *Methods in enzymology*, Vol. 276 : *Macromolecular crystallography, Part A* (C. W. Carter, Jr & R. M. Sweet, eds) pp. 307-326, Academic Press, NY.

<sup>98</sup> Flechner, A., Gross, W., Martin, E. F. & Schnarrenberger (1999) chloroplast class I and class II aldolases are bifunctional for fructose-1,6-biphosphate and sedoheptulose-1,7-biposphate cleavage in the Calvin cycle, *FEBS letters* 447, 200-202.

<sup>99</sup> <http://www.cmpharm.ucsf.edu/~nomi/nnpredict.html>

<sup>100</sup> Kyte, J. & Doolittle, R. F. (1982) A simple method for displaying the hydropathic character of a protein, *J. Mol. Biol.* 157, 105-132.

<sup>101</sup> Van den Bergh, E. R., Baker, S. C., Raggars, R. J., Terpstra, P., Woudstra, E. C., Dijkhuizen, L. & Meijer, W. G. (1996) Primary structure and phylogeny of the Calvin cycle enzymes transketolase and fructosebisphosphate aldolase of *Xanthobacter flavus*, *J Bacteriol.* 178, 888-893.

<sup>102</sup> Conway, T. (1992) The Entner-Doudoroff pathway : history, physiology and molecular biology, *Microb.rev.* 103, 1-28.

<sup>103</sup> Bergfors, T. (1999) The use of dynamic light scattering as a predictive tool for protein crystallisation: how good is it?, Résumé du congrès de l'IUCr 1999, Glasgow, UK.

<sup>104</sup> Lewin, R. A. (1965) Isolation and some physiological features of *Saprospira thermalis*, *Can. J. Microb.* 11, 77-86.

<sup>105</sup> Blom, N. & Sygusch, J. (1997) Product binding and role of the C-terminal region in Class I D-fructose 1,6-bisphosphate aldolase, *Nature structural biology* 4, 36-39.

<sup>106</sup> Von der Osten, C. H., Sinskey, A. J., Barbas III, C. F., Pederson, R. L., Wang, Y.-F. & Wong C.-H. (1989) Use of a recombinant bacterial fructose-1,6-diphosphate aldolase in aldol reactions: preparative syntheses of 1-deoxynojirimycin, 1-deoxymannojirimycin, 1,4-dideoxy-1,4-imino-D-arabinol, and fagomine, *J. Am. Chem. Soc.* 111, 3924-3927.